A Waitlist Randomized Implementation Trial of Classroom Pivotal Response Teaching for Students With Autism

Aubyn C. Stahmer, PhD[1,2], Jessica Suhrheinrich, PhD[2,3], Sarah R. Rieth, PhD[2,3], Scott Roesch, PhD[2,3], Sarah Vejnoska, PhD[1,2], Janice Chan, MA, BCBA[2,3], Allison Nahmias, PhD[4] [iD], and Tiffany Wang, MA[2,5]

Abstract

Classroom Pivotal Response Teaching (CPRT) is a community-partnered adaptation of a naturalistic developmental behavioral intervention identified as an evidence-based practice for autistic children. The current study evaluated student outcomes in a randomized, wait-list controlled implementation trial across classrooms. Participants included teachers ($n = 126$) and students with autism ($n = 308$). Teachers participated in 12 hours of didactic, interactive training and additional in-classroom coaching. Generalized Estimating Equations accounted for clustering. Adjusted models evaluated the relative effects of training group, CPRT fidelity, and classroom quality on student outcomes. Results indicate higher CPRT fidelity was associated with greater increases in student learning. Having received CPRT training predicted increased student engagement and greater decreases in reported approach/withdrawal problems. These differences may be linked to the theoretical foundations of CPRT of increasing student motivation and engagement and collaborative adaptation to increase feasibility in schools. Overall, results suggest CPRT may be a beneficial approach for supporting autistic students.

Keywords

autism, school age, classroom, naturalistic intervention

Autism is a neurodevelopmental condition characterized by differences in social communication along with restricted, repetitive, and stereotyped patterns of behavior (American Psychiatric Association, 2013). The pervasive nature of the unique learning needs associated with autism, in conjunction with increasing rates of autism identification, presents public school systems with significant challenges in providing intensive, individualized programming. Delivery of high-quality educational services requires contextual supports, skilled providers, and effective use of evidence-based autism intervention and instructional strategies. A lack of evidence-based autism interventions designed specifically for use in school programs further complicates the process of educating children with autism (Kasari & Smith, 2013; McGee & Morrier, 2005; Stahmer, 2007).

Systematic reviews have identified evidence-based practices (EBPs) for autistic individuals (Steinbrenner et al., 2020; C. Wong et al., 2015); however, most practices were designed for, and tested in, one-on-one or highly controlled settings. Teachers attempting to use these programs in classrooms report barriers related to staffing, training, and the fit of the model for their setting and a broad range of students with heterogeneous learning needs (Stahmer et al., 2005, 2012; Suhrheinrich et al., 2021; Wilson & Landa, 2019). Only a few research trials examining autism EBPs have

taken place in the school context (Odom et al., 2022), and randomized trials of EBPs implemented by teachers indicate varying success in student outcomes. In the few controlled trials examining autism intervention in schools, students tend to make progress across both experimental and control groups on standardized tests with more positive or mixed results regarding the advantage of specific EBPs when more proximal measures are examined (Mandell et al., 2013; Sam et al., 2021; Young et al., 2016). School research suggests that teacher fidelity to the intervention (the degree to which the intervention is being applied as specified in the treatment manual) has some relationship with student outcomes (Boyd et al., 2014; Mandell et al., 2013; Strain & Bovey, 2011; V. Wong et al., 2018). Additionally, high classroom quality overall may support

[1]UC Davis MIND Institute, Sacramento, CA, USA
[2]Child and Adolescent Services Research Center, CA, USA
[3]San Diego State University, CA, USA
[4]Stony Brook University, NY, USA
[5]University of California, San Diego, USA

Corresponding Author:
Aubyn C. Stahmer, Departments of Psychiatry and Behavioral Science and Psychology, UC Davis MIND Institute, 2825 50th Street, Sacramento, CA 95817, USA.
Email: astahmer@ucdavis.edu

improved fidelity in the specific practice (Boyd et al., 2014). More research on the use of EBPs in classrooms with school-age children during academically focused activities is needed.

Decades of research in child psychotherapy and more recent autism studies indicate outcomes may not be as positive when EBPs developed in research settings are translated to the community (Kurtines et al., 2004; Nahmias et al., 2019; Shelton et al., 2018). Therefore, researchers have suggested limiting the development and testing of new autism interventions that may not translate well to the community (Boyd et al., 2021). Rather, EBPs need to be adapted in collaboration with the community to fit the context. Evidence-based practices can be systematically adapted to improve fit with student and classroom characteristics while maintaining the active ingredients of the EBPs (Stahmer et al., 2019). Such adaptations should improve teachers' fidelity of the intervention and thus facilitate better outcomes for students (e.g., Brookman-Frazee et al., 2020; Durlak & DuPre, 2008; Sanetti & Kratochwill, 2009; Stahmer et al., 2020). However, adapted EBPs need to be tested in community settings to determine effectiveness.

One identified autism EBP is Pivotal Response Training (PRT; Steinbrenner et al., 2020). The PRT was designed based on a series of studies identifying treatment components that increase motivation to learn in autistic children and is one of several Naturalistic Developmental Behavioral Interventions (Schreibman et al., 2015) considered best practice for autism. Several systematic reviews have found PRT to be effective for improving multiple skills, most consistently language and social skills (Bozkus Genc & Vuran, 2013; Bozkus Genc & Yucesoy-Ozkan, 2016). A recent systematic review of randomized controlled trials of PRT found only five qualifying studies, three assessing PRT effectiveness in a parent training model and two using expert therapists in a clinic (Ona et al., 2020).

The limited evidence of how teachers implement PRT indicates significant modification (Stahmer, 2005) and low fidelity (Suhrheinrich et al., 2007). Observational studies of teachers using PRT have identified specific areas of strength and difficulty across training methods and settings (Stahmer et al., 2015). However, data from community studies also support the importance of PRT for improving student outcomes (Pellecchia et al., 2015).

Classroom Pivotal Response Teaching (CPRT) is an adaptation of PRT for classroom use (Stahmer et al., 2016) developed using systematic mixed-methods, community-partnered approach evaluating intervention components and teacher use of strategies in the classroom coupled with input from teachers and a community advisory board (see Suhrheinrich et al., 2020 and Stahmer et al., 2012 for further description). Like its predecessor, CPRT focuses on increasing student motivation, initiation, and responsivity

(Suhrheinrich et al., 2020) with the idea that dedicating effort toward improving these areas will result in positive change for other skills. By increasing overall student engagement in the learning environment through enhanced motivation, students will be able to access more learning opportunities. Thus, motivation supports should increase engagement and skill development. Rather than being a curriculum with a scope and sequence, CPRT is a framework by which to target educational goals in a way that is motivating to students at both individual and class-wide levels. Thus, teachers using CPRT are encouraged to integrate the intervention into their existing classroom activities.

The CPRT has several components to be implemented in each lesson (see Table 1). The primary components are the same as PRT components, and most adaptations provide examples of how to use these components in group settings and to meet Individualized Education Program (IEP) goals. Adaptations included simplifying complex components based on student developmental level, adapting direct reinforcement for use in group activities (e.g., using a group reward related to the activity), differential use of turn taking based on student language level and target skills, and providing examples of group implementation to meet educational goals (see Stahmer et al., 2012 for description of adaptations). For example, teachers indicated the use of conditional discrimination (red pen vs. red pencil or blue pen) was not developmentally appropriate for many of their students, which experimental studies confirmed (Reed et al., 2013). Therefore, CPRT provides recommendations for when to implement this component based on student characteristics.

Initial pilot data of the adapted protocol indicated positive outcomes for both teachers and students with participating teachers demonstrating fidelity to CPRT and use of CPRT associated with improved student engagement (Stahmer et al., 2016). More recently, as part of the randomized trial reported here, the majority (70%) of teachers met mastery criteria for CPRT fidelity and reported using CPRT 3 days a week for 50 min, on average (Suhrheinrich et al., 2020). These encouraging findings support further investigation of how CPRT impacts autistic students.

The current study evaluated the effectiveness of CPRT implemented in public schools on student outcomes. Specifically, the aims of the current study were to examine

1. the effects of teacher CPRT training on social and communication skills, adaptive behavior, educational goals, and classroom engagement during the training and follow-up year;
2. the relationship between teacher fidelity to CPRT and student outcomes; and
3. the role of overall classroom quality on student outcomes.

Table 1. CPRT Components and Descriptions.

Type	CPRT component	Description
Antecedent components	Incorporates choice	• The teacher provides specific choices to the student, either within or between activities.
	Follows student interest	• The teacher responds to student interest within the context of the lesson, activity, or learning objective and adjusts activities accordingly.
	Incorporates preferred materials	• The teacher sets up appealing activities or uses personalized, interesting, or unusual materials in the learning activities.
	Takes or facilitates turns	• The teacher or a student models a related play, academic, or communication skill *at or just above* the students' level.
	Presents opportunities (maintenance/acquisition)	• The teacher clearly *intersperses* tasks that are easy with tasks that are difficult for the students.
	Gains attention before cue	• The student is attending to the teacher *before* the teacher presents a cue.
	Provide clear cues	• A clear cue indicates to the students how they should respond and is at or slightly above the students' response level.
	Provides varied cues	• The teacher alters the type or form of the instruction or opportunity to respond related to each learning goal.
Consequence components	Provides contingent consequences	• The teacher provides immediate, appropriate feedback to the student based on their response (correct, incorrect, or attempt).
	Provides reinforcement	• The teacher rewards the student for appropriate responding and behavior with something other than praise.
	Provides direct reinforcement	• The teacher uses rewards that are directly related to the teaching activity and/or the behaviors required.
	Reinforces goal-directed attempts	• The teacher should provide reinforcement after most of the students' reasonable, goal-directed attempts.

Note. CPRT = Classroom Pivotal Response Trial.

Method

Design

The study tested the effectiveness of CPRT implementation on student outcomes in a hybrid Type 1, randomized, wait-list controlled trial across classrooms (Registry of Effectiveness Trials ID: 430.1). Classrooms were randomized to one of three training cohorts: Cohort A received CPRT training in Year 1. Cohort B was observed in Year 1 and received CPRT training in Year 2. Cohort C was observed in Year 2 and received CPRT training in Year 3. Serving as comparison groups, Cohorts B and C represented current autism services as usual during the years they were observed (Years 1 and 2, respectively). As recommended by Rhoads (2010), an independent consultant randomized participants (to reduce selection bias) at the classroom level to increase efficiency and statistical power. Teachers in each classroom identified two students being served under the educational classification of autism for enrollment. Students participated for the duration of their time in the participating classroom. Teachers identified new students if needed during subsequent school years.

Participants

The research team contacted eligible public school districts in Southern California. Inclusion criteria for districts specified they serve at least 15 students, ages 3 to 12 years with an educational classification of autism. Of the 35 school districts in San Diego County, 21 districts were eligible; administrators from 17 (81%) districts agreed to participate.

Teachers. Inclusion criteria specified that teachers had not received prior CPRT training and had at least two students with autism in their classroom. School district staff shared study information with eligible teachers. Research staff met with interested teachers to explain the study. A total of 126 teachers consented to participation. After initial enrollment during Year 1, each teacher was randomly assigned to Cohort A (*n* = 36), B (*n* = 34), or C (*n* = 33). Due to additional teacher interest and to meet project enrollment goals, an additional 23 teachers enrolled in Year 2 and were randomly assigned to Cohort B (*n* = 9) or C (*n* = 14). Teachers were primarily female (93%) and White (86%; 18% reporting Hispanic ethnicity), held a master's (61%) or bachelor's (38%) degree, and had a wide range of special education

Table 2. Characteristics of the Students in the Classroom Pivotal Response Trial Study (*n* = 308).

Characteristic/clinical data		Measure
Sex (*n* = 277 reported): *n* (%)		
Female		41 (13.31)
Male		236 (76.62)
Average age at study entry (in years): $\bar{x}$ (*SD*)		5.94 (2.28)
Ethnicity (*n* = 178 reported): (*SD*)		
Hispanic/Latino		120 (38.96)
Non-Hispanic/Latino		158 (51.30)
Race: $\bar{x}$ (*SD*)		
African American $\bar{x}$		22 (7.14)
Asian /Pacific Islander		18 (5.84)
Caucasian/White		133 (43.18)
Native American		5 (1.62)
Filipino/a American		2 (0.65)
Other		52 (16.88)
Mixed		29 (9.42)
Not reported		44 (14.29)
Clinical data: *n*, $\bar{x}$ (*SD*)		
Nonverbal IQ (Standard Score)	286	59.49 (34.26)
Autism Diagnostic Observation Schedule, Second Edition Severity	292	7.27 (1.83)

and autism experience. A total of 31 (25%) teachers reported receiving any prior training in PRT, with only 6 of those teachers reporting receiving coaching with feedback. Additional detail and description of participating teachers can be found in Suhrheinrich et al. (2020).

Students. Each participating teacher enrolled two students with whom they would practice CPRT during the training period and for data collection. Inclusion criteria for students included being between 3 and 12 years of age, enrollment in a participating teacher's classroom, and receiving services under the educational classification of autism. A total of 308 students participated for as long as they remained enrolled in a participating teacher's classroom (see Table 2). Participating students had a mean age of 5.94 (*r* = 3–2 years) and were predominantly male (76.62%). Parents reported student race/ethnicity that reflects the diversity of the large, urban, Southern California region of the study (38.96% Hispanic/Latino, 51.30% Non-Hispanic/Latino White, 9.74% Not Reported). Upon enrollment, the research team conducted a student assessment to confirm autism using the Autism Diagnostic Observation Schedule–2 (ADOS-2; Lord et al., 2012) and characterized student cognitive skills using the Mullen Scales for Early Learning (MSEL; Mullen, 1995) or Differential Ability Scales–II (DAS-II; Elliot, 2007) based on age and developmental level (see Assessment section). The mean ADOS severity score (1–10 scale) was 7.27 (*SD* = 1.83). The mean nonverbal standard score (*M* = 100, *SD* = 15) across both cognitive assessments was 59.49 (*SD* = 34.26).

Figure 1 presents a consort diagram of participants with a focus on students. A teacher consort diagram is available in Suhrheinrich (2020).

Procedures

CPRT Training. Detailed information on CPRT training for teachers is available in Suhrheinrich et al. (2020). Briefly, the training involved active learning and practice-based instructional strategies, modeling of CPRT components, and ongoing coaching throughout the school year with data-based feedback on CPRT fidelity. Within each training Cohort (A, B, and C), teachers received training in small, collaborative groups within a district or by combining districts in close proximity. Training involved 12 hr of didactic, interactive lecture, delivered 2 hr per week over 6 weeks (with slight variability based on district schedules). Teachers watched video examples of each CPRT component and several longer lessons, participated in role-playing, practiced fidelity recording, and data collection during the didactic sessions. After the first three sessions, teachers began to receive individual, in-classroom CPRT coaching weekly. Coaches evaluated CPRT fidelity using the CPRT Assessment (see *Teacher CPRT Fidelity section*) and provided structured feedback using a standardized format. Upon successfully maintaining mastery criteria over two sequential visits, coaching was faded from weekly, to biweekly, and then monthly for the duration of the school year. Teachers were aware of the fading contingencies and sessions were not increased based on teacher performance.
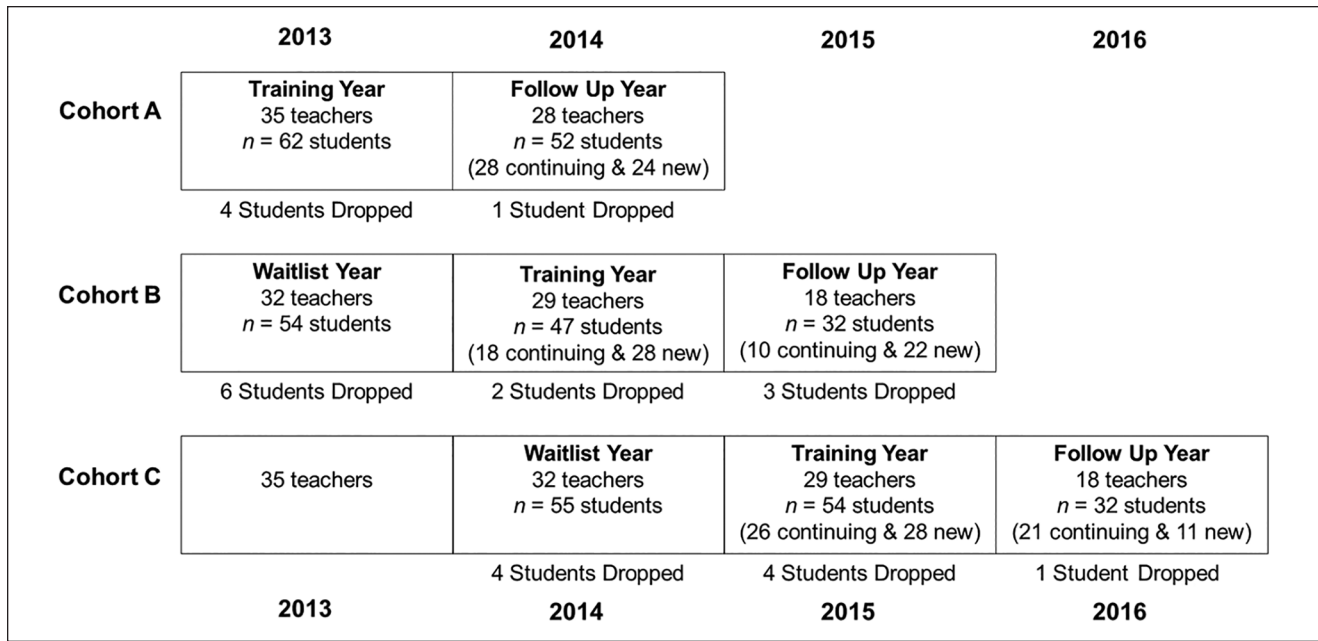
Figure 1. Consort Diagram of Study Participants With a Focus on Students.
Note. Teacher randomized to training group (see Suhrheinrich et al., 2020, for teacher consort diagram) and enrolled 2 students (with replacement).

During the follow-up year, teachers received two coaching sessions scheduled at their convenience.

Data Collection Procedures. Data were collected for teachers and students at study enrollment and throughout observation, training, and follow-up years. Teacher-completed measures were distributed via email with survey links. Teachers also had the option to complete measures using paper forms or by a phone interview. Parents received a packet with instructions, surveys, and a prepaid, preaddressed envelope from their child's teacher. Parents were asked to complete the surveys and return them to their child's teacher or mail them back to the research team. If the parents did not complete the paperwork, a member of the research team scheduled an interview to complete measures over the phone. Standardized student characterization measures were completed in school by the research team. At the beginning of the school year, teachers selected two classroom activities to be followed throughout the school year. Video recordings of these two activities were recorded 4 times during the control and training years and 2 times during the follow-up year by the research team. Activities recorded included a range of typical classroom tasks, including small group academics (e.g., math, reading, and language arts), individual instruction, play-based activities (e.g., puzzles, choice time), and whole group lessons (e.g., Circle time). Observation videos were, on average, 25.23 ($SD = 5.64$) min long. A total of 957 (77%) videos were group activities and 288 (23%) were 1:1 activities. These videos were used to code CPRT fidelity and student engagement.

Measures

Trained research staff naive to project hypotheses and condition completed standardized assessments.

Characterization Measures—Students

Autism Diagnostic Observation Schedule, Second Edition. The ADOS-2 is a semi-structured observational assessment of autism characteristics that includes four modules to evaluate social interactions, play, communication and stereotyped behaviors, and restricted interests based on the child language level (Lord et al., 2012). ADOS-2 has high reliability and validity (Gotham et al., 2008; Kamp-Becker et al., 2011). The research team administered ADOS-2 at enrollment. The majority of students received Module 1 (47.60%) with 31.85% of students receiving Module 2 and 20.55% of students receiving Module 3. The severity score was used to characterize the sample.

Mullen Scales of Early Learning / Differential Ability Scales-II. Each student received a cognitive measure based on age and developmental functioning and the hierarchy of use recommended by the Autism RUPP network (Arnold et al., 2000) and convergent validity of the measures (Bishop et al., 2011). Most students ($n = 161$) received the DAS-II. The DAS-II can be administered to students from age 30 months to 17 years 11 months. Four scales (Verbal, Nonverbal Reasoning, Spatial Reasoning, and Special Nonverbal Composite) are used to compute a composite score (General Conceptual Ability [GCA]; $M = 100$, $SD = 10$). For

students unable to obtain a basal score on one of the core subtests on the DAS, the examiner administered the MSEL. A total of 125 students received the MSEL, which can be administered to students from birth to 68 months of age. Five scales (Gross Motor, Visual Reception, Fine Motor, Expressive Language, and Receptive Language) are used to compute an overall standard score, the Early Learning Composite (ELC; $M = 100$, $SD = 15$). Construct, concurrent, and criterion validity are all verified by independent studies. The MSEL ELC standard score and the DAS GCA standard score were used to characterize the sample (Elliot, 2007; Mullen, 1995).

Dependent Measures

PDD Behavior Inventory. Teachers and parents completed the PDD Behavior Inventory (PDDBI; Cohen et al., 2003) rating scale at the beginning and end of the observation, training, and follow-up years to assess characteristics of autism and examine presentation over time. However, parent response rates for the PDDBI and the Vineland were quite low (30.08% and 27.27%, respectively); therefore, only teacher assessments were analyzed to ensure adequate power. The PDDBI is appropriate for children ages 2 to 12 years. Subscales measure maladaptive (sensory/perceptual approach behaviors; fears; arousal problems; aggressiveness /behavior problems; social pragmatic problems) and adaptive behaviors (social approach; learning, memory and receptive language; phonological skills; pragmatic ability). A summary "autism score" is provided. The PDDBI has good internal consistency (α coefficients range = .73–.97) for both the parent and the teacher versions. The measure has good construct, developmental, and criterion-related validity (Cohen, 2003). The Autism Composite, Approach/Withdrawal Problems Composite, and Receptive/Expressive Social Communication Abilities Composite T scores were used to examine student change over time.

Vineland Adaptive Behavior Scales, Second Edition. The Vineland Adaptive Behavior Scales, Second Edition (VABS-II; Sparrow et al., 2005) measures personal and social skills and has been validated in children with developmental disabilities and is applicable to children from birth through 18 years 11 months of age. The VABS-II assesses four domains: Communication, Daily Living Skills, Socialization, and Motor Skills. An Adaptive Behavior Composite (ABC) summarizes across the four adaptive behavior domains. Internal consistency for the ABC is 0.94. Test–retest reliability is .88 and interrater reliability is .74. Teachers completed the Teacher Rating Form for each student at the beginning and end of each school year. The ABC standard score was used to examine student change over time.

Goal Attainment Scaling. Goal Attainment Scaling (GAS) was used to assess progress toward IEP objec-

tives. The GAS is an evaluation technique for developing individualized, scaled descriptions for outcome measures (e.g., Oren & Ogletree, 2000). The GAS provides a scalable assessment of IEP goals to allow comparison across goals, students, and classrooms. Based on procedures for autistic students (Ruble et al., 2012), each teacher identified two IEP goals to target for each student. The research team and teacher developed objective, behavioral descriptors delineating observable estimates of degrees of progress toward the goal to help ensure consistency and that data were completed prior to training. Goals were scaled along a 5-point scale (-2 = *student's present levels of performance*, -1 = less than expected progress, 0 = expected level of outcome, 1 = somewhat more than expected, 2 = *much more than expected*). Teachers identified present levels of performance (-2), estimated the expected level of progress for the end of the year (0), and what to expect if the student greatly exceeded expectations ($+2$). Research staff then determined equidistant expectations for the -1 and $+1$. The principal investigator and a consulting researcher who is a developer of the GAS coding system (Ruble et al., 2012) rated each goal along a 3-point Likert-type scale in terms of equality, difficulty, and measurability. Goals that did not meet standards (average of 2.5) were updated by the teacher until they met standards.

At the end of each school year, teachers described the student's level of performance. Research staff naive to the training condition then determined the GAS Score to reduce bias. Prior studies have found a difference between teacher and research GAS ratings (Ruble et al., 2012). Research staff reviewed any data that supported the teacher's report (e.g., IEP progress reports; anecdotal notes; and data sheets). Teacher-supplied supplemental data were available for 28% of goals. Additionally, research staff observed the child completing the skill for 50.47% of goals. Of these observed goals, teacher reports matched the research staff observation 84.76% of the time. Two GAS goals were used in the analyses and classified as either *primary* (goal with the most progress) or *secondary*.

Classroom engagement. Classroom videos were continuously coded for student engagement using "The Observer Video-Pro" software (Noldus Information Technology, Inc.). Engagement codes included off-task (e.g., student demonstrates behavior that disrupts the activity lasting longer than 3 s, such as tantrums, leaving the table) and an active composite (e.g., student is engaged in the activity appropriately such as working independently or with a peer or passively observing the teacher). Complete codes are available from the authors.

Teacher CPRT fidelity. The CPRT fidelity evaluation form included 12 items (see Table 1 for CPRT components; Stahmer et al., 2012, 2016) rated on a 1 to 5 Likert-type scale

(1 = *teacher correctly implements the component less than 30% of the observation* through 5 = *teacher implements the component correctly through the entire observation [100% of opportunities]*), where a 4 or 5 was considered "correct." The measure addresses both adherence to (i.e., the degree to which proscribed procedures are utilized) and quality (i.e., skill used in delivery) of teachers' CPRT use (Sanetti & Kratochwill, 2009). Contact the authors for the full coding definitions. The percentage of components teachers used correctly (scored 4 or 5) were used in analyses. In analyses examining student change across the school year, the teacher's best fidelity score for that year was used. For student engagement analyses, teacher fidelity for the specific session was used.

Coding Procedures

Coders. Research assistants naive to the condition coded CPRT Fidelity or student engagement. Coders were trained to a reliability criterion of 80% agreement across two consecutive videos. Following initial training, interrater reliability was examined on an ongoing basis to reduce coder drift. If coders had two consecutive videos below 80% agreement, they retrained to the original criterion across two consecutive videos before further independent coding.

Interrater reliability. For *child engagement*, a total of 41% of videos (*n* = 508) were randomly selected for evaluation by two independent coders, and the percentage of agreement was calculated by the Observer Video-Pro software to assess interrater reliability. The average percentage of agreement across all videos was 79%, with a range of 60% to 95%. For *teacher fidelity*, 33% of videos (*n* = 354) was evaluated by two independent coders. Intraclass correlation coefficients (ICCs) were calculated to assess interrater reliability. The ICC mean was .81 (range = .67–.89). Contingent consequences and varied cues were below 0.7 and were not considered reliable and dropped from analysis (Cicchetti, 1994).

Classroom Moderators

Classroom quality. The quality of the classroom for supporting autistic students was evaluated using the Professional Development Assessment (PDA; Boyd et al., 2014). The PDA includes a 2-hr observation, a 30-min teacher interview, and a record review (i.e., review of IEPs and lesson planning documents). The PDA was completed by research staff naïve to training condition in the fall and spring of each school year. Detailed information on PDA training is available in Suhrheinrich et al. (2020); however, reliability was attained for each assessor across one video-recorded observation and one in vivo PDA prior to independent administration as per developer instructions. The PDA includes 50 items across seven domains. Item ratings were completed on a Likert-type scale (1 = *minimal/no implementation* to 5 = *full implementation*). Only the Classroom Environment scale was utilized in the current analyses, as items in this scale were most closely aligned with the components of CPRT and with teacher fidelity (Suhrheinrich et al., 2020). These included the use of clear and developmentally appropriate instructions, varied types of opportunities to respond, the use of natural/direct reinforcement, opportunities for students to make choices, provision of contingent consequences, opportunities to generalize skills, the use of individualized reinforcers, and incorporation of student interests and strengths in learning activities.

Ethics Approval. This research was approved by the Institutional Review Board at the University of California, San Diego with reliance from Rady Children's Hospital, and the University of California, Davis.

Data Analyses. To assess the impact of the CPRT training on student outcomes, we initially planned to use hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) due to the nested structure of the data: repeated measures (time: Fall, Spring) [Level 1], nested within students [Level 2], nested within teachers [Level 3], nested within districts [Level 4]. However, the HLM models did not converge, so Generalized Estimating Equations were used instead to account for clustering. First, we assessed for the effect of time (Fall vs. Spring) on GAS scores, communication, and social behaviors (PDDBI), and adaptive behavior (VABS). To assess the impact of CPRT training group (planned comparisons: observation vs. training, observation vs. follow-up), CPRT fidelity, and classroom quality on the change in student outcomes over the course of the academic year, we calculated change scores (Spring minus Fall) for all measures of interest. We then ran preliminary unadjusted models for CPRT training group, CPRT fidelity, and classroom quality (without covariates). For the final adjusted models, time since Fall assessment was used as a Level 1 predictor, and teacher fidelity to CPRT strategies and classroom quality were included as Level 2 predictors. For the observational measure of student engagement, all available observations were included in the model, and teacher fidelity to CPRT strategies and classroom quality were included as predictors.

Results

Time Effects

Significant effects for time (Fall vs. Spring assessment) were found for GAS, PDDBI Autism Composite, PDDBI Receptive/Expressive Social Communication Abilities Composite, and VABS Adaptive Behavior Composite (all *p* values <.027).

Table 3. Adjusted Models: Change in Teacher Report Composite Scores.

| | PDDBI composites | | | | | | | | | | | |
| | Autism | | | Approach/withdrawal problems | | | Receptive/expressive social communication abilities | | | VABS-II: Adaptive behavior composite | | |
Parameter	B	SE	p	B	SE	p	B	SE	p	B	SE	p
Intercept	6.21	5.05	.22	6.53	5.72	.25	−1.23	2.50	.62	6.41	4.51	.16
Time btwn assessments	0.05	0.20	.79	0.06	0.22	.79	0.10	0.12	.40	−0.27	0.17	.12
PDA classroom environ.	−1.37	1.14	.23	−1.49	1.34	.27	0.17	0.56	.76	−0.72	1.10	.52
CPRT fidelity (%)	−4.87	3.24	.13	−1.90	4.01	.64	3.03	2.09	.15	0.95	3.26	.77
Group: Training vs. observation	−2.19	1.15	.06	**−2.69**	**1.31**	**.04**	0.60	0.65	.36	1.06	1.00	.29
Group: Follow-up vs. observation	0.63	1.25	.62	0.10	1.40	.94	−0.71	0.71	.32	−0.63	1.02	.54

Note. Bold text indicates $p < .05$. PDDBI = Pervasive Developmental Disorder Behavior Inventory; VABS = Vineland Adaptive Behavior Scales (2nd ed.); PDA = Professional Development Assessment (Boyd et al., 2014); CPRT = Classroom Pivotal Response Trial.

Unadjusted Nested Models

CPRT Training Group. Appropriate Student Engagement was significantly higher in training ($M = 91.29\%$) and follow-up years ($M = 92.98\%$) compared with the observation year ($M = 89.29\%$, p-values $< .05$). Off-task engagement was lower in the follow-up year ($M = 5.33\%$) compared with the observation year ($M = 6.90\%$) at a marginal level of significance ($p < .07$). Group differences were nonsignificant ($p > .07$) for GAS, VABS Adaptive Behavior Composite, and all PDDBI composite scores.

CPRT Fidelity. Higher teacher CPRT fidelity was associated with higher appropriate student engagement ($B = 0.18, p < .0001$) and lower off-task engagement ($B = -0.12, p < .001$). Higher teacher CPRT fidelity was also associated with greater increases in GAS scores ($B = 1.64, p = .02$, secondary GAS score: $B = 1.59, p = .01$). The CPRT fidelity did not significantly predict ($p > .07$) changes on the VABS Adaptive Behavior Composite or any PDDBI composite scores.

Classroom Quality. Higher classroom quality was associated with higher appropriate student engagement ($B = 0.03, p < .01$) and lower off-task engagement ($B = -0.023, p < .01$). Higher classroom quality was associated with greater increase in secondary GAS goal scores ($B = 0.41, p = .02$). Classroom quality was not significantly associated (p values $> .07$) with changes in the VABS Adaptive Behavior Composite or any PDDBI composites.

Adjusted Models

This section describes the relative effects of CPRT training group, CPRT fidelity, and classroom quality, adjusting for time between assessments.

Goal Attainment Scaling. Change in GAS score did not significantly differ by group. However, higher observed teacher fidelity to CPRT strategies was associated with greater increases in both GAS scores ($B = 1.64, p = .02$), controlling for time between assessments, group, and classroom quality.

PDDBI (Teacher Report). Changes in the PDDBI Approach/Withdrawal Problems Composite significantly differed by group (see Table 3) such that students whose teachers were in the training group demonstrated a greater decrease in approach/withdrawal problems compared to students whose teachers were in the observation group (–2.79 points vs. –0.10 points, $B = -2.69, p < .05$) when controlling for time between assessments, teacher fidelity, and classroom quality. Changes in PDDBI Autism Composite differed by group at a marginal level of significance. Students whose teachers were in the training group demonstrated a greater decrease in Autism Composite scores when compared to students whose teachers were in the observation group (−3.70 points vs. –1.51 points, $B = -2.19, p < .06$), controlling for time between assessments, teacher fidelity, and classroom quality. No other PDDBI composite scores had significant predictors of change (all p values $> .1$).

Adaptive Behavior. There were no significant predictors of change on the teacher report VABS.

Student Engagement. When controlling for other factors, appropriate student engagement was higher in students whose teachers were in follow up compared with observation (see Table 4; 92.48% vs. 89.77%, $B = 0.03, p < .01$). Higher appropriate engagement was also predicted by higher classroom quality ($B = 0.02, p < .04$) and higher CPRT fidelity ($B = 0.14, p < .0001$). Off-task engagement

Table 4. Adjusted Models: Student Engagement.

Parameter	Active composite plus			Off-task		
	B	SE	p	B	SE	p
Intercept	0.74	0.04	.00	0.17	0.04	.00
PDA classroom environment	**0.02**	**0.01**	**.04**	−0.01	0.01	.08
Teacher fidelity to CPRT strategies (%)	**0.14**	**0.04**	**.00**	**−0.10**	**0.03**	**.00**
Group: Training vs. observation	0.02	0.01	.11	−0.01	0.01	.53
Group: Follow-up vs. observation	**0.03**	**0.01**	**.01**	−0.01	0.01	.29

Note. Bold text indicates $p < .05$. PDA = Professional Development Assessment (Boyd et al., 2014); CPRT = Classroom Pivotal Response Trial.

was associated with lower CPRT fidelity ($B = 0.10$, $p = .001$), controlling for group and classroom quality.

Discussion

This is one of the first studies of a collaboratively adapted autism EBPs conducted in classrooms with teacher implementation. Teacher participation in CPRT training significantly and positively affected student classroom engagement during CPRT lessons (proximal outcome) and the Approach/Withdrawal Composite on the PDDBI (distal outcome) even when controlling for classroom quality and CPRT fidelity. In contrast, we found no significant overall effect of teacher participation in CPRT training on student adaptive behavior or educational goals. These findings are consistent with the mixed outcomes from other classroom-based autism intervention effectiveness trials (e.g., Boyd et al., 2014; Mandell et al., 2013; Sam et al., 2021; Young et al., 2016). In our study, student engagement data were collected through video recordings of student/teacher interactions in their regular learning environment. This may suggest proximal measures of student behavior are more sensitive to change related to teacher intervention delivery (Boyd et al., 2014), especially when teachers are learning the intervention during the measurement period or that students did not make significant enough gains in the domain areas assessed by our distal measures.

Both significant group findings may be linked to CPRT theoretical foundations: increasing student motivation and engagement. Teachers had better fidelity to the antecedent strategies (see Table 1) that aim to increase motivation and engagement (Suhrheinrich et al., 2020). In our early qualitative studies, teachers reported that they do not typically use or have training in these strategies (Stahmer 2007). Therefore, increased use of these strategies may have encouraged student motivation and engagement even if overall fidelity did not increase. The expectation is that student motivation and engagement will eventually lead to improved long-term learning. We did not see evidence of those types of changes in the current short-term study. Additionally, engagement results are tempered by measurement procedures and student

baseline norms. We collected video lesson samples that reflected the variability of classroom reality. Although our findings related to the impact of student engagement are statistically significant, they also may have limited clinical significance given the high level of engagement of students overall. Therefore, future studies may consider a more fine-grained analysis of student participation in the lesson, which may require more sophisticated recording strategies.

Intervention fidelity has been identified as a driver of effective implementation and a key facilitator of student/patient outcomes (e.g., Zitter et al., 2021). In the current population, teachers exhibited substantial variability in CPRT fidelity (Suhrheinrich et al., 2019). This variability supported the evaluation of fidelity as a moderator of outcomes. Results indicated that higher teacher CPRT fidelity predicted improvement in student progress toward individualized goals, suggesting that student access to CPRT strategies improved learning. Additionally, higher CPRT fidelity was associated with better student engagement and reduced off-task behavior during the teaching activity. These associations remained even when controlling for classroom quality. This suggests additional work to support teacher CPRT fidelity will effectively support autistic students in the classroom. While additional studies are needed, these results suggest that a focus of future research should be implementation strategies that improve teacher fidelity to CPRT more quickly during the school year to increase intervention quality and dosage which may have a greater effect on academic outcomes.

Some data support the idea that improved classroom quality facilitates reaching fidelity for new EBPs more quickly in autism classrooms (Odom et al., 2013). This may be partially due to the overlap between high-quality classrooms and some key ingredients in many autism EBPs (Boyd et al., 2014). For example, some of the strategies associated with autism EBPs (including CPRT) such as antecedent and reinforcement strategies have been found to be critical for classroom management and quality (Allday et al., 2013). Therefore, we examined the relationship between classroom quality and child outcomes. In our

sample, higher classroom quality predicted better student engagement in learning activities and greater progress on GAS goals even when controlling for CPRT fidelity and treatment group. This is important because disparities in classroom quality exist in both general education (Hirsch, 2007) and special education (Billingsley & Bettini, 2017) in districts serving historically marginalized student populations. Teachers may benefit from training in developing and maintaining a high-quality classroom environment for autistic students prior to training in specific EBPs.

More broadly, outcomes indicated significant effects of time across a majority of student outcomes. This is encouraging and indicates that students in the public schools in our study receiving services for autism are making improvements over time. These findings also align with other school-based randomized trials of autism interventions. For example, Young et al. (2016) found that preschool students in autism classrooms made improvements in all outcome areas regardless of being in the treatment or control group, with the treatment group making greater improvement in receptive language and social skills. Similarly, data from a randomized trial of the STAR program (Arick et al., 2004) indicate a limited association between intervention fidelity and student outcomes as measured by changes in overall cognitive ability, with a fidelity of only one of the three interventions (PRT) associated with student outcomes (Pellecchia et al., 2015). Of course, teachers in these studies across groups were motivated enough to participate in a research program. In a randomized trial of the Classroom Social, Communication, Emotional Regulation, and Transactional Support Intervention (Morgan et al., 2018), the intervention group did show better outcomes on measures of classroom active engagement and more distal measures such as the VABS-II communication scale and the Social Responsiveness Scale, but differences were somewhat modest. Recently, a meta-analysis of intervention outcomes for young children with autism indicated little to no positive associations when methodological concerns were accounted for (Sandbank et al., 2020). Interestingly, the one area measured in this study that did not show significant positive change across the school year for all groups, Approach/Withdrawal Problems on the PDDBI is a domain for which CPRT appeared to make a difference. Students in the training group demonstrated a significantly greater reduction than students in the observation group. This finding suggests that CPRT may have beneficial impacts on areas that are not currently addressed by existing educational services.

Limitations

These limited positive results based on treatment group may be affected by several issues. Teachers were still learning the strategies throughout the school year, with some teachers not meeting CPRT fidelity until late in the school year. This may limit intervention dosage during the measurement period as students may have received high-quality CPRT for only a month prior to completing post-assessments, thus limiting the full impact of CPRT on student outcomes as measured here. Efforts to track students for the year following CPRT training in each classroom were complicated by changing classroom assignments. Additionally, CPRT fidelity decreased in the follow-up year (Suhrheinrich et al., 2020) which may mean changes to training and long-term coaching are needed to maintain fidelity and maximally impact student outcomes. Significant changes in student adaptive and academic behavior may require greater intervention dosage in conjunction with a high level of fidelity, and future research should explore the question of dosage and quality relative to impact.

Additionally, this research took place in a large urban county that has generally high-quality educational services. Our student sample represents the diversity of the region, but the quality of educational programs for autistic students may not be reflective of school services more broadly across the United States. Future research should target broader geographic areas to include a more diverse participant sample.

Future research should also expand upon this work in several ways with specific attention to measurement concerns. To more accurately evaluate student progress, it may be important to include additional proximal measures of student behavior. The measure of student engagement, while useful in revealing group differences, may be improved by including additional categories of student behavior. Additionally, teacher CPRT fidelity may be better conceptualized as a composite of accuracy and dosage of intervention delivery after initial fidelity criteria are met, thus accounting for varied rates of learning across teachers.

Conclusion

In summary, this trial contributes to the limited but growing body of research on the effectiveness of autism interventions delivered in school programs. Based on the challenges of implementing EBPs in schools in a way that clearly differentially impacts academic outcomes future research may need a greater focus on implementation strategies that support effective EBPs' training and use. Based on the implications that improved student outcomes are more likely when teachers use CPRT strategies with fidelity in classroom activities, which is consistent with previous work, future effectiveness trials should focus on how to improve fidelity to EBPs. Implementation science provides a framework for addressing concerns such as increasing teacher release time for training, improving access to ongoing supervision and coaching, and increasing leadership support for EBPs' use which has the potential to support school-based results (Boyd et al., 2021). The encouraging news is that our results

support the idea that autistic students are making significant gains throughout the school year which may point to a need for research focusing on key elements of school programs that lead to the best outcomes. An additional challenge with effectiveness research is understanding the validity and sensitivity of measures that can detect clinically relevant changes during the intervention period. Overall, school providers may require significant support to develop high-quality classrooms that facilitate EBPs' use that leads to improved outcomes for autistic students.

Declaration of Conflicting Interests

Funding

ORCID iD

Allison Nahmias ![ORCID icon] https://orcid.org/0000-0003-4678-0367

References

Allday, R. A., Neilsen-Gatti, S., & Hudson, T. M. (2013). Preparation for inclusion in teacher education pre-service curricula. *Teacher Education and Special Education*, *36*(4), 298–311. https://doi.org/10.1177/0888406413497485

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

Arick, J., Loos, L., Falco, R. A., & Krug, D. (2004). *The STAR program (strategies for teaching based on autism research)*. Pro Ed.

Arnold, L., Aman, M., Martin, A., Collier-Crespin, A., Vitiello, B., Tierney, E., Asarnow, R., Bell-Bradshaw, F., Freeman, B., Gates-Ulanet, P., Klin, A., McCracken, J. T., McDougle, C. J., McGough, J. J., Posey, D., Scahill, L., Swiezy, N., Ritz, L., & Volkmar, F. (2000). Assessment in multisite randomized clinical trials of patients with autistic disorder: The Autism RUPP Network. *Journal of Autism and Developmental Disorders*, *30*(2), 99–111. https://doi.org/10.1023/a:1005451304303

Billingsley, B. S., & Bettini, E. (2017). Improving special education teacher quality and effectiveness. In J. M. Kauffman, D. P. Hallahan, & P. C. Pullen (Eds.), *Handbook of special education* (2nd ed., pp. 501–520). Routledge.

Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the Differential Ability Scales in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, *116*, 331–343. https://doi.org/10.1352/1944-7558-116.5.331

Boyd, B. A., Hume, K., McBee, M. T., Alessandri, M., Gutierrez, A., Johnson, L., Sperry, L., & Odom, S. L. (2014). Comparative efficacy of LEAP, TEACCH and non-model-specific special education programs for preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *44*, 366–380. https://doi.org/10.1007/s10803-013-1877-9

Boyd, B. A., Stahmer, A. C., Odom, S. L., Wallisch, A., & Matheis, M. (2021). It's time to close the research gap in autism: The need for implementation science. *Autism*, *26*(3), 569–574. https://doi.org/10.1177/13623613211064422

Bozkus Genc, G., & Vuran, S. (2013). Examination of studies targeting social skills with pivotal response treatment. *Educational Sciences: Theory & Practice*, *13*, 1730–1742. https://doi.org/10.12738/estp.2013.3.1549

Bozkus Genc, G., & Yucesoy-Ozkan, S. (2016). Meta-analysis of pivotal response training for children with autism spectrum disorder. *Education and Training in Autism and Developmental Disabilities*, *51*, 13–26. https://eric.ed.gov/?q=peer+reviewed+only&ff1=dtysince_2003&pg=10&id=EJ1090294

Brookman-Frazee, L., Chlebowski, C., Villodas, M., Roesch, S., & Martinez, K. (2020). Training community therapists to deliver an individualized mental health intervention for autism spectrum disorder: Changes in caregiver outcomes and mediating role on child outcomes. *Adolescent Psychiatry*, *60*(9), 355–266. https://doi.org/10.1016/j.jaac.2020.07.896

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instrument in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cohen, I. L. (2003). Criterion-related validity of the PDD Behavior Inventory. *Journal of Autism and Developmental Disorders*, *33*(1), 47–53. https://doi.org/10.1023/a:1022278420716

Cohen, I. L., Schmidt-Lackner, S., Romanczyk, R., & Sudhalter, V. (2003). The PDD Behavior Inventory: A rating scale for assessing response to intervention in children with pervasive developmental disorder. *Journal of Autism and Developmental Disorders*, *33*(1), 31–45. https://doi.org/10.1023/a:1022226403878

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal Community Psychology*, *41*(3), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Elliot, C. D. (2007). *Differential Ability Scales* (2nd ed.). Harcourt Assessment.

Gotham, K., Risi, S., Dawson, G., Tager-Flusberg, H., Joseph, R., Carter, A., Hepburn, S., Mcmahon, W., Rodier, P., Hyman, S. L., Sigman, M., Rogers, S., Landa, R., Spence, M. A., Osann, K., Flodman, P., Volkmar, F., Hollander, E., Buxbaum, J., Pickles, A., & Lord, C. (2008). A replication of the autism diagnostic observation schedule (ADOS) revised algorithms. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*, 642–651. https://doi.org/10.1097/CHI.0b013e31816bffb7

Hirsch, D. (2007). *Experiences of poverty and educational disadvantage*. Joseph Rowntree Foundation.

Kamp-Becker, I., Ghahreman, M., Heinzel-Gutenbrunner, M., Peters, M., Remschmidt, H., & Becker, K. (2011). Evaluation of the revised algorithm of Autism Diagnostic Observation Schedule (ADOS) in the diagnostic investigation of high-functioning children and adolescents with autism spectrum disorders. *Autism*, *17*, 87–102. https://doi.org/10.1177/1362361311408932

Kasari, C., & Smith, T. (2013). Interventions in schools for children with autism spectrum disorder: Methods and recommendations. *Autism*, *17*(3), 254–267. https://doi.org/10.1177/1362361312470496

Kurtines, W. M., Silverman, W. K., & Hoagwood, K. (2004). Research progress on effectiveness, transportability, and dissemination of empirically supported treatments: Integrating theory and research. *Clinical Psychology: Science and Practice*, *11*, 295–299. https://doi.org/10.1093/clipsy.bph084

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule* (2nd ed.). Western Psychological Services.

Mandell, D. S., Stahmer, A. C., Shin, S., Xie, M., Reisinger, E., & Marcus, S. C. (2013). The role of treatment fidelity on outcomes during a randomized field trial of an autism intervention. *Autism*, *17*, 281–295. https://doi.org/10.1177/1362361312473666

McGee, G. G., & Morrier, M. J. (2005). Preparation of autism specialists. In F. R. Volmar, P. Rhea, A. Kline, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders. Vol. II: Assessment, interventions, policy* (3rd ed., pp. 1123–1160). John Wiley.

Morgan, L., Hooker, J. L., Sparapani, N., Reinhardt, R. A., Schatschneider, C., & Wetherby, A. (2018). Cluster randomized trial of the classroom SCERTS intervention for elementary students with autism spectrum disorder. *Journal of Consulting and Clinical Psychology*, *86*, 631–644. https://doi.org/10.1037/ccp0000314

Mullen, E. (Ed.). (1995). *Mullen scales of early learning: AGS edition*. AGS Publishing.

Nahmias, A., Pellecchia, M., Stahmer, A. C., & Mandell, D. S. (2019). Effectiveness of community-based early intervention for children with autism spectrum disorder: A meta-analysis. *Journal of Child Psychology and Psychiatry*, *60*(11), 1200–1209. https://doi:10.1111/jcpp.13073

Odom, S. L., Cox, A., & Brock, M. (2013). Implementation science, professional development, and autism spectrum disorder. *Exceptional Children*, *79*, 233–251. https://doi.org/10.1177/0014402913079002

Odom, S. L., Sam, A. M., & Tomaszewski, B. (2022). Factors associated with implementation of a school-based comprehensive program for students with autism. *Autism*, *26*(3), 703–715. https://doi.org/10.1177/13623613211070340

Ona, H. N., Larsen, K., Nordheim, L. V., & Brurberg, K. G. (2020). Effects of pivotal response treatment (PRT) for children with autism spectrum disorders (ASD): A systematic review. *Review Journal of Autism and Developmental Disorders*, *7*, 78–90. https://doi.org/10.1007/s40489-019-00180-z

Oren, T., & Ogletree, B. (2000). Program evaluation in classrooms for students with autism: Student outcomes and program processes. *Focus on Autism and Other Developmental Disabilities*, *15*(3), 170–175. https://doi.org/10.1177/10883576000150030

Pellecchia, M., Connel, J. E., Beidas, R. S., Xie, M., Marcus, S. C., & Mandell, D. (2015). Dismantling the active ingredients of an intervention for children with autism. *Journal of Autism and Developmental Disorders*, *45*, 2917–2927. https://doi.org/10.1007/s10803-015-2455-0

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 2). SAGE.

Reed, S. R., Stahmer, A. C., Suhrheinrich, J., & Schreibman, L. (2013). Stimulus overselectivity in typical development: Implications for teaching children with autism. *Journal of Autism and Developmental Disorders*, *43*(6), 1249–1257. https://doi.org/10.1007/s10803-012-1658-x

Rhoads, C. H. (2010). The implications of "contamination" for experimental design in education. *Journal of Educational and Behavioral Statistics*, *36*(1), 76–104. https://doi.org/10.3102/1076998610379133

Ruble, L., McGrew, J. H., & Toland, M. D. (2012). Goal attainment scaling as an outcome measure in randomized controlled trials of psychosocial interventions in autism. *Journal of Autism and Developmental Disorders*, *42*, 1974–1983. https://doi.org/10.1007/s10803-012-1446-7

Sam, A. M., Odom, S. L., Tomaszewski, B., Perkins, Y., & Cox, A. W. (2021). Employing evidence-based practices for children with autism in elementary schools. *Journal of Autism and Developmental Disorders*, *51*, 2308–2323. https://doi.org/10.1007/s10803-020-04706-x

Sandbank, M., Bottema-Beutel, K., Crowley, S., Cassidy, M., Dunham, K., Feldman, J. I., Crank, J., Albarran, S. A., Raj, S., Mahbub, P., & Woynaroski, T. G. (2020). Project AIM: Autism intervention meta-analysis for studies of young children. *Psychological Bulletin*, *146*(1), 1–29. https://doi.org/10.1037/BUL0000215

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, *38*(4), 445–459. https://eric.ed.gov/?id=EJ867973

Schreibman, L., Dawson, G., Stahmer, A., Landa, R. J., Rogers, S. J., McGee, G. G., Kasari, C., Ingersoll, B., Kaiser, A. P., Bruinsma, Y., McNerney, E., Wetherby, A. M., & Halladay, A. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(8), 2411–2428. https://doi.org/10.1007/s10803-015-2407-8

Shelton, R. C., Cooper, B. R., & Stirman, S. W. (2018). The sustainability of evidence-based interventions and practices in public health and health care. *Annual Review of Public Health*, *39*, 55–76. https://doi.org/10.1146/annurev-publhealth-040617-014731

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales* (2nd ed.). AGS Publishing.

Stahmer, A. C. (2005, February). *Teaching professionals and paraprofessionals to use pivotal response training: Differences in training methods*. Paper presented at the annual meeting of the California Association for Behavior Analysis, Dana Point, CA, United States.

Stahmer, A. C. (2007). The basic structure of community early intervention programs for children with autism: Provider descriptions. *Journal of Autism and Developmental Disorders*, *37*(7), 1344–1354. https://doi.org/10.1007/s10803-006-0284-x

Stahmer, A. C., Collings, N., & Palinkas, L. (2005). Early intervention practices for children with autism: Descriptions from community providers. *Focus on Autism and Other Developmental Disabilities*, *20*, 66–78. https://doi.org/10.1177/10883576050200020301

Stahmer, A. C., Dababnah, S., & Rieth, S. R. (2019). Considerations in implementing evidence-based early autism spectrum disorder interventions in community settings. *Pediatric Medicine*, *2*, 18. https://doi.org/10.21037/pm.2019.05.01

Stahmer, A. C., Reed, S., Lee, E., Reisinger, E., Connell, J., & Mandell, D. (2015). Training teachers to use evidence-based practices for autism: Examining procedural fidelity. *Psychology in the Schools*, *52*, 181–195. https://doi.org/10.1002/pits.21815

Stahmer, A. C., Rieth, S. R., Dickson, K. S., Feder, J., Burgeson, M., Searcy, K., & Brookman-Frazee, L. (2020). Project ImPACT for toddlers: Pilot outcomes of a community adaptation of an intervention for autism risk. *Autism*, *24*(3), 617–632. https://doi.org/10.1177/1362361319878080

Stahmer, A. C., Suhrheinrich, J., Reed, S. R., & 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*, Article 709861. https://doi.org/10.1155/2012/709861

Stahmer, A. C., Suhrheinrich, J., & Rieth, S. R. (2016). A pilot examination of the adapted protocol for Classroom Pivotal Response Teaching. *Journal of the American Academy of Special Education Professionals*, Winter, 119–139.

Steinbrenner, J., Hume, K., Odom, S., Morin, K. L., Nowell, S. W., Tomaszewski, B., Szendrey, S., McIntyre, N. S., Yucesoy-Ozkan, S., & Savage, M. N. (2020). *Evidence-based practices for children, youth, and young adults with autism*. Frank Porter Graham Child Development Institute, National Clearinghouse on Autism Evidence and Practice Review Team.

Strain, P., & 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

Suhrheinrich, J., Melgarejo, M., Root, B., Aarons, G. A., & Brookman-Frazee, L. (2021). Implementation of school-based services for students with autism: Barriers and facilitators across urban and rural districts and phases of implementation. *Autism*, *25*(8), 2291–2304. https://doi.org/10.1177/13623613211016729

Suhrheinrich, J., Rieth, S., Dickson, K., Roesch, S., & Stahmer, A. (2020). Classroom pivotal response teaching: Teacher training outcomes of a community efficacy trial. *Teacher Education and Special Education*, *43*(3), 215–234. https://doi.org/10.1177/0888406419850876

Suhrheinrich, J., Stahmer, A., & Schreibman, L. (2007). A preliminary assessment of teachers' implementation of pivotal response training. *Journal of Speech, Language Pathology, and Applied Behavior Analysis*, *2*(1), 8–20. https://doi.org/10.1037/h0100202

Wilson, K. P., & Landa, R. J. (2019). Barriers to educator implementation of a classroom-based intervention for preschoolers with autism spectrum disorder. *Frontiers in Education*, *4*, Article 27. https://eric.ed.gov/?id=ED593866

Wong, C., Odom, S., Hume, K., Cox, A., Fettig, A., Kucharczyk, S., Brock, M., Plavnick, J., Fleury, V., & Schultz, T. (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. https://doi.org/10.1007/s10803-014-2351-z

Wong, V., Ruble, L. A., McGrew, J. H., & Yu, Y. (2018). An empirical study of multidimensional fidelity of COMPASS consultation. *School Psychology Quarterly*, *33*(2), 251–263. https://doi.org/10.1037/spq0000217

Young, H. E., Falco, R., & Hanita, M. (2016). Randomized, controlled trial of a comprehensive program for young students with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*, 544–560. https://doi.org/10.1007/s10803-015-2597-0

Zitter, A., Rinn, H., Szapuova, Z., Avila-Pons, V. M., Coulter, K. L., Stahmer, A. C., Robins, D. L., & Vivanti, G. (2021). Does treatment fidelity of the early start Denver Model impact skill acquisition in young children with autism? *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-021-05371-4