Regular Article

Hammill Institute on Disabilities

Assessment for Effective Intervention
2019, Vol. 44(2) 123–134

DOI: 10.1177/1534508417749873
aei.sagepub.com
SAGE

Examining the Concurrent Criterion-Related Validity of Direct Behavior Rating–Single Item Scales With Students With Social Competence Deficits

Stephen P. Kilgus, PhD[1], T. Chris Riley-Tillman, PhD[1], Janine P. Stichter, PhD[1], Alexander M. Schoemann, PhD[2], and Sarah Owens, PhD[1]

Abstract

A line of research has supported the development and validation of *Direct Behavior Rating–Single Item Scales* (DBR-SIS) for use in progress monitoring. Yet, this research was largely conducted within the general education setting with typically developing children. It is unknown whether the tool may be defensibly used with students exhibiting more substantial concerns, including students with social competence difficulties. The purpose of this investigation was to examine the concurrent validity of DBR-SIS in a middle school sample of students exhibiting substantial social competence concerns (n = 58). Students were assessed using both DBR-SIS and systematic direct observation (SDO) across three target behaviors. Each student was enrolled in one of two interventions: the Social Competence Intervention or a business-as-usual control condition. Students were assessed across three time points, including baseline, mid-intervention, and postintervention. A review of across-time correlations indicated small to moderate correlations between DBR-SIS and SDO data (r = .25–.45). Results further suggested that the relationships between DBR-SIS and SDO targets were small to large at baseline. Correlations attenuated over time, though differences across time points were not statistically significant. This was with the exception of academic engagement correlations, which remained moderate–high across all time points.

Keywords

autism spectrum disorder, progress monitoring

Over the last several decades and as the American education system has embraced an evidence-based practice (EBP) model, there has been a logical explosion of intervention research yielding numerous evidence-based interventions (EBIs; Burns, Riley-Tillman, & VanDerHeyden, 2012). Regardless of the conceptualization of the school-wide problem-solving model (e.g., positive behavioral interventions and supports), each model is founded upon three core components: (a) EBI, (b) evidence-based assessment (EBA), and (c) evidence-based analysis. Although intervention receives the most attention within the literature, problem-solving models all require defensible assessment data sources to inform intervention decisions. For instance, educators require progress-monitoring data to gauge the effectiveness of any intervention. It should be well understood at this point that EBIs, even those that have been selected and implemented with integrity, might be ineffective for a target student. The EBI might not be matched to a student's needs (e.g., not aligned with the function of the student's problem behavior) or it might not be of sufficient intensity (Kilgus, Collier-Meek, Johnson, & Jaffery, 2014). Given the potential variability in effectiveness across students, it is essential to collect psychometrically defensible (i.e., reliable and valid) data for any intervention case and use that data to gauge the appropriateness and effectiveness of an intervention. As a result, educational professionals require progress-monitoring tools that are technically adequate, efficient, and useful to guide the problem-solving process (Fuchs & Fuchs, 2006).

Although the necessity of progress-monitoring data is embedded in current concepts of best practice, it does create a dilemma when examining the landscape of available tools. Although there are a variety of defensible progress-monitoring options available for academics (many representing curriculum-based measurement tools), there are few options for

[1]University of Missouri, Columbia, USA
[2]East Carolina University, Greenville, North Carolina, USA

Corresponding Author:
Stephen P. Kilgus, University of Missouri, 16 Hill Hall, Columbia, MO 65203, USA.
Email: kilguss@missouri.edu

social behavior. As an illustration of this imbalance, the National Center on Intensive Intervention (2017) has "tools charts" for both academic and behavior progress monitoring, which summarize expert panel ratings of methods submitted for review. As of March 2017, the academic progress-monitoring tools chart includes well over 100 different methods, whereas the behavior tool chart has 17. Furthermore, of the available measures, only three were rated as possessing convincing reliability and validity evidence. This dearth of acceptable and quality methods supports the need for research to identify progress-monitoring tools that are reliable, sensitive, and feasible for use in a problem-solving model (Hintze, 2005; Hintze & Matthews, 2004).

Within the last decade, a great deal of empirical attention has been focused on a particular method of social behavior assessment, termed *Direct Behavior Rating–Single Item Scales* (DBR-SIS; Chafouleas, 2011; Miller, Riley-Tillman, Chafouleas, & Schardt, 2017). DBR-SIS is a hybrid method that combines aspects of behavior-rating scales and systematic direct observation (SDO). Similar to behavior-rating scales, a teacher or other educational professional rates a target student on an operationally defined behavior using an 11-point unipolar graphic-rating scale (0%–100%). Like SDO, DBR-SIS ratings are made in close temporal proximity to the observed behavior. For instance, teachers can complete ratings almost immediately following their observation or within the next several hours. Furthermore, like SDO, DBR-SIS is intended to be used repeatedly over time to establish a stream of reliable data.

The DBR-SIS method was designed with two primary goals in mind. First, DBR-SIS was intended to represent specific instrumentation and associated procedures for which psychometric procedures could be established. In this vein, there is a substantial research base (with more than 30 peer-reviewed publications to date) that establishes DBR-SIS as a defensible, flexible, efficient, and repeatable (i.e., capable of being administered on a frequent basis) method of classroom behavior assessment for use within the general education setting. DBR-SIS includes three core target behaviors: academic engagement (AE), disruptive behavior (DB), and respectful behavior (RB). DBR-SIS has been found to be consistent with SDO, suggesting the former method's criterion-related validity (Chafouleas, McDougal, Riley-Tillman, Panahon, & Hilt, 2005; Riley-Tillman, Chafouleas, Sassu, Chanese, & Glazer, 2008). For instance, Chafouleas, Sanetti, Kilgus, and Maggin (2012) found correlations (r) between DBR-SIS and SDO to be equal to .34 and .29 for AE and DB, respectively. In addition, DBR-SIS and SDO have been found to present a similar formative presentation of target behaviors (e.g., time series single-case line graph; Chafouleas, Riley-Tillman, Sassu, LaFrance, & Patwa, 2007), and result in similar intervention decisions when considered by school psychologists (Riley-Tillman, Chafouleas, Briesch, & Eckert, 2008). The second goal for DBR-SIS was to provide a method that was feasible for classroom use. In relation to this goal, DBR-SIS has been found to be a highly acceptable method of social behavior assessment for teachers and school psychologists (Chafouleas, Riley-Tillman, & Sassu, 2006; Riley-Tillman et al., 2008). The research base for DBR-SIS has resulted in its inclusion within the National Center for Intensive Intervention's tools chart, with ratings of convincing evidence for reliability, validity, and sensitivity to change.

Social Behavior Assessment for Students With Disabilities

Although the global need for behavior assessment methods is critical, the situation is even more concerning in regard to students with disabilities. This is in part due to both (a) the severity of these students' concerns and thus the need for accurate monitoring of treatment response, and (b) the increase in students receiving special education services. Although a wide range of students receive such services, the recent increase in students identified as eligible for special education as a result of needs related to social competence deficits has been particularly unprecedented. For instance, the number of students identified and served as eligible for services related to autism spectrum disorders (ASDs) in the United States has increased from 5,413 to 370,011 cases between the 1991–1992 and 2010–2011 academic school years (Volker, 2012). Such increases in special education enrollment have resulted in great demand for targeted curricula and practices, as well as assessment methods for evaluating response to intervention (Odom, Cox, & Brock, 2013).

The aforementioned concept of *social competence* is defined as the impact of social behavior rather than the behavior itself (Gresham, 1981). That is, students with deficits in social competence may be capable of displaying certain social behaviors, but do so with less skill and fluency, thus decreasing their probability of completing social tasks. Although particularly characteristic of students with ASD, such impairments in social competence are characteristic of several conditions, including emotional disturbance, learning disabilities, and attention-deficit/hyperactivity disorder (Gresham, Sugai, & Horner, 2001). As such, general and special educators, administrators, parents, and support service personnel agree that assessment and intervention for social competence must be a focus if students are expected to achieve increased success and independence (Brown, Odom, & Conroy, 2001; Rogers, 2000).

Although the importance of social competence has been widely agreed upon, the development of this concept into both EBI and EBA has had varied attention. Although there has been progress in the realm of EBI, there are no fully vetted EBA methods with utility in progress monitoring for

students with substantial social competence difficulties. Organizations such as What Works Clearinghouse (WWC) and the National Professional Development Center (NPDC) on ASD have recently identified EBIs for students with specific disabilities, such as ASD (Cook & Cook, 2013). This effort outlines scientifically based interventions that practitioners can implement to address the unique needs of specific populations (e.g., the deficits evidenced by students with ASD). One example of an intervention with a growing evidence base is the Social Competence Intervention (SCI). SCI is a suite of curricula designed to address social competence deficits consistent with students identified with high-functioning ASD or related social needs (Schultz, Schmidt, & Stichter, 2011; Stichter, O'Connor, Herzog, Lierheimer, & McGhee, 2012). The SCI curriculum has multiple age-appropriate programs for elementary (SCI-E), adolescent (SCI-A), and high school (SCI-H) students. The SCI continuum of curricula has been tested for more than 8 years and has received numerous Institute of Education Sciences (IES) and National Institutes of Health (NIH) funding awards with initially promising results (Stichter, Herzog, Owens, & Malugen, 2016; Stichter et al., 2010; Stichter et al., 2012; Stichter, Herzog, Kilgus, & Schoemann, 2018). Although the research on SCI and other emerging EBI for children who have deficits in social competence is promising, it is critical for such EBI to also be paired with feasible EBA options for progress monitoring of both the targeted and generalized effects of intervention.

The most likely candidate for utility in progress monitoring is arguably SDO, which possesses a substantial psychometric base (Chafouleas, Riley-Tillman, & Sugai, 2007) and has been historically used as the core progress-monitoring method for children with social competence concerns. However, in judging the feasibility of SDO for monitoring student performance, it is important to differentiate between higher and lower functioning students. Low-functioning students tend to be instructed in alternative placements with high student-to-teacher ratios and the presence of multiple-student support staff (e.g., special educators, psychologists, behavior analysts). In such a setting, it is feasible to use SDO to track intervention effectiveness, as staff are available to conduct such an assessment. In contrast, high-functioning students tend to be integrated into regular education classes and instructed in group settings with fewer staff, making the collection of SDO data less feasible. In this case, DBR-SIS may present a necessary alternate option.

Utility of DBR for Measuring Social Competence

In 2013, the current team began to explore the use of DBR-SIS in monitoring the progress of students with social competence deficits in response to the SCI. This occurred within the context of a broader randomized controlled trial examining SCI efficacy. It was hypothesized that DBR-SIS would be an ideal candidate to monitor SCI in applied settings given (a) the measure's high level of utility and (b) its relevance to social behavior within the general education setting. Building on the DBR-SIS research base, the "Big 3" scales (i.e., AE, DB, and respectfulness) were considered and adopted as an initial set of target behaviors. Although these behaviors are not specifically focused on social competence, it is likely that increased social competence will result in higher levels of AE and respectfulness, as well as reduced DB. Such logic is consistent with existing theory. For instance, Walker, Irvin, Noell, and Singer (1992) posited that correlates of child social-behavioral competence within the school setting should include (among others) behaviors related to respect (e.g., rule following, compliance, responsiveness to corrections), engagement (e.g., listening carefully to teachers, production of acceptable quality work), and disruption (e.g., defiance, disturbing others, tantrum). This logic is also consistent with prior research, which has suggested social competence is related to both externalizing problems (e.g., DB; Burt & Roisman, 2010) and AE (Wentzel & Watkins, 2002).

The manner in which the Big 3 behaviors are related to social competence could be similar to how oral reading fluency (ORF) is related to general reading proficiency. ORF is by no means synonymous with overall reading ability. However, research suggests ORF is highly predictive of reading proficiency, with changes in the former predicting changes in the latter (Deno, 2003). Accordingly, researchers have commonly referred to ORF as a *general outcome measure* of reading ability (Shinn, 1989). Although ORF is not always the target of reading interventions, it is nevertheless used as a target for progress monitoring given its capacity to predict reading ability. In recognition of the ORF literature, it is proposed that the Big 3 behaviors might serve as similar predictors of student social competence. Although not the target of interventions like SCI, change in the Big 3 behaviors may predict changes in the social skills targeted by such interventions. This article specifically considers this line of reasoning that the core DBR-SIS targets can be used to progress monitor children with substantial social competence deficits who are receiving a SCI in their regular education classes, given that improvements in these three areas tend to predict changes within social competence.

Purpose of the Study

A primary goal of this investigation was to expand upon prior DBR-SIS research, which has been predominantly conducted with students without disabilities in the general education setting. The current research represents the first study to examine DBR-SIS psychometric defensibility in a population of students with substantial concerns beyond

those characteristics of typically developing children. Two research questions were of interest.

Research Question 1: Are DBR-SIS scales concurrently valid indicators of SDO data?

It was hypothesized that each DBR-SIS scale would be associated with its corresponding and construct-relevant SDO target. For example, in the case of the DBR-SIS DB target, we would predict a correspondence with the SDO-DB target. Furthermore, it was hypothesized that these convergent relations would exceed those of expected discriminant relations; that is, correlations between nonconstruct-relevant DBR and SDO targets.

Research Question 2: Is the concurrent validity of DBR-SIS scales consistent over time across three assessment periods (i.e., pre, mid-, and postintervention)?

It was hypothesized that findings would support the relation between DBR-SIS and SDO scales and that this relation would be consistent over time.

Method

Participants and Setting

The current investigation was a part of a larger randomized control trial examining the efficacy of the SCI-A program relative to a business-as-usual (BAU) control. SCI-A is an intervention designed for youth that would be characterized with social competence needs consistent with high-functioning forms of ASD or similar social challenges. The SCI-A curriculum is comprised of content targeted toward a specific social phenotypic profile as opposed to a specific *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; American Psychiatric Association, 2013) diagnostic profile (Stichter et al., 2010, 2012). It is delivered in a small group setting of four to six students across 31 lessons, across five scaffolded content units targeting facial expressions, sharing ideas, conversational reciprocity, and social problem solving. Although beyond the scope of this article, the reader is referred to recent publications for additional information regarding SCI-A, as well as data regarding the fidelity with which SCI-A was delivered in this particular investigation (Stichter et al., in press).

Prior to data collection, project personnel met with school officials to describe the nature of the curriculum and the characteristics of students for whom SCI-A was designed. Schools identified students whom they believed had substantial social needs and challenges, such as those common to high-functioning ASD. Referrals were collected via a demographic form, which was used to gather information regarding each student's social needs and what services were being presently provided. All referred students also had to meet the following criteria: cognitive functioning within normal range (based on full-scale IQ scores), access to general education environments/peers, and absence of significant mental health diagnoses or major behavioral incidents. SCI-A's inclusion and exclusion criteria are consistent with other similar social interventions for this population (Solomon, Goodlin-Jones, & Anders, 2004) and are necessary given the cognitive-behavioral underpinnings and group nature of the intervention.

Parental consent for data collection was obtained for total of 63 students across six SCI-A (n = 33) and six BAU (n = 30) schools. After parent consent, two students were discovered as ineligible based on misreported IQ information (accurate scores were well below threshold). Thus, 61 students were assessed at preintervention (SCI-A n = 32 and BAU n = 29). After attrition due to student programming changes and changes in school, the final sample completing assessment at mid-intervention included 57 students (SCI-A n = 29 and BAU n = 28). The resulting sample includes 58 students in middle and junior high school. Of the 58, 57 students were present from pre to postintervention, with one student dropping midway through the study. For the purposes of this article, this student's data were still considered in final analyses.

Of the current sample, across both the SCI-A and BAU settings, 53 participants were male and five were female. Participants ranged from 11 to 14 years of age. As often is the case, the school determination of service eligibility of this social phenotype varied considerably. The majority of participants were eligible for special education services in the state of Missouri in the following categories: Autism 44.8%, Emotional Disturbance 24.1%, Other Health Impairment 19.0%, Specific Learning Disability 3.4%, and Speech/Language Impairment 1.7%. Of the remaining students, 3.7% did not have a current individualized education plan (IEP) and 1.7% had a Section 504 Plan only. The majority of students were Caucasian/White (74%), with two participants identifying as African American, three as bi or multiracial, one as Hispanic, and one as Native American. (Note: 16% of the sample did decline to report race/ethnicity data.) All participants met screening criteria for cognitive ability as measured by the Wechsler tests of intelligence, with a mean Full-Scale IQ of 95.28 (SD = 14.86).

In addition, 23 core curriculum general education teachers (across the 12 participating middle schools) were recruited to participate in the rating of target student behavior in the general education setting. The majority of teachers provided instruction to target students in social studies or science classes. These classes were chosen as settings to observe and rate students due to increased opportunities for observation of peer interactions.

Measures

DBR-SIS. The DBR-SIS is an efficient tool that may be used to collect information regarding student behavior(s) immediately following an observation session. DBR-SIS has been described as a hybrid instrument, incorporating elements of SDO and behavior-rating scales (Chafouleas, 2011). It is akin to SDO in that it is used to collect information regarding student behavior within the time and setting that it is displayed. It is also similar to a behavior-rating scale in that data collection requires completion of a brief rating, which is completed using an 11-point unipolar graphic-rating scale. The scale extends from 0 to 10, with each point represent increments of 10% and anchors affixed to the beginning (0%; *Never*), middle (50%; *Sometimes*), and end (100%; *Always*) points. Within the current investigation, teachers received a daily email and completed DBR-SIS ratings within the Qualtrics system, allowing for electronic data collection, storage, and reporting. Teachers were asked to consider and rate the behavior their student exhibited during a particular target activity each day (e.g., social studies, science). It is estimated that it took a teacher 30 to 120 s to complete all ratings for a given student following each observation.

Three DBR behavior scales were examined within this analysis. DBR-DB was defined as an action that interrupts regular school or classroom activity. Examples included being out of one's seat, fidgeting, playing with objects, acting aggressively, or talking or yelling about things that are unrelated to classroom instruction. DBR-AE was defined as actively or passively participating in the classroom activity. Examples included writing, hand-raising, answering a question, talking about a lesson, listening to the teacher, reading silently, or looking at instructional materials. DBR-RB was defined as compliant and polite behavior in response to adult directions and interactions with peers and adults. Examples included following teacher direction, prosocial interaction with peers, positive response to adult request, or verbal or physical disruption without negative tone/connotation.

SDO. The SDO tool upon which this study was founded corresponded to the observation of classroom functioning (OCF), a researcher developed SDO tool intended to measure and monitor student functioning in the general education classroom. One general education classroom was selected per student and all SDO (via the OCF) occurred in this classroom. Observation times occurred during the same class period throughout the semester; however, they were varied within the class period to attempt a sampling of classroom behaviors at the beginning, middle, and end of the period. Observations occurred in a wide variety of classroom types (e.g., science, social studies, language arts), and classroom activities also varied based on school identification of most optimal opportunities to witness engagement or interaction by the student (e.g., large-group instruction, individual work, peer collaboration were coded in each OCF but are not the focus of the current analysis). As such, study hall or independent work classes were not targeted.

SDO data were collected across 15-min observation sessions, each of which was divided into 30-s intervals. The OCF employed both partial interval recording and momentary time sampling recording to estimate percentage of time target students engaged in relevant classroom behaviors. Three OCF behaviors were considered as part of this study. SDO-AE was defined as physical orientation to the teacher or current stimuli or active participation in the lesson or social interaction. Examples included turning to look at the SmartBoard or verbal response to a peer initiation. SDO-DB was defined as purposeful engagement in behavior that interrupts the natural flow of academic instruction or classroom functioning. Examples included calling out, out of seat, conversation unrelated to task, and aggression. *Noncompliance* (SDO-NC) was defined as failure to follow/complete verbal or gestural behavioral directions provided by the teacher to a group or target student within 5 s. Examples included student failing to look at the board when prompted or running away when directed to "come here." Note that SDO-DB and NC were coded using partial interval recording (where a behavior was marked as having occurred if it was observed at any point within each 30-s interval), whereas SDO-AE was coded using momentary time sampling (where a behavior was marked as having occurred if it was observed at the end of each 30-s interval). Partial interval was deemed appropriate given the typically irregular and brief, albeit still interruptive, nature of both DB and NC. Momentary time sampling was also considered appropriate given the expectation of frequent and nearly continuous AE within the classroom (which rules out partial interval recording as an option, which would otherwise credit even brief AE within an interval), while still allowing for occasional disengagement (which rules out whole interval recording, which would otherwise only credit full engagement across each interval).

Procedures

DBR-SIS teacher training. Prior to the start of data collection, research staff met with teacher participants for 30 min to introduce the DBR-SIS form. The training was modeled after procedures employed in previous DBR-SIS-related investigations (e.g., Kilgus, Riley-Tillman, Chafouleas, Christ, & Welsh, 2014), and included the following components: (a) introduction of the DBR-SIS and target behaviors; (b) description of how to complete DBR-SIS ratings,

including review of how to first anchor rating estimates at 0%, 50%, or 100%, and then sliding up or down depending on the most likely range of student performance; (c) description of DBR-SIS rating interpretation (i.e., the extent to which behavior was displayed in the observed class time); and (d) instruction on how to access and complete DBR-SIS ratings on a daily basis using the Qualtrics-based scale delivered via a web-based link.

SDO observer training. Nine research staff members were trained to collect SDO data. Training was delivered by two expert team members involved in the development of the OCF tool and took place across multiple days. Training consisted of the following components: (a) direct instruction and introduction to target behavior definitions, (b) orientation and practice using the interval recording/time sampling systems and data collection sheets, (c) modeling of appropriate codes, and (d) opportunities to practice to viewing and coding exemplar classroom videos. Once observers met reliability levels of 90% or above, interobserver agreement (IOA) levels were monitored throughout the study to determine if retraining was necessary. Over the course of data collection, reliability was maintained at 90% or above, thus no retraining was necessary. To examine IOA, an additional trained observer was present to conduct simultaneous observations during 37.6% of all sessions, with IOA sessions roughly equally distributed across time points. IOA was defined as percent agreement, which was calculated by dividing the number of intervals observers agreed on the presence or absence of the behavior by the total number of opportunities for agreement, multiplied by 100. IOA levels across behaviors ranged from 94.9% to 99.8%, with a mean of 97.2%.

Data collection. Teachers were emailed a unique DBR link daily, which prompted them to complete all DBR-SIS ratings immediately following the target class for that given day (e.g., social studies, science). Teachers were given 24 hr to complete the rating each day. DBR-SIS data were collected on a daily basis for approximately 4 to 5 months for each student across the three phases of intervention implementation (i.e., pre, mid-, and postintervention). Although the timing of each phase varied by student in accordance with his or her study start date, preintervention generally corresponded to January, mid-intervention to February and March, and postintervention to April and May. For the purposes of this study, SDO data were also collected during these same three time points. Specifically, three SDO data points were collected within each phase, with each data point corresponding to a 15-min observation. All observations occurred within a span of 5 to 7 school days, with each observation being separated by at least 1 school day.

Data Analysis Plan

Missing data. The current investigation, which was part of a larger randomized controlled trial, employed a two-method planned missing-data methodology (Graham, Taylor, Olchowski, & Cumsille, 2006). As part of this procedure, each student was randomized to one of two measurement groups. As a result of this randomization, 33 students (59.65%) were assessed using both DBR-SIS and SDO, whereas the remaining 24 (40.35%) were assessed using DBR-SIS only. Using this procedure increases the efficiency of the measurement process because fewer observations are needed, which also reduces the burden of the observational procedures on both research assistants and classroom teachers. Missing data were handled using full information maximum likelihood (FIML), which allowed for the inclusion of all 58 students in subsequent analyses. FIML was considered appropriate given the assumption of data being missing completely at random (MCAR). The MCAR assumption was founded in the use of randomization via the two-method missing-data design, which resulted in the probability of SDO missingness being unrelated to other variables as well as unmeasured values of SDO itself (Enders, 2010). Multiple researchers have identified FIML as a "state of the art" missing data-handling technique, with prior studies suggesting use of FIML in the presence MCAR data results in power recovery and unbiased parameter estimates (Baraldi & Enders, 2010; Schafer & Olsen, 1998). Through FIML, all 58 students were included in analyses, allowing maximum power of all statistical significance tests compared with traditional missing-data techniques such as pairwise or listwise deletion.

Data organization. Prior to analysis, we removed data dependencies in the form of multiple DBR and SDO data per student. This process was accomplished through the following data selection steps. First, research assistants identified the final SDO score collected within each particular time point (i.e., pre, mid, post). Second, research assistants identified two DBR-SIS data points, including (a) the DBR-SIS score collected on the same day as the final SDO score, as well as (b) the DBR-SIS score collected on the day prior to the SDO score. We then calculated the arithmetic mean of these two DBR-SIS data points. The decision to use 2-day mean DBR-SIS scores was made in the interest of deriving more robust indicators of student behavior within each particular time period (i.e., the points directly corresponding to the SDO data point under consideration). Subsequent to this data selection process, each student possessed three SDO data points and three DBR-SIS mean scores (i.e., pre, mid, post). These scores then served as the basis of all subsequent analyses.

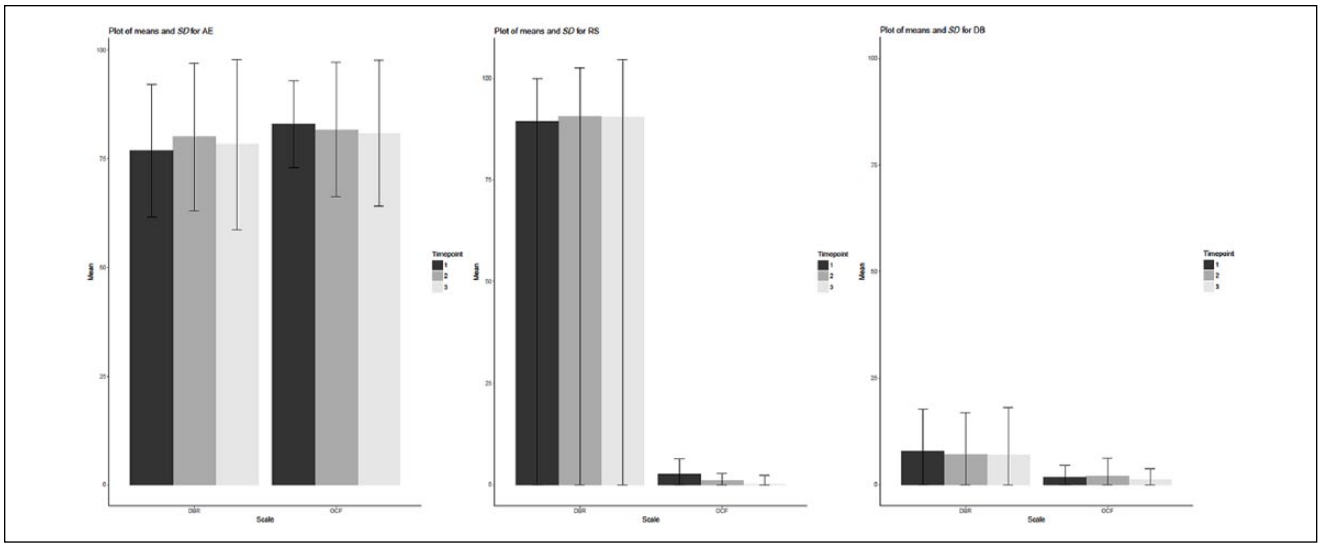

Figure 1. *Note.* AE = academic engagement; DBR = direct behavior rating; OCF = observation of classroom functioning; RB = respectful behavior; DB = disruptive behavior.

Bivariate correlations. Correlation coefficients were calculated to examine the relationship between each DBR and SDO target within each time point (i.e., pre, mid, post). Hypothesized convergent relations corresponded to the pairings of (a) DBR-AE and SDO-AE, (b) DBR-DB and SDO-DB, and (c) DBR-RB and SDO-NC. All other DBR–SDO pairings were hypothesized to be discriminant relations and were thus expected to be lower in magnitude relative to convergent relations. We followed Cohen's (1988) guidelines for effect size interpretations of correlation magnitudes, where $r \geq .10$ was considered small, $r \geq .30$ medium, and $r \geq .50$ large. In the interest of limiting overinterpretation of spurious or nonmeaningful relations, conclusions regarding the presence of concurrent criterion-related validity were limited to medium and large correlations.

Next, correlation coefficients were compared across time points within each DBR-SDO pairing to examine the extent to which correlation magnitude varied over time. This testing was accomplished via chi-square nested model comparisons between a model with correlations freely estimated across time to a model that specified correlational equivalence (H_0: $\rho_1 = \rho_2 = \rho_3$). Finally, a single overall correlation was estimated and evaluated within each DBR-SDO pair to evaluate the relationship between each measure across all time points. All correlations were estimated with Mplus v. 7.11 (Muthén & Muthén, 2012).

Results

See Figure 1 for plots of means and standard deviations across convergent DBR-SDO pairings and time points. A review of the bar graphs suggests DBR-SIS and SDO yielded scores that were similar in terms of level; such mean scores appeared similar within each DBR-SDO pairing and across time points. Standard decisions were notably different, however, with error bars proving to be smaller for SDO than DBR-SIS.

See Table 1 for a summary of correlational findings. With regard to the convergent DBR-AE/SDO-AE pairing, one within-time correlation fell in the small range (i.e., <.30 in the pre wave), one in the medium range (i.e., .30 ≤ .49 in the mid wave), and one in the large range (i.e., ≥.50 in the post wave). Similar findings occurred across the (a) DBR-DB/SDO-DB and (b) DBR-RB/SDO-NC, with 66% of within-time correlations falling in the medium range, 17% in the small range, and 17% below the small range. Despite such variance in correlational magnitude within each DBR-SDO pairing, the chi-square difference test findings indicated that the differences in the within-time correlations were not significantly different from zero within any of the five DBR-SDO pairings.

Overall, findings suggested variance across DBR-SDO pairings in terms of correlation magnitudes. Similar findings were noted for across-time correlations, which also varied between pairings. Specifically, DBR-DB/SDO-DB and DBR-AE/SDO-AE fell in the medium range, and DBR-RB/SDO-NC in the small range.

With regard to hypothesized convergent and discriminant relations, when examining across-time correlations, convergent relations exceeded those of expected discriminant relations for both DBR-AE/SDO-AE (relative to DBR-AE/SDO-DB and DBR-AE/SDO-NC) and DBR-DB/SDO-DB (relative to DBR-DB/SDO-AE and DBR-DB/

Table 1. Correlations Between DBR and SDO Targets Across and Within Time, as Well as Within-Time Correlation Chi-Square Tests.

DBR	SDO	Within-Time Correlations			Across-Time Correlation	Within-Time Correlation Equivalence Tests		
		Time 1	Time 2	Time 3		χ^2	*df*	*p*
AE	**AE**	**0.21***	**0.33***	**0.54***	**0.45***	**0.06**	**2**	**.97**
AE	DB	−0.14	−0.48*	−0.21	−0.33*	2.05	2	.36
AE	NC	−0.36*	−0.21	−0.28	−0.22	2.28	2	.32
DB	AE	−0.21	−0.07	−0.27	−0.17	0.96	2	.64
DB	**DB**	**0.25**	**0.32***	**0.37***	**0.32***	**3.19**	**2**	**.20**
DB	NC	0.49*	0.03	0.51*	0.31*	6.30	2	.04
RB	AE	0.14	0.29*	0.50*	0.32*	3.32	2	.19
RB	DB	−0.49*	−0.69*	−0.36*	−0.54*	4.32	2	.12
RB	**NC**	**–0.40***	**–0.02**	**–0.36***	**–0.25**	**2.37**	**2**	**.31**

Note. Bolded rows represent hypothesized convergent DBR-SDO relations. DBR = direct behavior rating; SDO = systematic direct observation; AE = academic engagement; DB = disruptive behavior; NC = noncompliance; RB = respectful behavior.
*$p < .05$.

SDO-NC). Of the nine possible within-time DBR-SDO comparisons (i.e., three time points × three convergent pairings), the convergent pairing exceeded the discriminant pairings only 22% of the time (i.e., two out of nine). These included the (a) DBR-AE/SDO-AE pairing within Time 3, which exceeded the DBR-AE/SDO-DB and DBR-AE/SDO-NC pairings, and (b) the DBR-DB/SDO-DB pairing within Time 2, which exceeded the DBR-DB/SDO-AE and DBR-DB/SDO-NC pairings.

Discussion

The broader purpose of the current investigation was to examine DBR-SIS concurrent criterion-related validity within a sample of students with substantial social competence difficulties, the majority of whom qualified for special education services. Analyses considered the extent to which each DBR-SIS target predicted its corresponding and construct-relevant SDO target. A review of findings indicated all correlations were in the expected direction. Whereas medium correlations were noted between SDO and DBR for AE and DB (.45 and .32, respectively), smaller correlations were found for respect (–.25). The current findings are commensurate with prior research in regard to correlational magnitude for AE and DB (Chafouleas et al., 2012). The pattern of findings is also similar to prior research, where AE and DB validity outperformed that of RB (Chafouleas et al., 2013; Kilgus, Riley-Tillman, et al., 2014). Thus, in consideration of the aforementioned criteria for evaluating correlation magnitude, findings afford somewhat weak preliminary support for the concurrent convergent validity of the DBR-AE and DBR-DB targets, and no support for the DBR-RB target.

Interestingly, the pattern of correlations did not conform to theory-driven expectations, as some expected discriminant relations exceeded those of expected convergent relations within certain time points and at the overall level. Accordingly, though the current findings do yield some support for DBR-SIS concurrent convergent validity (as described above), support for the more nuanced concept of construct validity is somewhat lacking given the discriminant validity evidence. Support for the DBR-RB target was found to be particularly lacking, as the expected convergent relation did not exceed discriminant relations within any of the time points or at the overall level.

Despite the somewhat equivocal construct validity findings, the current findings suggest some promise for the use of DBR-SIS within a high needs population. Given the extensive research base upon which the DBR-SIS "Big 3" targets are built, it would be expected each target would perform well in progress monitoring for a wide range of students, including those with more moderate and/or high levels of need. Despite having been designed and intended for the larger typically developing population, it is anticipated each of the three DBR-SIS targets would function well with students exhibiting substantial social competence deficits. This is given the expectation that the DBR-SIS targets would serve as social-behavioral general outcome measures for all students, regardless of level of need (Chafouleas, 2011). Given this foundation, it should be unsurprising that teachers are capable of yielding ratings of students with social competence difficulties that are moderately to highly associated with SDO data indicative of such social-behavioral functioning.

In consideration of these expectations for DBR-SIS performance, one might question the pattern of observed findings, where only two of the three targets were weakly

supported by the current evidence. One potential contributor to this finding might be differences in the observability of behaviors, with RB being potentially more nuanced and subjective, and thus more difficult for raters to evaluate. An additional contributor to the pattern of findings includes differences across DBR-SDO pairings in the correspondence between target definitions. A review of the DBR-SIS and SDO operational definitions of AE and DB revealed strong similarities in conceptualization and wording (with a notable caveat raised in relation to DB, with an important distinction of behavioral intentionality across the definitions). This suggested that both DBR-SIS and SDO were examining similar behavior. In contrast, correspondence for the DBR-SDO pairing of RB (DBR-SIS) and NC (SDO) was much weaker and indirect (and was the smallest correlation). Although not a perfect match, justification for the pairing of these two targets was founded in (a) recognition that the two behaviors share an antonymic relation, with noncompliant behaviors being incompatible with RB (and thus hypothetically inversely related), and (b) the history of DBR research, wherein "compliance" served as a target of interest and "Big 3" member prior to being recast as "respectful" (Christ, Riley-Tillman, & Chafouleas, 2009). Nevertheless, that this DBR-SDO relationship was more distal relative to the other pairings (with compliance representing only a portion of RB) suggests the need to examine the respectful DBR target relative to alternative SDO targets in future studies, such as direct measures of respect or narrower constituent behaviors (e.g., prosocial interactions with peers).

DBR-SIS and SDO data were then disaggregated and compared within time point. Results suggested that the relationships between DBR-SIS and SDO scales were small to moderate at baseline. Correlations were then found to (a) increase over time for DBR-SIS scales of DB and AE, and (b) vary in a somewhat curvilinear fashion for RB, with pre and postintervention proving similar and mid-intervention lower than the other time points. Follow-up analyses suggested these differences were not statistically significant. With that said, this null finding might be related to the limited power of the current design, which was founded upon only 58 students. Nevertheless, results suggest DBR-SIS performance is likely to vary over time. In considering reasons for such variance in performance, particularly with regard to RB, one potential explanation pertained to differences across assessment tools regarding sensitivity to change in student behavior in response to SCI-A. That is, it was of interest to determine whether the DBR-SIS scales did not document change in student behavior documented by SDO, thus creating discordance between the methods. Yet, a review of descriptive statistics associated with each assessment suggested similar patterns of change in means and standard deviations within each DBR-SDO pairings. Specifically, both measures documented modest changes in student behavior across time and in response to SCI-A. As such, though difference in sensitivity to change is an unlikely explanation for this discordance within certain DBR-SDO pairings, findings are promising given that sensitivity to change is a necessary and vital psychometric property in justifying applied use of a method for progress-monitoring purposes.

In sum, initial research regarding the use of DBR-SIS with students exhibiting more significant behavioral concerns has yielded inconsistent results. Findings are promising, in that for at least two of the three behaviors, teachers were able to generate brief ratings of student behavior that correlate meaningfully with SDO data collected by an external observer using an intensive coding procedure. At this stage of the research line, we can draw two conclusions. First, there is evidence that DBR-SIS has the potential to be used as a viable method for progress monitoring in relation to children with substantial social competence deficits who are most commonly examined with direct observation. This is, of course, pending the completion of additional research and the accumulation of promising findings.

Second, the current investigation yields large implications for future DBR-SIS research and practice. The design and data collection procedures are unique in that they led to the amassing of a sizable amount of data both within and across student participants. This volume of data has not been observed within DBR-SIS work to date, thus allowing this study to indicate the potential influence of DBR-SIS data within applied service delivery. Without sizable recourses, time, and effort, it is traditionally pragmatically difficult to collect the necessary amount of formative data on children with special needs across the general education environment. Research has suggested it is necessary to collect at least one SDO data point a day across 5 days within any given phase (e.g., baseline, intervention) to reliably estimate a student's behavior (Briesch, Chafouleas, & Riley-Tillman, 2010). One can imagine the organizational hardship likely to be encountered when attempting to coordinate such an extensive data collection endeavor across the four to six students comprising a SCI group. Relative to this scenario, the DBR-SIS represents a simple and efficient method to collect daily data on target students in their general setting without interrupting the class, pushing in an outside observer, or requiring the dedication of extensive time or resources. It also allows for the collection of data within settings where data collection often does not take place at all given the problems associated with assessment coordination. Anecdotally, teachers reported few difficulties with data collection and the capacity to include the assessment process into their instructional routine. The present investigation is therefore considered promising in that it suggests the potential for feasible formative assessment within relative to high needs students within general education settings.

Limitations

Limitations to this investigation should be noted. First, the student sample is considered somewhat restricted, with analyses corresponding to only 58 students with social competence difficulties. Although FIML supported a degree of power recovery, yielding a sample similar in size to previous DBR-SIS research (e.g., Chafouleas, Kilgus, & Hernandez, 2009; Chafouleas et al., 2012), the somewhat small sample both (a) limited the generalizability of findings and (b) restricted statistical power, thus increasing the potential for Type II errors.

Second, though teacher DBR-SIS training procedures were modeled upon previous work with DBR training, specific video-training modules developed in consideration of this research were not used given scheduling difficulties and time limitations. As such, it is possible that DBR-SIS may have functioned differently if administered in the presence of more systematic training. With that said, it should be noted that the level of training provided is similar to that used within previous investigations (Chafouleas et al., 2013; Kilgus, Riley-Tillman, et al., 2014), suggesting the current findings align well with previous research. Third, though anecdotal evidence suggests DBR-SIS was feasible for teachers, the absence of social validity evidence (e.g., via the Usage Rating Profile–Assessment; Briesch, Chafouleas, Neugebauer, & Riley-Tillman, 2013; Chafouleas, Riley-Tillman, Briesch, & Chanese, 2008) limits objective conclusion regarding the acceptability of the DBR-SIS method and associated data collection procedures. Fourth, due to the remote nature of DBR-SIS data collection, we were unable to examine the fidelity with which DBR-SIS ratings were completed. Future investigations should look to more directly evaluate this issue.

Future Directions for Research

Moving forward, additional research is needed to support the applied use of DBR-SIS in evaluating students with social competence difficulties. First, it is that recommended future research employ larger and more diverse samples of varying age ranges (e.g., early and upper elementary) in the interest of increasing the external validity of study findings. In addition, the environment should be varied to look both at general education settings (as in this study) and special education environments. In addition, researchers should look to control for level of functioning, permitting evaluation of the extent to which DBR-SIS functions similarly for students with varying levels of symptomatology. Second, DBR-SIS performance should be evaluated with students with different types of concerns, including those related to attentional or emotional problems. Such research would support further consideration of DBR-SIS defensibility within populations characterized by higher levels of need. Third, DBR-SIS scales should be considered relative to alternative criterion scales with greater correspondence to DBR definitions. Such research might consider SDO "appropriate interaction" targets with definitions representative of both contextual norms and behavior within social interactions. Fourth, researchers should consider the extent to which variance in DBR-SIS is attributable to the individual completing the ratings. Findings of such studies, including those founded upon generalizability theory (e.g., Briesch et al., 2010), would speak to the requisite number of raters required to achieve sufficient reliability, as well as the types of professionals who are best able to yield accurate ratings (e.g., classroom teachers, school psychologists). Finally, future researchers should examine DBR-SIS sensitivity to change in student behavior in response to intervention. The National Center for Intensive Intervention (intensiveintervention.org) has identified such psychometric evidence as necessary in justifying applied use of a measure, thus making its collection a priority within the DBR-SIS line.

Declaration of Conflicting Interests

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

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, *48*, 5–37.

Briesch, A. M., Chafouleas, S. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2013). Assessing influences on intervention use: Revision of the Usage Rating Profile. *Journal of School Psychology*, *51*, 81–96.

Briesch, A. M., Chafouleas, S. M., & Riley-Tillman, T. C. (2010). Generalizability and dependability of behavior assessment methods to estimate academic engagement: A comparison of systematic direct observation and Direct Behavior Rating. *School Psychology Review*, *39*, 408–421.

Brown, W. H., Odom, S. L., & Conroy, M. A. (2001). An intervention hierarchy for promoting young children's peer interactions in natural environments. *Topics in Early Childhood Special Education*, *21*, 162–175. doi:10.1177/027112140102100304

Burns, M. K., Riley-Tillman, T. C., & VanDerHeyden, A. (2012). *RTI Applications: Vol. 1. Academic and behavioral interventions*. New York, NY: Guilford Press.

Burt, K. B., & Roisman, G. I. (2010). Competence and psychopathology: Cascade effects in the NICHD Study of Early

Child Care and Youth Development. *Development and Psychopathology*, *22*, 557–567.

Chafouleas, S. M. (2011). Direct Behavior Rating: A review of the issues and research in its development. *Education and Treatment of Children*, *34*, 575–591. doi:10.1353/etc.2011.0034

Chafouleas, S. M., Kilgus, S. P., & Hernandez, P. (2009). Using Direct Behavior Rating (DBR) to screen for school social risk: A preliminary comparison of methods in a kindergarten sample. *Assessment for Effective Intervention*, *34*, 214–233. doi:10.1177/1534508409333547

Chafouleas, S. M., Kilgus, S. P., Jaffery, R., Riley-Tillman, T. C., Welsh, M. E., & Christ, T. J. (2013). Direct Behavior Rating as a school-based behavior screener for elementary and middle grades. *Journal of School Psychology*, *51*, 367–385. doi:10.1016/j.jsp.2013.04.002

Chafouleas, S. M., McDougal, J. L., Riley-Tillman, T. C., Panahon, C. J., & Hilt, A. M. (2005). What do Daily Behavior Report Cards (DBRCs) measure? An initial comparison of DBRCs with direct observation for off-task behavior. *Psychology in the Schools*, *42*, 669–676. doi:10.1002/pits.20102

Chafouleas, S. M., Riley-Tillman, T. C., Briesch, A. M., & Chanese, J. A. (2008). The development and validation of the Usage Rating Profile (URP). *Canadian Journal of School Psychology*, *23*, 175–189. doi:10.1177/0829573508323989

Chafouleas, S. M., Riley-Tillman, T. C., & Sassu, K. A. (2006). Acceptability and reported use of Daily Behavior Report Cards among teachers. *Journal of Positive Behavior Interventions*, *8*, 174–182. doi:10.117/10983007060080030601

Chafouleas, S. M., Riley-Tillman, T. C., Sassu, K. A., LaFrance, M. J., & Patwa, S. S. (2007). Daily Behavior Report Cards: An investigation of the consistency of on-task data across raters and methods. *Journal of Positive Behavior Interventions*, *9*, 30–37. doi:10.1177/10983007070090010401

Chafouleas, S. M., Riley-Tillman, T. C., & Sugai, G. (2007). *School-based behavioral assessment: Informing intervention and instruction*. New York, NY: Guilford Press.

Chafouleas, S. M., Sanetti, L. M. H., Kilgus, S. P., & Maggin, D. M. (2012). Evaluating sensitivity to behavioral change across consultation cases using Direct Behavior Rating Single-Item Scales (DBR-SIS). *Exceptional Children*, *78*, 491–505.

Christ, T. J., Riley-Tillman, T. C., & Chafouleas, S. M. (2009). Foundation for the development and use of Direct Behavior Rating (DBR) to assess and evaluate child behavior. *Assessment for Effective Intervention*, *34*, 201–213.

Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2nd ed.). New York, NY: Academic Press.

Cook, B. G., & Cook, S. C. (2013). Unraveling evidence-based practices in special education. *The Journal of Special Education*, *47*, 71–82. doi:10.1177/0022466911420877

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, *37*, 184–192.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to responsiveness-to-intervention: What, why, and how valid is it? *Reading Research Quarterly*, *4*, 93–99.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, *11*, 323–343. doi:10.1037/1082-989X.11.4.323

Gresham, F. M. (1981). Social skills training with handicapped children: A review. *Review of Educational Research*, *51*, 139–176.

Gresham, F. M., Sugai, G., & Horner, R. H. (2001). Interpreting outcomes of social skills training for students with high-incidence disabilities. *Exceptional Children*, *67*, 331–344.

Hintze, J. M. (2005). Psychometrics of direct observation. *School Psychology Review*, *34*, 507–519.

Hintze, J. M., & Matthews, W. J. (2004). The generalizability of systematic direct observations across time and setting: A preliminary investigation of the psychometrics of behavioral observation. *School Psychology Review*, *33*, 234–248.

Kilgus, S. P., Collier-Meek, M. A., Johnson, A. H., & Jaffery, R. (2014). Applied empiricism: Ensuring the validity of causal response to intervention decisions. *Contemporary School Psychology*, *18*, 1–12. doi:10.1007/s40688-013-0009-z

Kilgus, S. P., Riley-Tillman, T. C., Chafouleas, S. M., Christ, T. J., & Welsh, M. (2014). Direct Behavior Rating as a school-based behavior universal screener: Replication across sites. *Journal of School Psychology*, *52*, 63–82. doi:10.1016/j.jsp.2013.11.002

Miller, F. G., Riley-Tillman, T. C., Chafouleas, S. M., & Schardt, A. A. (2017). Direct Behavior Rating instrumentation: Evaluating the impact of scale formats. *Assessment for Effective Intervention*, *42*, 119–126.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

National Center on Intensive Intervention. (2017). *National center on intensive intervention*. Washington, DC: U.S. Department of Education, Office of Special Education Programs. Available from http://www.intensiveintervention.org/

Odom, S. L., Cox, A. W., & Brock, M. E. (2013). Implementation science, professional development, and autism spectrum disorders. *Exceptional Children*, *79*, 233–251.

Riley-Tillman, T. C., Chafouleas, S. M., Briesch, A. M., & Eckert, T. (2008). Daily Behavior Report Cards and systematic direct observation: An investigation of the acceptability, reported training and use, and decision reliability among school psychologists. *Journal of Behavioral Education*, *17*, 313–327.

Riley-Tillman, T. C., Chafouleas, S. M., Sassu, K. A., Chanese, J. A. M., & Glazer, A. D. (2008). Examining the agreement of direct behavior ratings and systematic direct observation for on-task and disruptive behavior. *Journal of Positive Behavior Interventions*, *10*, 136–143.

Rogers, S. J. (2000). Interventions that facilitate socialization in children with autism. *Journal of Autism and Developmental Disorders*, *30*, 399–409. doi:10.1023/A:1005543321840

Schafer, J. L., & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research*, *33*, 545−571.

Schultz, T. R., Schmidt, C. T., & Stichter, J. P. (2011). A review of parent education programs for parents of children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *26*, 96–104. doi:10.1177/1088357610397346

Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children*. New York, NY: Guilford Press.

Solomon, M., Goodlin-Jones, B. L., & Anders, T. F. (2004). A social adjustment enhancement intervention for high functioning autism, Asperger's syndrome, and pervasive developmental disorder NOS. *Journal of Autism and Developmental Disorders*, *34*, 649–668.

Stichter, J. P., Herzog, M. J., Kilgus, S. P., & Schoemann, A. M. (2018). Exploring the moderating effects of cognitive abilities on Social Competence Intervention outcomes. *Behavior Modification*, *42*, 84–107.

Stichter, J. P., Herzog, M. J., Owens, S. A., & Malugen, E. (2016). Manualization, feasibility, and effectiveness of the school-based Social Competence Intervention for Adolescents (SCI-A). *Psychology in the Schools*, *53*, 583–600. doi:10.1002/pits.21928

Stichter, J. P., Herzog, M. J., Visovsky, K., Schmidt, C., Randolph, J., Schultz, T., & Gage, N. (2010). Social Competence Intervention for youth with Asperger syndrome and high functioning autism: An initial investigation. *Journal of Autism and Developmental Disorders*, *40*, 1067–1079. doi:10.1007/s10803-010-0959-1

Stichter, J. P., O'Connor, K., Herzog, M. J., Lierheimer, K., & McGhee, S. (2012). Social Competence Intervention for elementary students with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders*, *42*, 354–366. doi:10.1007/s10803-011-1249-2

Volker, M. A. (2012). Introduction to the special issue: High-functioning autism spectrum disorders in the schools. *Psychology in the Schools*, *49*, 911–916.

Walker, H. M., Irvin, L. K., Noell, J., & Singer, G. H. S. (1992). A construct score approach to the assessment of social competence: Rationale, technological considerations, and anticipated outcomes. *Behavior Modification*, *16*, 448–474.

Wentzel, K. R., & Watkins, D. E. (2002). Peer relationships and collaborative learning as contexts for academic enablers. *School Psychology Review*, *31*, 366–377.