A·B·A·I
Association for Behavior Analysis International

REVIEW

Check for
updates

Social Validity in Behavioral Research: A Selective Review

Rachelle N. Huntington[1] · Natalie M. Badgett[2] · Nancy E. Rosenberg[3] ·
Kaitlin Greeny[3] · Alice Bravo[4] · Roxanne M. Bristol[5] · Young Hee Byun[6] ·
Madelynn S. Park[6]

Abstract

Through the application of behavioral principles, behavior analysts seek to produce socially meaningful behavior change, defined as alterations in behavior that yield important outcomes immediately beneficial for the direct consumers of interventions and key stakeholders. Behavioral practitioners and researchers often engage in assessment and reporting of the meaningfulness of behavior change using social validity assessments. These assessments ensure that target behaviors are appropriately selected, intervention procedures are acceptable, and satisfactory outcomes are produced. The purpose of this review is to identify the current state of social validity within behavioral literature. We reviewed eight peer-reviewed journals between 2010 and 2020. We found that 47% of the intervention studies reviewed included a social validity assessment. Social validity assessment across journals has increased over time, with a significant rise from 2019 to 2020. Implications of these findings and suggestions for future work are discussed.

Keywords acceptability · satisfaction · social importance · social significance · social validity · subjective

✉ Rachelle N. Huntington
     rachelle.huntington@nau.edu

Extended author information available on the last page of the article

Through the application of behavioral principles (e.g., reinforcement, punishment, extinction), behavior analysts seek to produce socially important behavior change. Baer et al. (1968, 1987) define *socially important* as changes in behavior that yield outcomes that are beneficial for consumers. Schwartz and Baer (1991) categorize consumers as direct (i.e., those participating directly in the intervention) and indirect (i.e., individuals who may have contact with direct consumers) and contend that social importance should extend to both. The early and consistent focus on socially important outcomes in applied behavior analysis (ABA) led to the guiding construct of social validity, defined by Wolf (1978) as the extent to which an intervention's goals, interventions, and outcomes are acceptable to the consumer.

Today, behavioral practitioners and researchers engage in assessment and reporting of social validity to ensure that target behaviors are appropriately selected, to evaluate the acceptability of assessment and interventions, and to assess the meaningfulness of the behavior change according to consumers (i.e., goals, procedures, and outcomes; Schwartz & Baer, 1991). Social validity assessment is an iterative and ongoing process in which future implementation is directly informed by feedback from key stakeholders; consumer feedback can alter interventions and those alterations can subsequently affect consumer feedback (Nicolson et al., 2020; Schwartz & Baer, 1991). Given that the purpose of ABA is to affect meaningful change in important human behavior, the effort to continuously seek feedback from those directly experiencing behavioral strategies is essential. In addition, although assessment of social validity has clear implications for the satisfaction of the specific consumer and for the success of an individual intervention, these data can also be valuable to the field of ABA. Given the recent and increasingly critical feedback from stakeholders in both formal (e.g., publications; Cumming et al., 2020; McGill & Robinson, 2020) and informal (e.g., blogs, position statements, social media; National Council on Independent Living, 2021) outlets, the assessment of social validity and emphasis on consumer experience is paramount.

Social validity has been the subject of several previous reviews within the behavioral literature (Carr et al., 1999; Ferguson et al., 2019; Kennedy, 1992) and closely related fields such as special education (Spear et al., 2013; Snodgrass et al., 2018) and early intervention (Ledford et al., 2016; Park & Blair, 2019). In the behavioral literature, Kennedy (1992) conducted a selective review of social validity in all research studies published in the *Journal of Applied Behavior Analysis* (JABA) from 1968 to 1990 and *Behavior Modification* (BMOD) from 1977 to 1990. Kennedy reported assessment type (i.e., subjective measurement or normative comparison), focus (i.e., goals, interventions, and/or outcomes), and timing of assessment (i.e., pre-, or postintervention). This early review found that 20% of studies included a social validity assessment, and that most assessments were subjective evaluations implemented post-intervention. Carr et al. (1999) also selectively examined social validity trends within JABA, extending the range of studies reviewed to include those published between 1968 and 1998. They examined trends in assessment of two social validity domains (i.e., procedural acceptability and satisfaction with outcomes) as well as whether frequency of reporting was different across intervention settings (i.e., naturalistic or analog).

Results indicated similar levels of social validity assessment as those reported by Kennedy (1992); 13% of studies included outcome and acceptability measures. Ferguson et al. (2019) recently updated the review of social validity assessment in JABA, examining studies between 1999 and 2016. Like Kennedy (1992), they reviewed studies for the inclusion of a social validity measure, assessment type, and focus. They also noted if social validity assessment was recommended as an area for future research. This selective review found that 12% of studies included social validity assessment, with frequency ranging from 3% to 22% across years in the selected period. Subjective assessments were most utilized and social validation of interventions was the most common focus.

Noticeably absent in these reviews is a discussion of the language used to describe social validity assessment. No existing review of behavioral literature has reported specific search terms or phrases used to identify social validity assessments, resulting in a lack of clarity regarding social validity verbiage. In response to this issue, Nicolson et al. (2020) suggested that common terms like *social significance* that are often used in place of *social validity* may inherently have a different meaning. The lack of an agreed upon lexicon for describing social validity may be related to the low level of reporting described in previous selective reviews of the literature. It is possible that low reporting is an artifact related to disagreement among researchers related to which terms relate to social validity. This lack of attention to terminology may have also adversely affected researchers' ability to summarize and evaluate social validity measurement using systematic search methods as evidenced using selective strategies used across existing reviews related to social validity (Carr et al., 1999; Ferguson et al., 2019; Kennedy, 1992).

These existing reviews of behavioral literature have illuminated trends in reporting assessment of social validity, primarily focusing on publications in JABA. Although JABA is considered the "flagship" journal for the field of ABA (Kranak et al., 2020) and its publications are frequently assigned in undergraduate and graduate preparation programs (Ferguson et al., 2019; Frieder et al., 2018; Pastrana et al., 2016), it is not the only outlet for behavior analytic research. In a recent review of publication patterns in JABA, Alligood et al. (2019) found that out of 741 publishing instructors affiliated with graduate level verified ABA course sequences, only 260 published in JABA between 2000 and 2015. The implication of this finding is that 65% of researchers who are instructors in ABA-related coursework are publishing in journals other than JABA.

The purpose of this review is to identify the current state of social validity assessment within behavioral literature. We seek to add to the current understanding of reporting patterns by (1) expanding the scope of our search beyond JABA through the inclusion of additional journals that focus on the empirical evaluation of behavioral interventions to change human behavior; (2) exploring reporting trends, in comparison with JABA, to determine if it is representative of the behavioral literature as a whole in relation to inclusion of social validity assessment; and (3) including a wide range of terms that are synonymous with social validity or that describe common methods of social validity assessment.

This review was guided by the following questions:

1. What is the frequency of reporting of social validity assessment in the behavior analytic literature published between 2010 and 2020?
2. How does frequency of reporting differ from previous reviews?
3. What terms are currently being used to describe social validity assessment across journals?

Method

We conducted a review of behavioral intervention studies published in peer-reviewed journals between 2010 and 2020. A selective search strategy was used to identify trends in the publication of social validity assessment in behavioral intervention research. This strategy was chosen in alignment with previous reviews on this topic (Carr et al., 1999; Ferguson et al., 2019; Kennedy, 1992; Snodgrass et al., 2018) and to further evaluate the terminology used to describe this construct. Reviewed journals were selected using a systematic strategy to expand previous reviews to include additional journals that publish behavior analytic research.

We used two established definitions of social validity to guide our search strategy and to determine whether social validity was assessed in published behavioral intervention research, Wolf (1978) and Cooper et al. (2020). Wolf (1978) defined social validity as measuring the social significance of goals, social appropriateness of the interventions, and/or the social importance of the effects. Cooper et al. (2020) defined social validity as "the extent to which target behaviors are appropriate intervention with a human, procedures are acceptable, and important and significant changes in target and collateral behaviors are produced" (Cooper et al., 2020, p. 800). Both definitions were included to incorporate the foundational definition of the concept of social validity (i.e., Wolf, 1978) and the current definition from a widely recognized educational text (i.e., Cooper et al., 2020).

Journal Identification

We sought to identify journals that published intervention research relevant to ABA and behavioral intervention. We began with a list of 25 peer-reviewed journals that were identified by a university librarian as containing behavior analytic content. We then reviewed the provided list for relevance within our research team and invited experts in ABA to review and comment on the list. Then we cross-referenced the journal list with professional organization journal recommendations (e.g., Association for Behavior Analysis International). Through this process, we added 18 peer-reviewed journals for further review, resulting in a list of 43 peer-reviewed journals that publish research related to behavioral intervention. We then evaluated journal relevance to ABA by searching the mission statements of the 43 identified journals for the following terms: applied behavior analy*, ABA, behavior analy*, behavioral analy*, behavior modification, behavioral modification, and applications of the

experimental analysis of behavior. Journals were included in the search and screening process if (1) their mission statement included at least one of the search terms; (2) they were published primarily in English and in the United States; and (3) they published applied research. A total of eight journals were included in this review: *The Analysis of Verbal Behavior* (AVB), *The Behavior Analyst/Perspectives on Behavior Science* (TBA/POBS), *Behavior Analysis in Practice* (BAP), *The Behavior Analyst Today* (BAT)/*Behavior Analysis: Research and Practice* (BARP), *Behavioral Interventions* (BI), *Behavior Modification* (BMOD), *Journal of Applied Behavior Analysis* (JABA), and *The Psychological Record* (TPR).

Search and Screening

We used a four-step search and screening strategy to identify peer-reviewed studies that included an assessment of social validity: (1) identification of empirical research studies; (2) identification of terms and phrases related to social validity assessment within empirical research studies; (3) identification of intervention studies with human participants; and (4) confirmation that terms and phrases were used in the context of social validity assessment (see Table 1).

In the initial step of our search and screening process, we manually searched every issue from the included journals during the designated period to identify empirical research studies. Studies were considered empirical if they included both a method section and a results/findings section. One exception to the rule included "practice brief reports" published in BAP, which often include a description of results without a dedicated section header; studies in this category that included empirical research data were included. Examples of studies that did not qualify as empirical research included editorials, commentaries, and book reviews. We screened 3,846 study articles, of which 2,538 were identified as empirical with methods and results sections.

In the second step of the search and screening process, members of the research team identified terms and phrases related to social validity assessment within empirical research studies. To do this, coders used the "search" function (i.e., ctrl + F) to determine whether an article included a term or phrase related to social validity in

Table 1 Results of the Search and Screening Process

Journal	Empirical	Related Terms	Intervention	Social Validity
Journal of Applied Behavior Analysis	913	373	303	148
Behavior Modification	331	270	199	80
Behavior Analysis in Practice	271	158	111	65
Behavioral Intervention	291	163	141	89
The Behavior Analyst/Perspectives on Behavior Science	21	8	1	0
The Psychological Record	477	224	74	17
Analysis of Verbal Behavior	125	42	34	8
Behavior Analyst Today/Behavior Analysis: Research and Practice	109	61	36	18

the methods and/or results sections. Terms and phrases were identified by referencing previous reviews focused on social validity (Carr et al., 1999; Ferguson et al., 2019; Kennedy, 1992; Nicolson et al., 2020; Snodgrass et al., 2018) and included social validity, social validation, social significance, social importance, socially valid, socially significant, socially important, treatment validity, treatment acceptability, consumer satisfaction, satisfaction survey(s), Likert, acceptability scale(s), rating scale(s), rating profile(s), rating form(s), acceptability, validity, satisfaction, subjective, questionnaire(s), survey(s), and interview(s). Of the 2,538 empirical studies, 1,299 included a matching social validity term or phrase.

The purpose of the third step of the search and screening process was to identify studies evaluating the effects of a behavioral intervention (i.e., the introduction of an independent variable with the intent to change behavior) on human behavior for which an assessment of social validity would be appropriate. Studies focusing exclusively on behavior assessment, literature reviews, surveys, and studies with animal subjects were excluded. Of the 1,299 studies reviewed, 899 were determined to contain an intervention intended to change human behavior.

In the last step of the search and screening process, we confirmed that the terms and phrases were used to describe assessment of social validity. Members of the research team completed a full text search of each article to locate terms and phrases recorded in step 2 of the search and screening process. We then read the section containing the term or phrase and determined whether the text referred to assessment of social validity as defined by Wolf (1978) and Cooper et al. (2020). If the term was used to describe assessment of social validity, the coder recorded the primary term associated with the social validity assessment as described in the study. For example, if the coder found the term "satisfaction survey" in reference to the social validity assessment and it met the definition(s) of social validity, the term "satisfaction survey" was recorded. Of the 899 studies that measured the effects of a behavioral intervention on human behavior, 425 described assessments of social validity.

Interrater Reliability

During step 1 of the search and screening process, the research team consisted of three undergraduate research assistants, one master's level student studying ABA, and four doctoral-level behavior analysts. In the subsequent steps 2, 3, and 4, the coding team consisted of one undergraduate research assistant, two master's level students studying ABA, two doctoral students studying ABA, and four doctoral-level behavior analysts. All members of the research team participated in coding after being thoroughly trained. During training, team members reviewed studies published prior to 2010 (thus, outside the scope of the review) from each included journal and were required to meet a criterion of 85% or higher interrater reliability (IRR) with the first or second authors. During coding, IRR was calculated weekly. If a coder were to fall below 85% accuracy, they would be retrained before they continued with coding (this did not occur). Issues and studies selected for IRR were chosen at random but the ratio per journal was carefully monitored to ensure that accuracy checks were proportionally equivalent.

For each item selected for IRR, two different coders reviewed independently and coded for inclusion using the criteria described above. In step 1, we recorded agreement on the number of studies with a method and results section within each issue. In step 2, we recorded agreement of each paper that included a term or phrase. For step 3, we recorded agreement on whether the paper contained an intervention intended to change human behavior and finally, in step 4, we recorded agreement on whether coders agreed that the term and/or phrase met the definition(s) of social validity. In addition, we recorded agreement with the term listed to describe the social validity assessment. IRR was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. Any disagreements were reviewed and discussed by the first and second authors who came to a consensus. Steps 1 and 2 were calculated together with 30.5% ($n =$ 93) of issues reviewed for IRR with an agreement score of 92%. Steps 3 and 4 were calculated together with 30.7% ($n = 399$) of articles reviewed for IRR with an agreement score of 93%.

Results

Frequency and Trend of Reporting

We found social validity assessments in seven of the eight journals included in this review, with an overall average of 47% of studies containing a social validity measure. BI had the highest percentage of studies, with 63.10% ($n = 89$) including a social validity measure. In BAP, 58.60% of studies included a social validity measure ($n = 65$), BAT/BARP had 50% ($n = 18$), JABA had 48.80% ($n = 148$), BMOD had 40.20% ($n = 80$), the AVB had 23.50% ($n = 8$) and TPR had 23.00% ($n = 17$; see Figure 1). No studies in TBA/POBS were identified that included social validity assessment.

The inclusion of social validity assessment varied by year both within and across journals (see Figure 2). From 2010 to 2016, the inclusion rate showed a slight positive trend with a marked increase in 2017, and there was a sharp increasing trend from 2018 to 2020 with the highest rates observed in 2019 ($n = 72$) and 2020 ($n = 76$).

Terminology Used to Discuss Social Validity

The most frequently used term across all included studies was "social validity" ($n = 318$; 74.82%), including variations of the phrase (i.e., "social validation," "social validation assessment," "social validity assessment," "social validity measure," "social validity questionnaire," "social validity survey," "socially valid"; see Table 2). The term "acceptability," including related terms (i.e., "acceptability measure," "acceptability questionnaire," "acceptability survey," "treatment acceptability," "social acceptability," "intervention acceptability") was used to describe assessment of social validity in 54 studies (12.70%). "Satisfaction," or a variation on the term (i.e., "satisfaction questionnaire," "satisfaction rating," "consumer
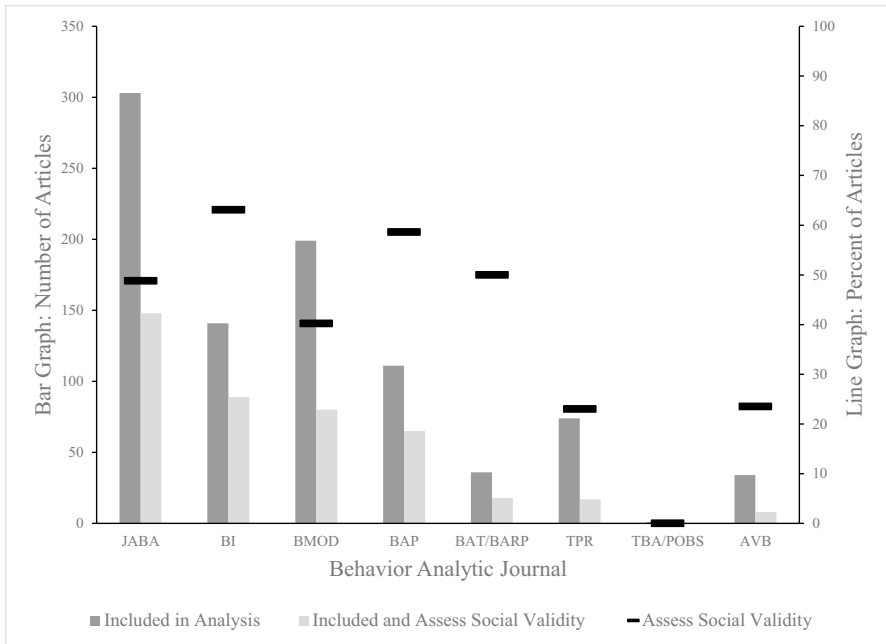
Fig. 1 Articles Included in Review by Journal. *Note. Behavioral Interventions* included the highest percent of studies that assessed social validity. 48.8% of studies (*n* = 303) published in *The Journal of Applied Behavior Analysis* assessed social validity, more than the percent of articles when all other journals are combined (43.1%). JABA = *Journal of Applied Behavior Analysis*; BMOD = *Behavior Modification*; BAP = *Behavior Analysis in Practice*; TBA = *The Behavior Analyst*; POBS = *Perspectives on Behavior Science*; TPR = *The Psychological Record*; AVB = *Analysis of Verbal Behavior*; BAT= *The Behavior Analyst Today*; BARP = *Behavior Analysis: Research and Practice*; BI = *Behavioral Interventions*

satisfaction," and "satisfaction survey") was used in relation to social validity assessment in 28 studies (6.58%). Terms and phrases that appeared at low frequency included "questionnaire(s)" (*n* = 8; 1.88%), "Likert" (*n* = 6; 1.41%), "rating(s)" (*n* = 3; 0.70% includes "rating profile," "rating scale"), "validity" (*n* = 2; 0.47%), "subjective reports" (*n* = 1; 0.23%), "survey(s)" (*n* = 1; 0.23%), "System Usability Scale" (*n* = 1; 0.23%; Brooke, 1996), and "self-reported preference" (*n* = 1; 0.23%).

Within JABA, the most used term was "social validity" (*n* = 123), followed by "acceptability" (*n* = 18), "satisfaction" (*n* = 3), "questionnaire(s)" (*n* = 1), and "Likert" (*n* = 1). Likewise, "social validity" was used more often than any other term combined in all journals except BMOD and TPR. Although "social validity" was the most used term in BMOD (*n* = 31; 38.75%), 49 studies in this journal used a different term or phrase to describe assessment of social validity (61.25%). Likewise, TPR frequently used the term "social validity" (*n* = 7; 41.17%), but a different term was used to describe assessment of social validity in most studies in this journal (*n* = 10; 58.82%).

"Terms analysis" was primarily focused on terms that appeared at least four times across the review. However, it should be noted that terms occurring fewer than four
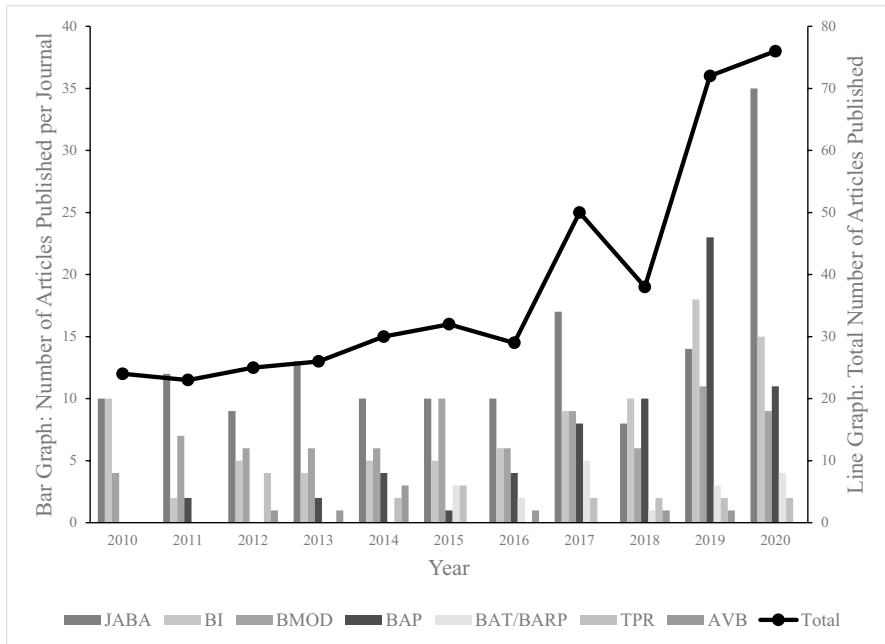
Fig. 2 Articles that Discuss Social Validity Assessment by Year. *Note*. 425 studies (47% of sample, *n* = 899) were found to have assessed social validity. Our findings indicate that an increasing number of published studies include an assessment of social validity, with a large increase in this number occurring in 2019 and sustaining into 2020. JABA = *Journal of Applied Behavior Analysis*; BMOD = *Behavior Modification*; BAP = *Behavior Analysis in Practice*; TBA = *The Behavior Analyst*; POBS = *Perspectives on Behavior Science*; TPR = *The Psychological Record*; AVB = *Analysis of Verbal Behavior*; BAT = *The Behavior Analyst Today*; BARP = *Behavior Analysis: Research and Practice*; BI = *Behavioral Interventions*.

times in the review, (e.g., "subjective report") were also used to describe assessment of social validity in the intervention literature. Across behavioral journals included in this review, low frequency terms appeared most often in studies published in BMOD (*n* = 3), followed by JABA (*n* = 2), BI, (*n* = 2), TPR (*n* = 2), and BAP (*n* = 1).

Discussion

In this review, we explored social validity assessment and identified trends in reporting across eight behavioral journals. We found that 47% of the intervention studies reviewed included a social validity assessment. This percentage represents a definitive increase from previous reviews, which primarily explored social validity in JABA (Carr et al., 1999; Ferguson et al., 2019; Kennedy, 1992) and BMOD (Kennedy, 1992). Social validity assessment across journals has increased over time, with a significant rise from 2019 to 2020. Our findings indicate a range of social validity assessment reporting across behavior analytic journals, with two

Table 2 Terms Related to Social Validity by Journal

Journal	Social Validity	Acceptability	Satisfaction	Questionnaire	Likert	Miscella-neous*
Journal of Applied Behavior Analysis	123	18	3	1	1	2
Behavioral Interventions	78	5	3	1	0	2
Behavior Modification	31	21	15	3	5	5
Behavior Analysis in Practice	57	7	0	0	0	1
Behavior Analyst Today / Behavior Analysis: Research and Practice	15	0	2	1	0	0
The Psychological Record	7	2	4	2	0	2
Analysis of Verbal Behavior	7	0	1	0	0	0
Total	318	53	28	8	6	12

Terms labeled "acceptability" include acceptability, treatment acceptability, and social acceptability. Terms labeled "satisfaction" include satisfaction, consumer satisfaction, and treatment satisfaction. Terms listed as "Miscellaneous" occurred fewer than four times across all journals and include rating/ rating profile/rating scale ($n = 3$), client satisfaction questionnaire ($n = 2$), validity ($n = 2$), intervention acceptability ($n = 1$), subjective reports ($n = 1$), survey, ($n = 1$), system usability scale ($n = 1$) and self-reported preference ($n = 1$)

journals (BI, 63.10%; BAP, 58.60%) publishing a higher percent of studies than JABA (48.80%) and other journals publishing a significantly lower percentage (e.g., TPR, 23.00%; AVB, 23.50%). In addition, we found that the language regarding assessment of social validity has been inconsistent in the behavior analytic literature. Although "social validity" was the most used phrase, its use varied across journals and phrases and/or terms other than social validity were used more frequently within some journals (BMOD, TPR).

This review expands our knowledge of social validity assessment in several ways. First, we added to the current knowledge base regarding trends over time by extending the most recent review of social validity (Ferguson et al., 2019) by 4 years to provide a comprehensive summary of reporting trends. During the most recent years inclusion of social validity assessments has increased dramatically. For example, of the 148 studies that included social validity assessment in JABA, 50% were published between 2017 and 2020. BI, BMOD, and BAP also had notably higher social validity rates in the years 2018–2020 compared to the period 2010–2017.

This review also expands our knowledge of the prevalence of social validity assessment across eight journals dedicated to publishing behavior analytic literature. Although still selective in scope and thereby missing behavior analytic research published in journals not exclusive to ABA, this review provides a significantly wider view of the frequency with which social validity is assessed in behavioral research. Given the wide range of social validity reporting across journals found

in this review (23.00%–63.10%), and that JABA accounted for only 33.74% of the studies reviewed, we suggest that previous reviews may not be representative of the behavior analytic literature. By expanding the scope of this review to a broader set of behavioral literature, we have identified limitations that may exist in generalizing social validity reviews focused on JABA. Our results support that an expanded review, as suggested by Ferguson et al. (2019), was warranted and provides valuable insight regarding the state of social validity assessment in ABA research.

Finally, this review contributes unique information regarding the prevalence of terms used to describe social validity assessment across selected journals. Although "social validity" was identified as the most used phrase across all journals, other terms (e.g., "acceptability" and "satisfaction") were frequently used to describe social validity assessment (approximately 25% of the time). The search terms identified in this review may support future researchers and practitioners in identifying social validity assessment in journal articles. Our findings, however, highlight the fact that discussion of social validity assessment across the behavior analytic literature is at best varied and, at worst, inconsistent. Inconsistent and imprecise use of language to describe methods of social validity assessment contradicts the behavioral commitment to technological descriptions of methods (Baer et al., 1968) and can create barriers in general reporting, review, and use of social validity data in both research and practice. By increasing consistency in language referring to social validity assessment, researchers may promote understanding of social validity and encourage implementation of assessment by consumers of behavioral research (i.e., practitioners).

To achieve more consistency in the behavioral literature, we suggest that future researchers use the term "social validity" consistently to describe these assessments. By systematically labeling and referencing social validity as such, we can further the efforts to define social validity, identify quality indicators for assessing it, and provide avenues for future review. In addition, because the body of behavioral literature is currently inconsistent in its use of terminology, it is difficult to systematically analyze social validity assessment. In other words, without consistent terminology, it is difficult to identify and correct major methodological issues such as addressing some but not all aspects of the social validity construct. Journal editors can support this cause by requiring that social validity data be termed as such. Future researchers may also consider consistent language for each aspect of the social validity construct they are measuring (i.e., goals, procedures, outcomes) if doing so in isolation.

Although this review identifies important trends in reporting of social validity assessment in the selected behavioral literature, it does not provide a detailed analysis of how researchers are implementing social validity assessment or with whom. Future research should investigate the types, interventions, settings, participants, and timing of social validity assessments across the behavior analytic literature to better understand the state of social validity assessment and how assessment data are being used to inform implementation (Schwartz & Baer, 1991). Future researchers should also explore how behavioral practitioners use social validity assessment in their practice. Understanding how, when, how often, and in what ways practitioners utilize social validity assessment is critical information for future researchers looking to improve upon these practices. Finally, the behavior analytic community could

further define the types of studies that should include a social validity assessment. In this review for example, we determined that the introduction of an independent variable with the intention to change human behavior constituted the need for a social validity assessment. Others may argue that social validity assessment is not appropriate for human behavior studies considered more basic or translational in nature. A standard for inclusion of social validity assessments would support their increased use and further promote the behavior analytic commitment to change behavior in a meaningful and socially important manner.

There are several limitations of the data presented in this review. First, this is a selective review and does not represent a fully comprehensive view of the behavior analytic literature. Further, the review includes only journals that publish primarily in English and in the United States. Although this selection of journals inherently limits the perspective of this review, it is a significant extension of previous attempts to review and summarize implementation of social validity assessment in behavioral intervention research. Consistent reporting would facilitate systematic reviews of the literature focused on social validity (Moher et al., 2015).

Our results are also limited by the terms and phrases that were selected to identify social validity assessment within studies, as well as by the method that we used to search for the terms. Although we created an extensive list and completed a manual search to ensure we could identify the terms and phrases, it is possible that we missed terms that describe social validity assessment. Again, a commitment to using consistent terms in social validity assessment among behavioral researchers and practitioners would address this issue.

In addition, our review required the terms or phrases used to describe social validity assessment to be in the methods or results sections of the study. It is possible that studies where social validity assessment was described or reported in other sections of the article (e.g., discussion) were missed and not included within the scope of this review. Social validity assessment is an essential element of behavior analytic work that provides insight for effectiveness (i.e., socially meaningful change; Baer et al., 1968) of interventions. The assertion from Schwartz and Baer (1991) that social validity assessment should be a minimum requirement of applied behavioral research remains valid and should be considered. We encourage researchers to consider social validity as a vital component of their method and data collection processes and include technological descriptions in the methods section of their studies.

Social validity assessment in ABA is critical to ensuring that behavior change is meaningful, important, and worthwhile for those experiencing it (Baer et al., 1968; Horner et al., 2005; Wolf, 1978). Although this review demonstrates an increase in reporting in recent years, social validity is still being assessed in fewer than 50% of publications across the journals we reviewed. This finding is concerning, particularly in light of the influx of negative public feedback (i.e., social validity data) that ABA has received in relation to the appropriateness of intervention goals (e.g., claims that goals target conformity for autistic consumers instead of autonomy; Ne'eman, 2010), acceptability of assessment and interventions (e.g., reports that behavior change interventions are abusive; Sandoval-Norton et al., 2021), and dissatisfaction with intervention outcomes (e.g., assertions that ABA is ineffective and/or has long-term adverse impacts on mental health; Sandoval-Norton et al., 2021).

We hope this review will serve as encouragement for behavior analysts to continue the upward trend and regularly incorporate social validity assessment into their research and practice.

Compliance with Ethical Standards

Conflicts of Interests The authors have no conflicts of interest to report

References

Alligood, C., Anderson, C., & McGee, H. (2019). Casting a wider net: An analysis of scholarly contributions of behavior analysis graduate program faculty. *Behavior Analysis in Practice, 12*(2), 466–472. https://doi.org/10.1007/s40617-018-00281-x

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91–97. https://doi.org/10.1901/jaba.1968.1-91

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 20*(4), 313–327. https://doi.org/10.1901/jaba.1987.20-313

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds.), *Usability evaluation in industry*. Taylor & Francis.

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. https://doi.org/10.1002/(SICI)1099-078X(199910/12)14:4<223::AID-BIN37>3.0.CO;2-Y

Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis* (3rd ed.). Pearson Education.

Cumming, T., Strnadová, I., Danker, J., & Basckin, C. (2020). "I was taught that my being was inherently wrong:" Is applied behavioural analysis a socially valid practice? *International Journal of Arts Humanities & Social Sciences Studies, 5*(12), 72.

Ferguson, J. L., Cihon, J. H., Leaf, J. B., Van Meter, S. M., McEachin, J., & Leaf, R. (2019). Assessment of social validity trends in the journal of applied behavior analysis. *European Journal of Behavior Analysis, 20*(1), 146–157. https://doi.org/10.1080/15021149.2018.1534771

Frieder, J. E., Zayac, R. M., Ratkos, T., Donahue, N., Ware, M., & Paulk, A. (2018). Essential readings for undergraduate students in behavior analysis: A survey of behavior analytic faculty and practitioners. *Behavior Analysis in Practice, 11*(4), 327–333. https://doi.org/10.1007/s40617-018-0260-x

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*(2), 165–179. https://doi.org/10.1177/001440290507100203

Kennedy, C. H. (1992). Trends in the measurement of social validity. *The Behavior Analyst, 15*(2), 147–156. https://doi.org/10.1007/BF03392597

Kranak, M. P., Falligant, J. M., Bradtke, P., Hausman, N. L., & Rooker, G. W. (2020). Authorship trends in the *Journal of Applied Behavior Analysis*: An update. *Journal of Applied Behavior Analysis, 53*(4), 2376–2384. https://doi.org/10.1002/jaba.726

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*(4), 223–233.

McGill, O., & Robinson, A. (2020). "Recalling hidden harms:" Autistic experiences of childhood applied behavioural analysis (ABA). *Advances in Autism, 7*(4), 269–282. https://doi.org/10.1108/AIA-04-2020-0025

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group. (2015). Preferred reporting items for systematic review and

meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Review, 4*(1), 1–9. https://doi.org/10.1186/2046-4053-4-1

National Council on Independent Living. (2021). *Resolution opposing applied behavioral analysis (ABA). The Advocacy Monitor.* https://advocacymonitor.com/ncil-resolution/resolution-opposing-applied-behavioral-analysis-aba/. Accessed 16 Jul 2021

Ne'eman, A. (2010). The future (and the past) of autism advocacy, or why the ASA's magazine, *The Advocate*, wouldn't publish this piece. *Disability Studies Quarterly, 30*(1). https://dsq-sds.org/index.php/dsq/article/view/1059/1244

Nicolson, A. C., Lazo-Pearson, J. F., & Shandy, J. (2020). ABA finding its heart during a pandemic: An exploration in social validity. *Behavior Analysis in Practice, 13*(4), 757–766. https://doi.org/10.1007/s40617-020-00517-9

Park, E. Y., & Blair, K. S. C. (2019). Social validity assessment in behavior interventions for young children: A systematic review. *Topics in Early Childhood Special Education, 39*(3), 156–169.

Pastrana, S. J., Frewing, T. M., Grow, L. L., Nosik, M. R., Turner, M., & Carr, J. E. (2016). Frequently assigned readings in behavior analysis graduate training programs. *Behavior Analysis in Practice, 11*(3), 267–273. https://doi.org/10.1007/s40617-016-0137-9

Sandoval-Norton, A. H., Shkedy, G., & Shkedy, D. (2021). Long-term ABA therapy is abusive: A response to Gorycki, Ruppel, and Zane. *Advances in Neurodevelopmental Disorders, 5*(2), 126–134. https://doi.org/10.1007/s41252-021-00201-1

Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis, 24*(2), 189–204. https://doi.org/10.1901/jaba.1991.24-189

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 & Special Education, 34*(6), 357–370.

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*(2), 203–214. https://doi.org/10.1901/jaba.1978.11-203

Authors and Affiliations

Rachelle N. Huntington[1] ⓘ **· Natalie M. Badgett[2] · Nancy E. Rosenberg[3] · Kaitlin Greeny[3] · Alice Bravo[4] · Roxanne M. Bristol[5] · Young Hee Byun[6] · Madelynn S. Park[6]**

Natalie M. Badgett
Natalie.badgett@unf.edu

Nancy E. Rosenberg
nancyr@uw.edu

Kaitlin Greeny
kloesk@uw.edu

Alice Bravo
bravoa@spu.edu

Roxanne M. Bristol
bristolr@hawaii.edu

Young Hee Byun
yb3ts@virginia.edu

Madelynn S. Park
mp4mw@virginia.edu

[1]   College of Education, Northern Arizona University, Eastburn Education 207c, 639 Riordan Ranch Street, Flagstaff, AZ 86001, USA

[2]   College of Education and Human Services, University of North Florida, Jacksonville, FL, USA

[3]   College of Education, University of Washington, Seattle, WA, USA

[4]   School of Education, Seattle Pacific University, Seattle, WA, USA

[5]   College of Education, University of Hawaii at Manoa, Honolulu, HI, USA

[6]   School of Education and Human Development, University of Virginia, Charlottesville, VA, USA