Exceptional Children
2020, Vol. 86(2) 137–154

DOI: 10.1177/0014402919855321
journals.sagepub.com/home/ecx
SAGE

Moving Beyond Fidelity: Assessing Implementation of a Comprehensive Treatment Program for Adolescents With Autism Spectrum Disorder

Jessica R. Steinbrenner[1], Samuel L. Odom[1], Laura J. Hall[2], and Kara Hume[1]

Abstract

Assessing implementation of comprehensive treatment programs is a multifaceted process that should extend beyond measuring solely treatment fidelity. The purpose of this article is to describe and demonstrate a thorough process for assessing implementation and receipt of a comprehensive treatment program for students with autism spectrum disorders in high schools. The Center on Secondary Education for Students With Autism Spectrum Disorders (CSESA) developed a profile and index approach to assess implementation of seven features related to the CSESA intervention process and content. In a cluster-randomized trial, the CSESA implementation profile captured information about diverse features of implementation, and this index successfully differentiated between intervention and control schools. A multifeatured implementation profile and index is an important step forward for examining and implementing comprehensive treatment programs in educational settings.

Youth with autism spectrum disorder (ASD) are entering high schools at an increasing rate (U.S. Department of Education, 2017). Comprehensive treatment programs for children and youth with ASD have been developed to address learning needs (National Research Council, 2001). Such programs, when employed in educational contexts, are effective only if implemented fully by school personnel and delivered to identified students. The Center on Secondary Education for Students With Autism Spectrum Disorders (CSESA), a research and development center funded by the Institute of Education Science, has developed a comprehensive treatment program for adolescents with ASD. A challenge for researchers of such programs, including CSESA, has been to measure all aspects of implementation, such as staff training and support, staff delivery of the intervention components, and the degree to which the students with ASD in a school receive those intervention components. The purposes of this article are to demonstrate the use of an implementation index assessment approach developed for the CSESA program and to examine

[1]University of North Carolina–Chapel Hill
[2]San Diego State University

Corresponding Author:
Jessica R. Steinbrenner, Frank Porter Graham Child Development Institute, University of North Carolina–Chapel Hill, 517 S. Greensboro Street CB #8040, Chapel Hill, NC 27599-8040.
E-mail: jessica.dykstra@unc.edu

its utility in differentiating between schools that did and did not implement the CSESA program.

In the field of special education, the combination of the developmental features of adolescence (Caskey & Anfara, 2014), the range of characteristics that occur across the autism spectrum (Volkmar, Reichow, Westphal, & Mandell, 2014), and the multifarious nature of high school programs converge to create a "perfect storm of complexity" (Odom, Duda, Kucharczyk, Cox, & Stabel, 2014). This situation has risen in seriousness for the U.S. educational system for several reasons. First, there has been a 189% increase in the number of high school students with ASD qualifying for special education services from 2006 to 2015 (U.S. Department of Education, 2017). Second, the outcomes for young adults with ASD who transition out of high school programs are among the worst of any disability group (Shattuck et al., 2012). Third, teachers do not feel prepared to deliver effective instructional interventions that could result in more positive outcomes for students with ASD (Knight, Huber, Kuntz, Carter, & Juarez, 2019).

To address the variety of learning needs exhibited by children and youth with ASD, researchers and program developers have designed comprehensive treatment programs (National Research Council, 2001). Such programs are termed "comprehensive" because they focus on multiple learning needs typically occurring for children and youth with ASD. Other characteristic features of these programs are an identifiable conceptual framework, sustained implementation across time (e.g., a year or more), program intensity (e.g., occurrence multiple days of the week), and "manualization," or written protocols that document the intervention procedures (Odom et al., 2014; Odom, Boyd, Hall, & Hume, 2010). Other publications have differentiated comprehensive treatment programs from focused intervention practices (i.e., practices focusing on single skills; Odom, Collet-Klingenberg, Rogers, & Hatton, 2010) or modular interventions (i.e., a package of practice focusing on a single skills area; Kasari, 2015). Examples of comprehensive treatment programs are the Early Start Denver Model (Rogers & Dawson, 2010), the LEAP (Learning Experiences and Alternative Program for Preschoolers and Their Parents) program (Strain & Bovey, 2011), Pivotal Response Treatment (Koegel & Koegel, 2006; Stahmer, Suhrheinrich, Reed, & Schreibman, 2012), and the variety of early intensive behavioral intervention programs that is most typified by the Lovaas model (Lovaas, 2003).

The CSESA Program

Comprehensive treatment programs focusing on ASD have been developed primarily for infants and toddlers, preschool children, and children of elementary school age (Odom et al., 2014). To date, there has been no publication of a comprehensive treatment program designed for adolescents with ASD attending public high schools. To address this void, the Institute of Educational Sciences funded CSESA to develop such a program. The first 2 years were devoted to developing and pilot testing the implementation of the program, and the last 3 years consisted of a randomized control trial (RCT) examining implementation and effects of the intervention (i.e., the efficacy analysis is currently being conducted and will appear in a separate paper). Initial reviews of the literature indicated that the greatest learning needs for adolescents with ASD were in literacy (Fleury et al., 2014), social competence (Carter et al., 2014), independence (Hume, Boyd, Hamm, & Kucharczyk, 2014), and transition that also involves families (Test, Smith, & Carter, 2014). National experts in each respective area collaborated with CSESA staff to adapt or develop interventions in each specific area, design fidelity measures, and conduct pilot tests. During this 1st year, focus groups of constituents and stakeholders were also conducted to provide advice about the interventions being developed (Kucharczyk et al., 2015). Individual interventions were revised based on the information in the 1st year. In the 2nd year, the CSESA researchers conducted a pilot study of multiple combinations of

interventions, which included trialing the implementation measures.

The resulting CSESA model was based on the work from the National Professional Development Center for Autism. The model has overarching professional development for school personnel in the form of training and coaching support provided by CSESA coaches (i.e., research team members) throughout the 2 years of active study implementation. The CSESA coaches support the school staff in a process of (1) assessment (i.e., of school programs for students with ASD and individual students with ASD), (2) planning (i.e., for school rollout of the CSESA model and selecting goals and matching interventions for individual students), (3) intervention (i.e., implementing CSESA interventions, described in detail in the subsequent paragraph), and (4) monitoring outcomes (i.e., tracking progress of school rollout, staff fidelity of intervention delivery and individual student goals). At the end of the 2 years of active implementation in the current study, CSESA coaches worked with staff at each school to develop a sustainability plan for continuing the CSESA model outside the context of the study.

The CSESA intervention covers four domains: academic (i.e., literacy focused), social competence and peer relationships, independence and behavior, and transition and families. Because of the range of abilities of adolescents with ASD, different interventions had to be developed for students who followed the general education curriculum to receive a diploma (i.e., typically with lesser degrees of autism severity) and those who followed a modified curriculum (i.e., typically with higher degrees of autism severity that often includes intellectual disability). In the *literacy* domain, the Alternative Achievement Literacy (Browder, Ahlgrim-Delzell, Flowers, & Baker, 2012) was adapted for students on the modified diploma track who were learning to read, and the Collaborative Strategic Reading (Reutebuch, El Zein, Kim, Weinberg, & Vaughn, 2015) approach was adapted for high school students with ASD who were reading to learn but had challenges with reading comprehension and were primarily on the diploma track. For *social competence*, the Social Competence Intervention (Stichter, Herzog, Owens, & Malugen, 2016) was adapted for high schoolers as a social skills training program designed for students on the diploma track. For all students, the peer-mediated work of Carter and colleagues (2017) was adapted to promote social interactions and relationships with neurotypical peers. For *independence and behavior*, CSESA staff developed the PRISM intervention that used the Secondary School Success Checklist to identify student needs (Hume et al., 2018) and select specific evidence-based practices (Wong et al., 2015) to address learning goals. The *transition and families* domain drew on the work of Test and colleagues (2014) to design tools focusing on community resource mapping, transition planning, student involvement in individualized education programs (IEPs), and work-based learning. The Transitioning Together intervention for families (DaWalt, Greenberg, & Mailick, 2018) was also a part of this domain of CSESA. The CSESA program has been described in greater detail in other papers (Hume & Odom, 2019; Odom et al., 2014).

CSESA Program Implementation Rationale and Assessment

The development of focused intervention practices, modular treatment programs, and comprehensive treatment programs is the first step in a process toward building the provision and quality of learning opportunities for youth with ASD. Implementation of the programs in authentic settings, such as high schools, is a next and challenging step in the process. For years, leaders in the field have recommended that implementers document program procedures (Gresham, Gansle, & Noell, 1993). In many cases, the assessment of implementation has focused specifically on the service provider employing procedures as specified by the purveyor of the intervention, which has been described as *treatment adherence, treatment integrity*, or *treatment fidelity* (Wolery, 2011). In fact, researchers of comprehensive treatment programs have often

used treatment fidelity (Colombi et al., 2018; Stock, Mirenda, & Smith, 2013; Strain & Bovey, 2011) or quantitative measures of how much treatment is delivered (e.g., number of hours of therapy; Dawson & Rogers, 2011) as single indices of implementation. Although fidelity and amount of treatment time is important, it may underrepresent the complexities of implementation of comprehensive treatment programs in school settings.

Although fidelity and amount of treatment time is important, it may underrepresent the complexities of implementation of comprehensive treatment programs in school settings.

The emergence of implementation science (Eccles & Mittman, 2006) has highlighted variables within organizational systems, such as schools, that establish implementation as multidimensional (Wong, Ruble, McGrew, & Yu, 2018) or multilevel (Stahmer, Suhrheinrich, Schetter, & Hassrick, 2018). One dimension of implementation, as noted previously, consists of delivery of the treatment with fidelity (O'Donnell, 2008), with some researchers proposing that measurement of fidelity itself consists of multiple dimensions (Sutherland, McLeod, Conroy, & Cox, 2013). Other researchers note that implementation of a program consists of not only the fidelity of delivery or adherence to the program but also the supports that are necessary for the provider to achieve fidelity. This dimension of implementation could include initial training for the provider (Brookman-Frazee, Drahota, & Stadnick, 2012), implementation planning and school consultation (Sanetti, Williamson, Long, & Kratochwill, 2018), and active in-school or remote coaching by a purveyor or support personnel after training occurs (Artman-Meeker, Hemmeter, & Snyder, 2014; Wong et al., 2018).

Yet another dimension of implementation in school-based programs is the degree to which students receive (i.e., engage in) the program (Low, Van Ryzin, Brown, Smith, & Haggerty, 2014). In individual therapist–client intervention programs, the degree to which the client receives an intervention can be documented by therapy hours or hours in treatment, as noted previously. In most school-based programs, teachers or related services professionals provide services to several students, and it is possible that not all students will receive all components of the intervention program. This is even more of a reality in high school settings, when students may be switching classes or have a wide range of teachers during the school year. The degree to which the students receive all components of the program is an important dimension of implementation (Bishop et al., 2014; Dane & Schneider, 1998; Low et al., 2014).

For comprehensive educational programs delivered in schools, Cordray and colleagues (Cordray, Poin, Brandt, & Molefe, 2013; Nelson, Cordray, Hulleman, Darrow, & Sommer, 2010) proposed that evaluation of implementation is by necessity multidimensional. That is, it should assess the training and preparation provided to the *school* for implementing the program, the elements of the intervention program that the *school staff* actually delivers to students, and the degree to which the *students* eligible for the intervention receive the intervention components. For the CSESA program, we adapted the approach that Cordray and colleagues have employed by including the multiple dimensions just noted.

Because this method of implementation assessment consists of multiple levels (i.e., school level, staff level, and student level), we employed an implementation profile approach in which assessment of individual features are displayed within levels (described in the next section) and also aggregated into a single index score (Table 1). The advantages of having both a profile of individual indicators of implementation and a single index are that one can (a) compare schools on the indicators of implementation that occurred or did not occur in individual schools, (b) have a summary score that displays the range of implementation that occurs among schools that have adopted a treatment model, and (c) discriminate between intervention and control

Table 1. Features of the CSESA Implementation Index.

Feature	Data source	Areas of process and content	Metric	Informant	Notes
SCHOOL: *Key Questions: Is the CSESA process being implemented at the school? Is the school receiving the CSESA model?*					
CSESA component trainings	Training log	Professional development; component areas	Number of trainings received and number of domains covered	Coach	Training provided to and received by key personnel
CSESA coaching	Coaching log	Professional development	Number of hours of coaching across 2-year project	Coach	
STAFF: *Key Questions: Is the staff implementing the CSESA interventions as designed? Is the staff following the CSESA process?*					
School-level planning	CSESA school planning document	Assessment; planning	Completion of CSESA school plan	Artifact	Dynamic document completed with coach and autism team
Intervention quality	Fidelity forms	Intervention; outcomes; component areas	Mean fidelity score across implemented components	Coach	Fidelity forms contain 10–14 items scored on a scale from 0 to 3
Teaming	Autism Program Evaluation Rating Scale (APERS)	Professional development; planning; intervention	APERS Teaming subdomain at posttest	APERS administrator	APERS item scores range from 1 to 5; Teaming subdomain is 7 items
STUDENT: *Key Questions: Is the CSESA process being implemented at the student level? Is the student receiving the interventions as designed?*					
Student-level planning	CSESA student planning documents	Planning	Mean of proportion of CSESA student plans completed each year	Artifact	Dynamic document completed with coach and key school staff
Student-level component dosage	Student intervention matrix	Intervention; component areas	Mean of individual student participation across CSESA components based on expected dosage	School staff	Start and stop dates for each component, converted into weeks of intervention per component

Note. CSESA = Center on Secondary Education for Students With Autism Spectrum Disorders.

schools. The profile and index scores can be used to address research and applied program evaluation questions. Although the CSESA implementation profile and index is conceptually multilevel, this term differs in meaning from the statistical multilevel designs that employ hierarchical linear modeling, a method used in educational efficacy research (Muthén, 1992).

The purpose of the current study is to demonstrate use of a procedure for capturing the implementation of a comprehensive treatment program for adolescents with ASD enrolled in public high school programs. The specific aims of current article are to (a) descriptively examine the range of scores for the features of the CSESA implementation index and the full implementation index score, (b) determine if the full CSESA implementation index differentiates between CSESA and services-as-usual (SAU) schools, and (c) determine if the "common features items," a subset of data that captured CSESA-like implementation at SAU schools, also differentiates between CSESA and SAU schools.

Method

This study of implementation assessment took place within a larger, multisite RCT that was approved by the institutional review board. Sixty high schools in three states were blocked

by district and randomly assigned in equal numbers to one of two groups: CSESA intervention or SAU. In each school, one or two school staff were designated as autism team (A-team) leaders. These individuals served as the key point of contact between CSESA researchers and the school. CSESA coaches (who were members of the research team) provided the training and coaching support in the schools. The CSESA coaches received in-depth training for the procedures for CSESA component trainings and coaching. The principal investigators assessed coaching fidelity for each CSESA coach.

Participants

A-team leaders recruited school personnel to be members of the team, which always included at least one special education teacher and one member of the school administration. A total of 579 school staff participated in the study (CSESA, $n = 392$; SAU, $n = 187$). Because of the implementation, the A-teams in the CSESA schools included a broader array of school personnel (e.g., special education teachers, related service providers, paraprofessionals). In SAU schools, the role of the A-team members focused on supporting the coordination of data collection, so A-teams were much smaller. On average, the school staff had 11.5 years of experience working with individuals with ASD and 12.8 years of experience in school settings. The school staff were mostly female (81%) and primarily White (90%) and non-Latino (96%).

Also after randomization, 547 students were recruited, consented, and enrolled in the study (CSESA, $n = 303$; SAU, $n = 244$), with a range of four to 12 students enrolled at each school. Although data were collected on a maximum of 12 students at each school, other students (with and without ASD) could receive CSESA interventions as part of the educational programming offered by the school. The students had a wide variety of characteristics, with nonverbal IQ as measured by the Leiter scale ranging from 30 to 141 ($M = 85.5$, $SD = 27.2$), Vineland Adaptive Behavior Composite standard scores ranging from 20 to 131 ($M = 75.7$, $SD = 16.7$), and autism severity as measured by Social Responsiveness Scale t scores ranging from 39 to 110 ($M = 70.4$, $SD = 12.2$). Additionally, the students were a racially and ethnically diverse sample, with 45.8% of the sample identifying as Hispanic or non-White. Initial analyses indicated that there were no significant differences between CSESA and SAU schools on these demographic variables.

CSESA Implementation Profile and Index

The CSESA researchers determined key features of the CSESA intervention model that reflected the three levels of implementation identified by Cordray et al. (2013): (1) delivery of training and preparation related to the CSESA model to the school by CSESA coaches (i.e., CSESA component training, CSESA coaching), (2) the delivery of the CSESA model by school staff to students (i.e., school-level planning, intervention quality, teaming), and (3) the reception of the CSESA model by students (i.e., student-level planning, student-level component dosage). Next, the CSESA researchers developed, revised, or determined measures and tools to assess each of the features. Several of the measures and tools were designed for use at both CSESA and SAU schools to document the CSESA-like features happening in SAU schools. Prior to the RCT, the researchers pilot-tested the measures and tools in the initial years of the project and revised as needed. In the current study, CSESA staff collected data for the final implementation index, including seven implementation profile features (see Table 1) in CSESA and SAU schools across the 2 years of active study involvement. The raw data for each feature of the implementation index data was converted to scaled scores (i.e., ratings of 0–3). Specifically, the researchers assigned a numerical value of 3 for ideal or complete implementation, assigned a value of 0 for poor or no implementation, and determined

Table 2. Scoring Metrics for the Scaled Scores of Features in the CSESA Implementation Index.

	0 *Limited or no implementation*	**1** *Fair implementation*	**2** *Good implementation*	**3** *Ideal implementation*
Feature				
SCHOOL				
CSESA component trainings	4 or fewer components	At least 5 components and at least 2 domains	At least 7 components and at least 3 domains	At least 9 components and all 4 domains
CSESA coaching	No coaching	1–108 hr	109–215 hr	216 hr or more
STAFF				
School-level planning	No indication of planning for CSESA components	—	—	Completed with indication of planning for CSESA components
Intervention quality	0–0.99	1.00–1.99	2.00–2.49	2.50–3.00
Teaming	1.00–2.00	2.01–3.00	3.01–4.00	4.01–5.00
STUDENT				
Student-level planning	0%–49%	50%–74%	75%–99%	100%
Student-level component dosage[a]	0–0.49	0.50–1.24	1.25–1.99	2.00–3.00

Note. CSESA = Center on Secondary Education for Students With Autism Spectrum Disorders.
[a]Process for calculating student-level component dosage: Student dosage was converted to scaled scores (0–3) per component based on expected dosage, scaled scores were used to calculate a student mean, then student mean scores were used to create a school mean. The academic components had different target student populations, so those scores were combined. The Social Competence Intervention was expected to be used only for a subset of students, so the dosage scaled score was calculated at the school level and combined into the school mean.

gradations of implementation between those for scaled scores of 1 and 2. Decisions about anchors for each item were based on the pilot studies and the conceptually ideal implementation of each implementation feature (see Table 2). In the subsequent sections, we describe the procedures for designing, collecting, and scoring each of the features of the CSESA profile and calculating the CSESA implementation index.

School: CSESA component training. Training logs that documented dates, length of training, and participation of "key personnel" (i.e., any school staff delivering the intervention component) were collected by CSESA staff at CSESA and SAU schools to track participation in the CSESA-delivered trainings. The raw data provided the number of components for which training was provided (out of 10) and the number of domains addressed by the trainings (out of four). A combination of these raw scores was used to determine the scaled score (see Table 2).

School: CSESA coaching. CSESA researchers developed a coaching log to document the coaching aspect of professional development. The coaching log included data on the number of staff members coached (individual, small group, large group), the roles of the staff members coached (e.g., special education teacher, psychologist, speech language pathologist), the modality of coaching (observation, discussion, action, check-in), the components and evidence-based practices coached, and the length of time coaching was provided (10 min or less to 91 min or more). CSESA coaches recorded data for each coaching session. For the implementation index, the total amount of coaching time delivered during the 2 years was summed and then converted to the scaled score (see Table 2 for details). No CSESA coaching occurred at SAU schools.

Staff: School-level planning. The CSESA researchers developed a school planning menu to track school-level planning for the CSESA model. On the basis of their local context, A-teams identified the order and time period (over the 2 years) for implementation of each CSESA component. The A-team and coach updated the school planning menu during the study, and the completed planning menu was turned in at the end of the study. The raw data reflected whether or not the school planning menu was completed. Because it was a dichotomous raw score, the scaled score was either a 0 or a 3. CSESA school planning menus were not completed at SAU schools.

Staff: Intervention quality. CSESA researchers developed fidelity measures with a relatively uniform format for components during the pilot years of the CSESA project. The fidelity measures assessed the adherence to component procedures originally identified and the quality of the delivery. The fidelity measures contained between 10 and 14 items grouped in five sections: (a) length and frequency of delivery of intervention session or practices as prescribed, (b) preparation and structure, (c) process, (d) strategies and content, and (e) evaluation and progress monitoring. CSESA researchers employed a 4-point rating scale (0= *not observed*, 1= *low fidelity*, 2= *mid fidelity*, and 3= *high fidelity*) with scoring guidelines for each of the levels of the rating. CSESA coaches collected fidelity on nine of the 10 interventions, with the exception of transition planning (i.e., time and logistics prevented observation of the planning process).

To collect fidelity data, the CSESA staff observed a school staff member implementing the CSESA intervention for at least 30 min or the full intervention session if shorter than 30 min. The CSESA staff member completed the fidelity form during the observation and later interviewed school staff if any items were not observable. The CSESA staff collected fidelity data three times during the initial semester of implementation and one time per semester for any subsequent semesters of implementation.

To collect comparable information at SAU schools, CSESA collected fidelity data during the middle of the 2nd year of study participation. First, the CSESA staff interviewed key school staff informants to determine if students at the school were receiving CSESA-like interventions. For example, for the peer networks component (a specific type of peer-mediated intervention), the probe was "Are any of the students enrolled in the study receiving support interacting with peers in non-classroom-based settings (e.g., lunch club/buddies, sports teams, clubs)?" If the key informants answered yes, then the CSESA staff probed further to identify all enrolled students who were receiving that intervention. Next, the CSESA staff scheduled fidelity observations for the CSESA-like interventions that were currently being implemented at the school for one student each from the diploma and nondiploma groups.

For scoring, the CSESA researchers calculated a mean item rating for each fidelity observation. CSESA staff then calculated, by school, the mean of the single fidelity observation scores for each intervention component (e.g., mean of all the fidelity observations for Alternate Achievement Literary, Collaborative Strategic Reading, peer support intervention). The fidelity scores were then converted to the CSESA implementation index scaled scores (see Table 2 for details).

Staff: Teaming. As part of the larger RCT study, CSESA researchers (i.e., not the CSESA coach assigned to the school) collected the Autism Program Environment Rating Scale–High School (APERS; Odom et al., 2018) for all schools at the beginning of the 1st year and end of the 2nd year. The APERS is a 66-item measure, organized by 11 domains (e.g., class structure, instruction, families, teaming) that utilizes a 5-point Likert-type scale for items. Raters based their ratings on observations across a wide array of school settings (academic and nonacademic, general education and special education), structured interviews with school staff and parents, and document analysis (review IEPs and transition plans). The strong evidence of APERS reliability and validity has been previously reported (Odom et al., 2018). The teaming domain score at posttest, which consists of seven

items that address team training, membership, communication, collaboration, decision making, and delivery of services, was used for the implementation index. The mean item rating was used to assign a scaled score.

Student: Student-level planning. CSESA researchers developed a student planning menu for school staff to use in collaboration with coaches to develop individualized intervention plans for students. The student planning menu included selection of the CSESA components appropriate for addressing student goals, identifying when and how they would be implemented, and identifying how outcomes would be assessed. The key school staff and coach updated the student planning menus periodically during the 2 years of the study, and completed student planning menus were turned in at the end of each year of the study. The raw score for the student planning menus was the mean of the proportion of student planning menus completed each year. The raw score was then converted to the scaled score (see Table 2 for details).

Student: Student-level component dosage. The CSESA researchers created a student intervention matrix to assess individual students' reception of (i.e., engagement in) CSESA components. CSESA personnel documented the start and stop dates for each CSESA intervention for each enrolled student. The data were collected across the semester during coaching sessions of CSESA interventions. Once all data were collected, the researchers calculated the number of weeks that a given intervention was implemented with each student. Adjustments were made to the number of weeks to account for longer school breaks (e.g., subtracting 1 week when the dates occurred in spring semester to account for spring break). The raw data were then converted to scaled scores at the student level, which were combined to calculate a scaled score at the school level (see note under Table 2 for additional detail).

CSESA scale scores. The raw scores for each implementation feature were converted into scaled scores, based on a 4-point Likert-type scale (i.e., 0 = *limited or no implementation*, 1 = *fair implementation*, 2 = *good implementation*, 3 = *ideal implementation*). Specific criteria for each of the ratings appear in Table 2 and were based on information collected in the pilot tests in the first 2 years and, as noted, ideal implementation. For example, for the coaching feature, it was estimated that coaches should spend about 4 hr per week coaching across three semesters of active implementation, so the ideal implementation was set at 216 hr (4 hr per week $\times$ 18 weeks per semester $\times$ 3 semesters). For the student-level component dosage feature, scaled scores were calculated first at the intervention component level and then combined to create school means. For example, work-based learning experiences (WBLE) were expected to be implemented for 12 weeks, so students received a WBLE scaled score of 3 for 12 weeks, 2 for 8 to 11 weeks, 1 for 6 to 7 weeks, and 0 for fewer than 6 weeks.

CSESA implementation index score. To calculate the total implementation index score, the researchers first weighted two of the seven profile features that were deemed to be more critical for implementation: intervention quality and student-level component dosage. These features had a weight of 4 (i.e., reflecting the four domains of the CSESA model). The remaining features (CSESA component training, CSESA coaching, teaming, and school- and student-level planning) were less intensive albeit still important aspects of the CSESA model and thus had a weight of 1. The weighting decisions were made in consultation with methodologists and statisticians as well as information collected from the first 2 years of pilot studies. The scaled scores for each feature were multiplied by their weight and divided by 13 (reflecting the total weighting of five features with a weight of 1 and two features with a weight of 4) to determine the total implementation index score on a scale from 0 to 3.

Common features. In research studies, it is important to determine as much as possible the interventions occurring in the control

condition (i.e., the SAU schools). Although the full implementation index provides an accurate picture of the breadth of intervention features in the SAU schools, it could provide a "depressed" score for schools because, by design, the CSESA staff did not provide features of the program (e.g., training), resulting in a 0 rating for SAU schools. To provide a closer examination of conditions in the SAU schools, CSESA staff examined interventions and program features in the SAU schools that were the same as or similar to the CSESA component features. For example, SAU schools may offer educational programming that is similar to a CSESA intervention (e.g., peer buddy program) even though they were not trained on the exact CSESA component (peer networks). CSESA staff collected information on the number of common interventions occurring in SAU schools, intervention fidelity when a common feature was implemented in the SAU school, and the APERS teaming score.

Data Analysis

Data analysis was completed with SPSS Statistics Version 25. To look at the variability within the CSESA group, we reported descriptive statistics for individual CSESA schools on the full implementation index and the features of the implementation index. We also examined the relationship between features and the full implementation index using correlations. To examine the validity of the full implementation index and look at differences between "common features," we used independent-sample *t* tests to examine group differences between CSESA and SAU schools.

Results

The three research aims for the current study were (a) to examine the range of scores for the CSESA implementation index profile and full implementation index score, (b) to determine if the full CSESA implementation index differentiates between CSESA and SAU schools, and (c) to determine if the "common features items," a subset of data that captured CSESA-like implementation at SAU schools, also differentiates between CSESA and SAU schools.

Range of Scores on the Implementation Profile and Index

For the total implementation index score in the CSESA intervention schools, four schools achieved a score of 2.50 or higher, 12 schools achieved a score between 2.00 and 2.49, 13 schools achieved a score between 1.50 and 1.99, and one school scored under 1.50, with a total range of scores from 1.31 to 2.87 (see Supplementary Materials in the online version of the article for additional data). Overall, there were high levels of good or ideal implementation in many high schools, but there was variability for almost all features with the exception that all CSESA intervention schools scored a 3 for school-level planning. Good or ideal implementation was achieved by 29 schools for intervention quality (fidelity), 26 schools for student-level planning, 23 schools for teaming, and 22 schools for CSESA component training. The scores were lower for CSESA coaching and student-level component dosage, with 17 of 30 schools achieving good or ideal implementation.

Although all features of the implementation index contributed conceptually to the overall measure and provided valuable information about individual variability among schools, there was variability in the degree to which they were statistically associated with the total implementation index. To examine the relationship between the individual profile feature ratings and the total implementation index score, we calculated Pearson product moment correlations. The correlation coefficients were significant for student dosage (.91, $p < .01$), intervention quality (.69, $p < .01$), training (.65, $p < .01$), and to a lesser extent, coaching (.34, $p < .05$). They were not significant for teaming (.25, *ns*) and student planning (.11, *ns*). The association for school planning profile feature and the implementation index metric was not run because there was no variability in the school planning rating (i.e., all had ratings of 3).

Difference Between CSESA and SAU Schools on Implementation Index

The mean implementation index for CSESA schools was 2.07 ($SD = 0.38$) compared to a mean of 0.47 ($SD = 0.17$) for SAU schools. For SAU schools the ratings were typically low across many of the implementation features, with limited variability. The difference between the implementation index total scores for CSESA schools compared to SAU schools was significant with a large effect size ($t = 28.13$, $p < .001$; $d_{Cohen} = 5.43$; confidence interval [CI] = [4.35, 6.53]).

Difference Between CSESA and SAU Schools on Common Features

The common features compared between the CSESA and SAU schools were teaming, intervention quality, and proportion of CSESA-like component areas addressed. There were significant differences between CSESA and SAU schools for these three common features. On average, SAU schools reported providing about half of the CSESA-like interventions, compared to CSESA schools, which provided just over two thirds of CSESA interventions on average ($t = 3.65$, $p = .001$), with a Cohen's d of 0.95 (CI = [0.42, 1.49]). The mean intervention quality was 0.68 points higher (on a 0-to-3 scale) in the CSESA schools ($t = 9.49$, $p < .001$), with a Cohen's d of 2.44 (CI = [1.77, 3.11]). The mean of the Teaming subscale from the APERS was 0.36 points higher (on a 1-to-5 scale) in the CSESA schools ($t = 2.65$, $p = .01$), with a Cohen's d of 0.77 (CI = [0.18, 1.23]).

Discussion

Measurement of program implementation has several purposes in research studies and a more practical purpose in school-based adoption of comprehensive treatment programs for youth with ASD. Its primary purpose is to document the degree to which features of the program have been implemented as planned. In this study, we demonstrated the use of an index approach to assess individual features of implementation that then can be used to compute an overall implementation index score. We propose that this multilevel profile and index assessment approach (based on Cordray et al., 2013) is appropriate for comprehensive programs implemented with, or adopted by, school districts where teachers and school staff will deliver the program. This approach extends the concept and assessment of implementation in special and general education beyond only measurement of treatment fidelity or amount of treatment delivered (O'Donnell, 2008).

The data from this study reflected variation in implementation among schools in the CSESA efficacy study. Although the mean implementation index score was in the *good* range (i.e., 2.07 out of 3.0), individual school total index scores varied from 1.31 to 2.85, with 18 of the 30 schools having scores above 2.0. Alternately, this means that 12 schools fell into the *fair* or lower implementation range. To some extent, this variability in implementation reflects the challenges of working in public school settings and supporting local service providers delivering the intervention. Implementers of interventions in schools (Sylvan, 2018) and other disciplines (Hamilton et al., 2018) have also noted the variability of implementation in community-based settings.

To some extent, this variability in implementation reflects the challenges of working in public school settings and supporting local service providers delivering the intervention.

There are several advantages of the profile approach. It allowed examination of individual features of implementation. Five of the seven individual features had mean ratings between 2.00 and 3.00. Coaching (discussed in a subsequent paragraph) and student dosage (1.67) were at a lower level than anticipated. In addition, the association of implementation profile features with the

total index score also allowed authors to examine empirically profile features. As noted, the index scores were conceptualized and calibrated through 2 years of pilot testing and in consultation with national experts in the field. Intervention fidelity and student dosage were most strongly associated with overall index score, which would be expected because they were weighted more heavily than other variables. Training was strongly associated with the overall index score, but surprisingly, teaming, as generated by the APERS teaming domain, was not. In other analyses, investigators found that APERS teaming scores increased across time and were significantly different from teaming scores in the SAU schools (Odom, Hume, DaWalt-Smith, Hall, & Kraemer, 2018). It could well be that the specific items from the APERS teaming domain did not map well onto the specific features of the A-team process that was central to the CSESA model. Similarly, student planning was not significantly associated with the total index score. Examining circumstances within these schools and perhaps removing these scores as outliers, if circumstances dictated, or using a quality rating approach to measure student-level planning could have resulted in a stronger relationship between planning and total index score.

These findings in this study also reflect that coaching is a complicated variable. The coaching profile score correlated significantly but modestly with the total index score and also fell below the level planned for implementation. This could be surprising given the importance of coaching often discussed in the literature (Cusumano, Preston, & Ward, 2018; Hershfeldt, Pell, Sechrest, Pas, & Bradshaw, 2012). In fact, it may well be that coaching is one of the nonlinear relationships to overall implementation that McGaghie (2011) has noted as a characteristic of comprehensive programs in health care services. A linear relationship would suggest that higher amounts of coaching would be associated with higher fidelity scores. This may be true for some schools. However, for "low-implementing" schools that may have been experienced challenges (e.g., staff turnover), more coaching may have been required to support even their low level of implementation. Conversely, for "high-implementing" schools, the CSESA staff may not have needed to provide as much coaching. Regardless, coaching is a complex process that deserves much greater scrutiny in future implementation studies.

In addition to documenting implementation, researchers also employ implementation assessment to determine that the treatment delivered in the intervention group (CSESA) was different from the conditions occurring in the control group (SAU; i.e., there is not contamination of treatment conditions; Keogh-Brown et al., 2007; Magill et al., 2018). The implementation index score did document, to a significant degree and with a high effect size, that the specific features of CSESA were not present in the SAU schools (i.e., little to no contamination). Also, from an instrument validity perspective, these significant differences establish the criterion-related validity of the assessment approach (Nunnally & Bernstein, 1994).

It was also possible, however, that SAU schools might have been using intervention components similar to CSESA components. For example, CSESA schools used the Social Competence Intervention to promote social competence, and an SAU school could have used the PEERS curriculum (Laugeson, 2014), another evidence-based intervention targeting social competence. An analysis of common intervention features (e.g., a social skills training program) revealed that such features were employed significantly less frequently and at a lower quality (i.e., fidelity) in the SAU schools. Researchers have emphasized the importance of documenting conditions occurring both in the treatment and control conditions of randomized control studies (Gersten et al., 2005), and the CSESA implementation index approach allowed for this comparison to occur in the larger RCT efficacy study context.

Researchers have emphasized the importance of documenting conditions occurring both in the treatment and control conditions of randomized control studies (Gersten et al., 2005), and the CSESA implementation index approach allowed for this comparison to occur in the larger RCT efficacy study context.

We did not examine the association between the degree of CSESA implementation and student outcomes. Dunst, Trivette, and Raab (2013) and others (e.g., Peters, Adam, Alonge, Agyepong, & Tran, 2013) have noted that research on implementation measurement and processes is a distinctly different from examinations of the association of implementation with student outcomes. The current study was clearly in the former camp. A next step in this program of research will be to analyze whether CSESA implementation moderates treatment effects for students and families.

With the movement to an "open science" perspective in education sciences (Cook, Lloyd, Mellor, Nosek, & Therrien 2018), the role of replication is critical. McBee, Makel, Peters, and Matthews (2018), quoting Simmons (2016), noted that "a scientist's #1 job is to differentiate what is true/replicable from what is not.... Replicability is not merely *a* consideration, but *the most important* consideration" (p. 374). A critical dimension of replication is documenting the degree to which the essential features of the intervention were employed in the replication study. For research on comprehensive treatment programs in special education delivered by school staff in a public school setting, assessment of implementation is the cornerstone of replication. We propose that it must extend beyond reporting only treatment fidelity and also include procedural features at the school, staff, and student levels. As noted, this level of implementation assessment detail is rarely applied in examinations of comprehensive programs in school settings, and the authors propose that this article contributes to the literature by providing just such a demonstration.

In addition to its scientific function, a multilevel implementation assessment approach as proposed by Cordray and colleagues (Cordray et al., 2013; Nelson et al., 2010), and exemplified in this study, has practical relevance. Such an assessment approach initially requires the purveyors or developers of the comprehensive treatment program to specify the activities leading to implementation at the school, staff, and student levels. This specification allows school system leaders to plan for activities required for implementation and use the data generated from this implementation assessment process in a formative way to monitor the degree to which implementation is occurring as planned. Sugai and Horner (in press) provide an example of how this happens for school adoptions of positive behavior interventions and supports, and other implementation scientists (Aarons, Hurlburt, & Horwitz, 2011; Fixsen, Blase, Metz, & Van Dyke, 2013) discuss the need for a continuous data feedback loop that provides such continuous monitoring of implementation. The CSESA implementation index provides an example of a process that could potentially have relevance in public school contexts.

In addition to its scientific function, a multilevel implementation assessment approach as proposed by Cordray and colleagues (Cordray et al., 2013; Nelson et al., 2010), and exemplified in this study, has practical relevance.

Limitations and Future Research

There are several limitations to note as well as next steps to advance this line of research and practice. First, the various informants across implementation data sources were not naive to study condition. Blind raters were not used, in part because the index requires a thorough understanding of the CSESA activities and intervention components. It is unlikely that a

comprehensive assessment of implementation could or should be expected to be completed by research staff naive to the intervention status; however, future research studies might incorporate the use of naive raters for portions of the index, such as fidelity assessments. Additionally, because of logistic constraints, the researchers did not collect interrater agreement on the implementation index data. Ideally, such checks would have been helpful. Last, the practical utility of a multilevel assessment would be important to determine. For example, if a school system decided to adopt the CSESA or other school-based comprehensive treatment programs, the feasibility of collecting the data identified in the current assessment approach would be important to examine.

Several future directions for this line of work have been identified, including examining the relationship between the degree of implementation and student outcomes and better understanding associations between various implementation features. Similarly, examining the relationship between demographic variables of the school, such as the geographic location (e.g., urban, suburban, rural) or socioeconomic status (e.g., rates of free and reduced lunch), or of the staff (e.g., student–teacher ratio, average years of experience); the implementation features; and overall index score may be helpful in examining readiness and factors that facilitate or inhibit implementation. Finally, conducting a factor analysis to more broadly examine the utility of the index and to potentially eliminate redundancy would be beneficial.

Conclusion

In addressing large issues in special education, experiences from other disciplines can be informative. In the health care services field, Hawe (2015) worked to provide health care to reduce malaria in Uganda. She proposed that for large and persistent societal problems, narrowly focused intervention approaches are not likely to be "disruptive" enough to change entrenched practices that previously led to poor outcomes. To affect such poor outcomes, comprehensive intervention approaches may be necessary, along with a systems-focused foundation to support them. The poor life outcomes for many individuals with ASD leaving the public school system exemplify just such a societal issue. The complexity of the factors that lead to those outcomes requires a comprehensive intervention approach, and implicit in that process is the need to assess implementation. The CSESA implementation profile and index in this article provides one such model for assessing the multiple features of implementation of a school-based comprehensive treatment program.

References

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 4–23. http://dx.doi.org/10.1007/s10488-010-0327-7

Artman-Meeker, K. M., Hemmeter, M. L., & Snyder, P. (2014). Effects of distance coaching on teachers' use of pyramid model practices: A pilot study. *Infants & Young Children*, *27*, 325–344. http://dx.doi.org/10.1097/IYC.0000000000000016

Bishop, D. C., Pankratz, M. M., Hansen, W. B., Albritton, J., Albritton, L., & Strack, J. (2014). Measuring fidelity and adaptation: Reliability of an instrument for school-based prevention programs. *Evaluation & the Health Professions*, *37*, 231–257. http://dx.doi.org/10.1177/0163278713476882

Brookman-Frazee, L. I., Drahota, A., & Stadnick, N. (2012). Training community mental health therapists to deliver a package of evidence-based practice strategies for school-age children with autism spectrum disorders: A pilot study. *Journal of Autism and Developmental Disorders*, *42*, 1651–1661. http://dx.doi.org/10.1007/s10803-011-1406-7

Browder, D., Ahlgrim-Delzell, L., Flowers, C., & Baker, J. (2012). An evaluation of a multicomponent early literacy program for students with severe developmental disabilities. *Remedial and Special Education*, *33*, 237–246. http://dx.doi.org/10.1177/0741932510387305

Carter, E. W., Common, E. A., Sreckovic, M. A., Huber, H. B., Bottema-Beutel, K., Gustafson,

J. R., . . . Hume, K. (2014). Promoting social competence and peer relationships for adolescents with autism spectrum disorders. *Remedial and Special Education*, *35*, 91–101. https://doi.org/10.1177/0741932513514618

Carter, E. W., Gustafson, J. R., Sreckovic, M. A., Dykstra Steinbrenner, J. R., Pierce, N. P., Bord, A., . . . Mullins, T. (2017). Efficacy of peer support interventions in general education classrooms for high school students with autism spectrum disorder. *Remedial and Special Education*, *38*, 207–221. https://doi.org/10.1177/0741932516672067

Caskey, M., & Anfara, V. A., Jr. (2014). *Developmental characteristics of young adolescents*. Retrieved from https://www.amle.org/BrowsebyTopic/WhatsNew/WNDet/TabId/270/ArtMID/888/ArticleID/455/Developmental-Characteristics-of-Young-Adolescents.aspx

Colombi, C., Narzisi, A., Ruta, L., Cigala, V., Gagliano, A., Pioggia, G., . . . Prima Peitra Team. (2018). Implementation of the Early Start Denver Model in an Italian community. *Autism*, *22*, 126–133. https://doi.org/10.1177/1362361316665792

Cook, B. G., Lloyd, J. W., Mellor, D., Nosek, B. A., & Therrien, W. J. (2018). Promoting open science to increase the trustworthiness of evidence in special education. *Exceptional Children*, *85*, 104–118. https://doi.org/10.1177/0014402918793138

Cordray, D. S., Pion, G. M., Brandt, C., & Molefe, A. (2013). *The impact of the Measures of Academic Progress (MAP) program on student reading achievement*. Paper presented at the spring conference of the Society for Research on Educational Effectiveness, Washington, DC. Retrieved from https://files.eric.ed.gov/fulltext/ED564093.pdf

Cusumano, D. L., Preston, A. I., & Ward, C. S. (2018). *Coaching for competence and confidence*. Retrieved from Global Implementation Initiative web site: https://globalimplementation.org/resources/coaching-for-competence-and-confidence/

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*, 23–45. https://doi.org/10.1016/S0272-7358(97)00043-3

DaWalt, L. S., Greenberg, J. S., & Mailick, M. R. (2018). Transitioning Together: A multi-family group psychoeducation program for adolescents with ASD and their parents. *Journal of Autism and Developmental Disorders*, *48*, 251–263. http://dx.doi.org/10.1007/s10803-017-3307-x

Dawson, G., & Rogers, S. J. (2011). *Early Start Denver Model for young children with autism: Promoting language, learning and engagement*. New York, NY: Guilford Press.

Dunst, C. J., Trivette, C. M., & Raab, M. (2013). An implementation science framework for conceptualizing and operationalizing fidelity in early childhood intervention studies. *Journal of Early Intervention*, *35*, 85–101. https://doi.org/10.1177/1053815113502235

Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. *Implementation Science*, *1*, 1–3. https://doi.org/10.1186/1748-5908-1-1

Fixsen, D. L., Blase, K., Metz, A., & Van Dyke, M. (2013). Statewide implementation of evidence-based practice programs. *Exceptional Children*, *79*, 213–230. https://doi.org/10.1177/001440291307900206

Fleury, V. P., Hedges, S., Hume, K., Browder, D. M., Thompson, J. L., Fallin, K., . . . Vaughn, S. (2014). Addressing the academic needs of adolescents with autism spectrum disorder in secondary education. *Remedial and Special Education*, *35*, 68–79. https://doi.org/10.1177/0741932513518823

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*, 149–164. https://doi.org/10.1177/001440290507100202

Gresham, F. A., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis*, *26*, 257–263. https://doi.org/10.1901/jaba.1993.26-257

Hamilton, A. B., Mittman, B. S., Campbell, D., Hutchinson, C., Liu, H., Moss, N. J., & Wyatt, G. E. (2018). Understanding the impact of external context on community-based implementation of an evidence-based HIV risk reduction intervention. *BMC Health Services Research*, *18*(1), 11. http://dx.doi.org/10.1901/jaba.1993.26-257

Hawe, P. (2015). Minimal, negligible and negligent interventions. *Social Science & Medicine*, *138*, 265–268. https://doi.org/10.1016/j.socscimed.2015.05.025

Hershfeldt, P. A., Pell, K., Sechrest, R., Pas, E. T., & Bradshaw, C. P. (2012). Lessons learned coaching teachers in behavior management: The PBIS plus coaching model. *Journal of Educational and Psychological Consultation*, *22*, 280–299. http://dx.doi.org/10.1080/10474412.2012.731293

Hume, K., Boyd, B. A., Hamm, J. V., & Kucharczyk, S. (2014). Supporting independence in adolescents on the autism spectrum. *Remedial and Special Education*, *35*, 102–113. https://doi.org/10.1177/0741932513514617

Hume, K., Dykstra Steinbrenner, J., Sideris, J., Smith, L., Kucharczyk, S., & Szidon, K. (2018). Multi-informant assessment of transition-related skills and skill importance in adolescents with autism spectrum disorder. *Autism*, *22*, 40–50. https://doi.org/10.1177/1362361317722029

Hume, K., & Odom, S. L. (2019). *A school-based treatment model for adolescents across the autism spectrum*. Manuscript submitted for publication.

Kasari, C. (2015). Update on behavioral interventions for autism and developmental disabilities. *Current Opinion in Neurology*, *28*, 124–129. http://dx.doi.org/10.1097/WCO.0000000000000185

Keogh-Brown, M. R., Bachmann, M. O., Shepstone, L., Hewitt, C., Howe, A., Ramsay, C. R., . . . Elbourne, D. R. (2007). Contamination in trials of educational interventions. *Health Technology Assessment*, *11*(43), 3. https://doi.org/10.3310/hta11430

Knight, V. F., Huber, H. B., Kuntz, E. M., Carter, E. W., & Juarez, A. P. (2019). Instructional practices, priorities, and preparedness for educating students with autism and intellectual disability. *Focus on Autism and Other Developmental Disabilities*, *34*, 3–14. https://doi.org/10.1177/1088357618755694

Koegel, R. L., & Koegel, L. K. (2006). *Pivotal response treatments for autism: Communication, social, and academic development*. Baltimore, MD: Brookes.

Kucharczyk, S., Reutebuch, C. K., Carter, E. W., Hedges, S., El Zein, F., Fan, H., & Gustafson, J. R. (2015). Addressing the needs of adolescents with autism spectrum disorder: Considerations and complexities for high school interventions. *Exceptional Children*, *81*, 329–349. http://dx.doi.org/10.1177/0014402914563703

Laugeson, E. A. (2014). *The PEERS curriculum for school-based professionals: Social skills training for adolescents with autism spectrum disorder*. New York, NY: Routledge.

Lovaas, O. I. (2003). *Teaching individuals with developmental delays: Basic intervention techniques*. Austin, TX: Pro-Ed.

Low, S., Van Ryzin, M. J., Brown, E. C., Smith, B. H., & Haggerty, K. P. (2014). Engagement matters: Lessons from assessing classroom implementation of steps to respect: A bullying prevention program over a one-year period. *Prevention Science*, *15*, 165–176. http://dx.doi.org/10.1007/s11121-012-0359-1

Magill, N., Graves, H., de Zoysa, N., Winkley, K., Amiel, S., Shuttlewood, E., . . . Ismail, K. (2018). Assessing treatment fidelity and contamination in a cluster randomised controlled trial of motivational interviewing and cognitive behavioural therapy skills in type 2 diabetes. *BMC family practice*, *19*, 60. http://dx.doi.org/10.1186/s12875-018-0742-5

McBee, M. T., Makel, M. C., Peters, S. J., & Matthews, M. S. (2018). A call for open science in giftedness research. *Gifted Child Quarterly*, *62*, 374–388. https://doi.org/10.1177/0016986218784178

McGaghie, W. C. (2011). Implementation science: Addressing complexity in medical education. *Medical Teacher*, *33*, 97–98. http://dx.doi.org/10.3109/0142159X.2011.550971

Muthén, B. O. (1992). Latent variable modeling in epidemiology. *Alcohol Research*, *16*, 286.

National Research Council. (2001). *Educating children with autism*. Retrieved from https://infiniteach.com/wp-content/uploads/2015/07/National-Research-Study-.pdf

Nelson, M. C., Cordray, D. S., Hulleman, C. S., Darrow, C. L., & Sommer, E. C. (2010). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *Journal of Behavioral Health Services & Research*, *39*, 374–396. http://dx.doi.org/10.1007/s11414-012-9295-x

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (McGraw-Hill Series in Psychology, Vol. 3). New York, NY: McGraw-Hill.

Odom, S. L., Boyd, B. A., Hall, L. J., & Hume, K. (2010). Evaluation of comprehensive treatment models for individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *40*, 425–436. http://dx.doi.org/10.1007/s10803-009-0825-1

Odom, S. L., Collet-Klingenberg, L., Rogers, S. J., & Hatton, D. D. (2010). Evidence-based practices in interventions for children and youth

with autism spectrum disorders. *Preventing School Failure*, *54*, 275–282. http://dx.doi.org/10.1080/10459881003785506

Odom, S. L., Cox, A., Sideris, J., Hume, K. A., Hedges, S., Kucharczyk, S., . . . Neitzel, J. (2018). Assessing quality of program environments for children and youth with autism: Autism Program Environment Rating Scale (APERS). *Journal of Autism and Developmental Disorders*, *48*, 913–924. http://dx.doi.org/10.1007/s10803-017-3379-7

Odom, S. L., Duda, M. A., Kucharczyk, S., Cox, A. W., & Stabel, A. (2014). Applying an implementation science framework for adoption of a comprehensive program for high school students with autism spectrum disorder. *Remedial and Special Education*, *35*, 123–132. https://doi.org/10.1177/0741932513519826

Odom, S. L., Hume, K., DaWalt-Smith, L., Hall, L. J., & Kraemer, B. (2018, May). *Efficacy of a school-based treatment: Models for adolescents with autism. Effects on school quality*. Paper presented at the annual meeting of the International Society for Autism Research, Rotterdam, Netherlands. Retrieved from https://insar.confex.com/insar/2018/webprogram/Paper26991.html

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, *78*, 33–84. https://doi.org/10.3102/0034654307313793

Peters, D. H., Adam, T., Alonge, O., Agyepong, I. A., & Tran, N. (2013). Implementation research: What it is and how to do it. *British Journal of Sports Medicine*, *347*, f6753. https://doi.org/10.1136/bmj.f6753

Reutebuch, C. K., El Zein, F., Kim, M. K., Weinberg, A. N., & Vaughn, S. (2015). Investigating a reading comprehension intervention for high school students with autism spectrum disorder: A pilot study. *Research in Autism Spectrum Disorders*, *9*, 96–111. https://doi.org/10.1016/j.rasd.2014.10.002

Rogers, S. J., & Dawson, G. (2010). *Early Start Denver Model for young children with autism: Promoting language, learning, and engagement*. New York, NY: Guilford Press.

Sanetti, L. M. H., Williamson, K. M., Long, A. C. J., & Kratochwill, T. R. (2018). Increasing in-service teacher implementation of classroom management practices through consultation, implementation planning, and participant modeling. *Journal of Positive Behavior Interventions*, *20*, 43–59. https://doi.org/10.1177/1098300717722357

Shattuck, P. T., Narendorf, S. C., Cooper, B., Sterzing, P. R., Wagner, M., & Taylor, J. L. (2012). Postsecondary education and employment among youth with an autism spectrum disorder. *Pediatrics*, *129*, 1042–1049. http://dx.doi.org/10.1542/peds.2011-2864

Simmons, J. P. (2016, September 30). What I want our field to prioritize [Web log post]. Retrieved from http://datacolada.org/53/

Stahmer, A. C., Suhrheinrich, J., Reed, S., & Schreibman, L. (2012). What works for you? Using teacher feedback to inform adaptations of pivotal response training for classroom use. *Autism Research and Treatment*, *2012*, 1–11. http://dx.doi.org/10.1155/2012/709861

Stahmer, A. C., Suhrheinrich, J., Schetter, P. L., & Hassrick, E. M. (2018). Exploring multi-level system factors facilitating educator training and implementation of evidence-based practices (EBP): A study protocol. *Implementation Science*, *13*, 3. http://dx.doi.org/10.1186/s13012-017-0698-1

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. https://doi.org/10.1002/pits.21928

Stock, R., Mirenda, P., & Smith, I. B. (2013). Comparison of community-based verbal behavior and pivotal response treatment programs for young children with autism spectrum disorder. Research in Autism Spectrum Disorders, *7*, 1168–1181. https://doi.org/10.1016/j.rasd.2013.06.002

Strain, P. S., & Bovey, E. H. (2011). Randomized, controlled trial of the LEAP model of early intervention for young children with autism spectrum disorders. *Topics in Early Childhood Special Education*, *31*, 133–154. https://doi.org/10.1177/0271121411408740

Sugai, G., & Horner, R. (2020). Sustaining and scaling positive behavioral interventions and supports: Implementation drivers, outcomes, and considerations. *Exceptional Children*, *86*, 120–136.

Sutherland, K. S., McLeod, B. D., Conroy, M. A., & Cox, J. R. (2013). Measuring implementation of evidence-based programs targeting young children at risk for emotional/behavioral disorders. *Journal of Early Intervention*, *35*, 129–149. https://doi.org/10.1177/1053815113515025

Sylvan, L. (2018). Tiers to communication success: How can SLPs join in the MTSS

framework many schools are adopting to catch students' special education needs earlier and provide levels of intervention? *ASHA Leader*, *23*(8), 44–53. https://doi.org/10.1044/leader.FTR1.23082018.44

Test, D. W., Smith, L. E., & Carter, E. W. (2014). Equipping youth with autism spectrum disorders for adulthood: Promoting rigor, relevance, and relationships. *Remedial and Special Education*, *35*, 80–90. https://doi.org/10.1177/0741932513514857

U.S. Department of Education. (2017). *39th annual report to Congress on the Implementation of the Individuals with Disabilities Education Act*. Retrieved from https://www2.ed.gov/about/reports/annual/osep/2017/parts-b-c/39th-arc-for-idea.pdf

Volkmar, F. R., Reichow, B., Westphal, A., & Mandell, D. S. (2014). Autism and the autism spectrum: Diagnostic concepts. In F. R. Volkmar, R. Paul, K. A. Pelphrey, & S. J. Rogers (Eds.), *Handbook of autism and pervasive developmental disorders* (4th ed., pp. 3–27). Hoboken, NJ: Wiley. https://doi.org/10.1002/9781118911389.hautc01

Wolery, M. (2011). Intervention research: The importance of fidelity measurement. *Topics in Early Childhood Special Education*, *31*, 155–157. https://doi.org/10.1177/0271121411408621

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., . . . Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of autism and developmental disorders*, *45*(7), 1951–1966.

Wong, V., Ruble, L. A., McGrew, J. H., & Yu, Y. (2018). An empirical study of multidimensional fidelity of COMPASS consultation. *School Psychology Quarterly*, *33*, 251–263. http://dx.doi.org/10.1037/spq0000217

Authors' Note

The work reported here was supported by the Institute of Education Sciences, U.S. Department of Education (Grant R324C120006) awarded to University of North Carolina–Chapel Hill. The opinions expressed represent those of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education. Thank you to the many schools, staff members, adolescents on the autism spectrum, and their families who made this study possible. Thanks also to the Center on Secondary Education for Students With Autism Spectrum Disorders (CSESA) research team working across three universities, which includes the authors and other key CSESA investigators: Leann Smith DaWalt, Bonnie Kraemer, Kate Szidon, and Christopher Brum.

Supplemental Materials

Supplementary materials are available in the online version of the article.

Manuscript received November 2018; accepted March 2019.