Original Article

autism

Autism
1–12

DOI: 10.1177/1362361318787802
journals.sagepub.com/home/aut
SAGE

The impact of implementation support on the use of a social engagement intervention for children with autism in public schools

Jill Locke[1], Wendy Shih[2], Christina D Kang-Yi[3], Julie Caramanico[3], Travis Shingledecker[3], Jordan Gibson[3], Lindsay Frederick[1] and David S Mandell[3]

Abstract

Several interventions have demonstrated efficacy in improving social outcomes for children with autism; however, few have been successfully implemented in schools. This study compared two implementation strategies to improve the use of a social engagement intervention for children with autism in public schools. In total, 31 children with autism in grades K-5 and 28 school personnel participated in a randomized controlled trial. Schools were randomized to (1) training in Remaking Recess, a social engagement intervention, or (2) training in Remaking Recess with implementation support. Linear regression with random effects was used to test the intervention effects on implementation fidelity and social outcomes (peer engagement, social network inclusion, and friendship nominations). In both groups, implementation fidelity improved after training but remained low. Children in the Remaking Recess with implementation support condition had significantly higher social network inclusion and received more friendship nominations than children in the Remaking Recess–only condition ($p = 0.03$). Children in both groups experienced reduced solitary engagement ($p < 0.001$) and increased joint engagement ($p < 0.001$). The results suggest that implementation supports may have an effect on outcomes above and beyond the intervention, and that further research is needed into the active intervention mechanisms.

Keywords

autism spectrum disorder, implementation, school, social engagement intervention

This study examined the effect of school-level implementation support on the (1) use of Remaking Recess (RR; Kretzmann et al., 2012), a social engagement intervention that uses both adult-facilitated (working with the student) and peer-mediated (working with typically developing peers) intervention strategies for children with autism, and (2) students' social outcomes (peer engagement, social network inclusion, and friendship nominations). Children with autism have poorer social outcomes than their typically developing classmates (Bauminger et al., 2010; Mendelson et al., 2016; Rotheram-Fuller et al., 2010). Several interventions have shown efficacy in improving these outcomes at school (Carter et al., 2010; Carter and Kennedy, 2006; Frankel et al., 2011; Kasari et al., 2012, 2016; Lane et al., 2015). While these interventions differ in strategy and format (e.g. parent- and peer-mediated interventions, positive behavioral intervention and supports, and social skills groups), their tests of efficacy all relied on expertly trained and highly supervised clinicians or graduate students rather than school personnel to deliver the interventions, which limits the generalizability of the study findings and the potential to sustain the use of the intervention once the research ends.

Licciardello et al. (2008) suggest that school personnel are ideal agents to implement social skills interventions, given that children with autism usually spend much of their day in school (Robertson et al., 2003). Many studies have demonstrated that training school personnel to implement autism-related interventions results in improved child outcomes (Howlin et al., 2007; Kamps et al., 2014;

[1]University of Washington, USA
[2]Loma Linda University, USA
[3]University of Pennsylvania, USA

Corresponding author:
Jill Locke, University of Washington, Box 354875, Seattle, WA 98195, USA.
Email: jjlocke@uw.edu

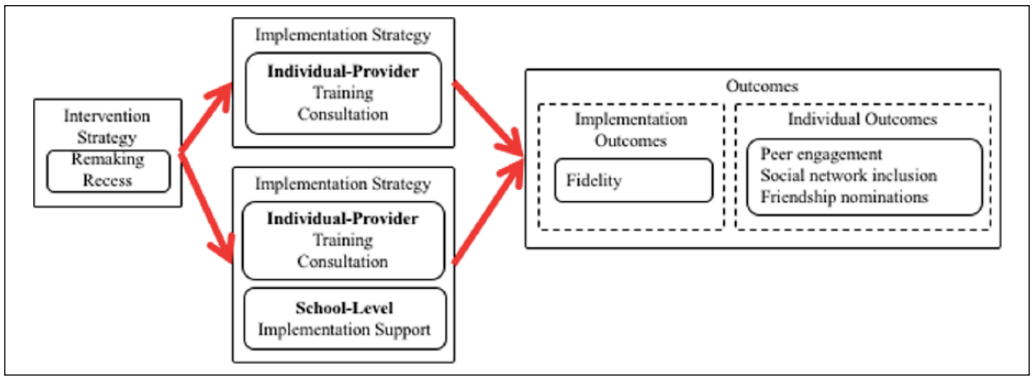

Figure 1. Conceptual framework.

Kretzmann et al., 2015; Mandell et al., 2013; Suhrheinrich et al., 2007). Most germane to this study, Kretzmann et al. (2015) demonstrated that training school personnel in RR improved playground peer engagement among elementary-aged children. Although school personnel were successfully trained to use RR, intervention implementation was not maintained at the 10-week follow-up observation. In a separate study, Locke et al. (2015) noted that school personnel could implement some components of RR successfully, but overall implementation fidelity—the degree to which the intervention was implemented as designed (Proctor et al., 2011)—was low.

Fidelity is an important measure of the success of intervention implementation in community practice (McLeod et al., 2013; Proctor et al., 2011). While a growing number of autism interventions have demonstrated efficacy when implemented by clinicians, few have been successfully implemented and sustained in schools by school personnel (Dingfelder and Mandell, 2011). When school personnel do implement evidence-based interventions, it is usually with low to moderate fidelity (Kretzmann et al., 2015; Locke et al., 2015; Mandell et al., 2013; Pellecchia et al., 2015; Stahmer et al., 2015), in large part due to limited resources and autism expertise (Durlak and DuPre, 2008). Research has identified many barriers to implementation when interventions are used in community settings (Locke et al., 2015; Mandell et al., 2013), such as inadequate training and insufficient support, which suggests that most schools will need additional support when implementing these practices.

We developed a strategy to increase the probability of successful implementation of RR in schools. Proctor et al. (2013) define an implementation strategy as a "method or technique used to enhance the adoption, implementation, and sustainability of a clinical program or practice" (p. 2). Powell et al. (2015) compiled a list of 73 possible implementation strategies. While implementation frameworks point to a myriad of strategies that map onto different phases (exploration, preparation, active implementation, and sustainment) and levels (organization, innovation, or individual) of the implementation process, most of these strategies are not relevant to schools. Because schools present a number of unique implementation challenges including educational timelines, professional characteristics, policies, and organizational constraints (Forman et al., 2013; Locke et al., 2015; Owens et al., 2014), few implementation strategies that have been designed to support clinical practice in healthcare settings can be easily applied in schools. We tested two implementation strategies most relevant to the school context: (1) *identify and prepare champions*, individuals who support and drive through an implementation and (2) *tailor strategies* to address barriers and leverage facilitators that were identified through earlier data collection (Powell et al., 2015).

The purpose of this study was to test the impact of implementation support that extends beyond the specific intervention or child outcomes being evaluated (McConachie and Fletcher-Watson, 2015). Specifically, this study examined whether providing school-level implementation support improves implementation fidelity and children's social outcomes (peer engagement, social network inclusion, friendship nominations) over and above training in RR alone. A growing body of research in implementation of other evidence-based mental health interventions in schools has identified school factors, such as culture and climate, that may affect implementation and sustainment of the use of evidence-based practices (Fixsen et al., 2005; Forman and Barakat, 2011; Forman et al., 2009, 2013; Hoagwood and Johnson, 2003; Langley et al., 2010; Lyon et al., 2014; Mandell et al., 2013; Massey et al., 2005; Owens et al., 2014). We augmented training in RR with implementation strategies designed to modify school factors such as culture and climate. Figure 1 depicts our theory of change. We draw from the Proctor et al.'s (2009) framework that distinguishes but links *implementation processes* and *implementation outcomes* to guide our study. Our model specifies that both the intervention (RR) and implementation strategy (training and consultation) affect implementation (fidelity) and individual outcomes (peer engagement, social network inclusion, and

friendship nominations) for children with autism. We hypothesized that school personnel in the RR with school-level implementation support would have higher implementation fidelity and children with autism in this condition will have better peer engagement, social network inclusion, and friendship nominations than those in the RR-only condition.

Methods

Participants

Participants included school-aged children with autism recruited from public elementary schools in the Northeastern United States over a 2-year period. The university institutional review board and each school district approved the study. Parents and guardians provided written consent and children assented.

Children were included if they (1) had a diagnosis of autism made by a community clinician; (2) were referred by school administrators; (3) had an intelligence quotient (IQ) of 65 or higher documented in their school records (to ensure that they had the *verbal* and *non-verbal* abilities to fully comprehend components of the intervention); and (4) were included in a kindergarten through fifth-grade general education classroom for at least 80% of the school day (to ensure that they had consistent access to general education peers). School records were used to document an autism diagnosis as opposed to independently validating an autism diagnosis using gold standard measures so that the study could be conducted within the same parameters and realistic constraints under which schools typically operate (e.g. schools do not have the ability to refuse intervention to identified children on their caseload). All children were performing at grade level at study enrollment, which suggests that they had appropriate verbal language abilities that allowed them to understand the intervention. In addition, all children independently and reciprocally engaged with others without assistive technologies or communication devices. A member of the research team confirmed this information with school personnel at the time of study enrollment. No children were excluded due to this parameter. These inclusion criteria were consistent with previous school-based trials of social engagement interventions (Kasari et al., 2012, 2016; Kretzmann et al., 2015; Locke et al., 2015). Children were excluded if they (1) were not expected to stay in the school or the classroom for the duration of the study and (2) did not have a participating school staff member on the playground during their recess period. These exclusion criteria were set to ensure that training would be completed as described within the study timeline.

A total of 55 schools in 27 school districts were contacted for participation. Of them, 40 schools were not interested in or did not have the capacity to participate in research. Three schools were deemed ineligible because they only had one child with autism who would qualify for the study, which would not allow for nesting (see section "Data analysis"). Hence, 12 schools in five districts participated. The average school size was 600 students (standard deviation (SD) = 176). On average, 50% of the students in each school received free/reduced price lunch, and the average class size was 24.2 students (SD = 4.5). From those schools, 35 children with autism were assessed for eligibility and consented to participate in the study. Four children with autism were excluded: one child moved during the baseline period and three children did not have a participating school staff member; thus, a total of 31 children were included. Another 322 typically developing classmates completed the Friendship Survey. See Figure 2.

All but three school personnel had one student with autism in their care; three school personnel served two students with autism. School personnel included 11 teachers and 17 other staff (classroom assistants, one-to-one assistants, and noontime aides). See Table 1 for demographic information.

Randomization

A randomized controlled design was used with three measurement points: start of implementation, post implementation, and 6 weeks post implementation. An independent data management core generated a random number sequence to assign randomization. Schools were randomly assigned to (1) training in RR or (2) training in RR with implementation support. Recess time was comparable for all students despite different grade levels; none of the school personnel dropped from training in RR.

Procedure and intervention

The research team met with the school district officials to obtain a list of eligible schools based on the inclusion criteria listed above. Subsequently, the research team met with the principal at each school to discuss the research activities and obtain a letter of agreement to conduct research on their campus. All recruitment materials (e.g. informational handouts, flyers) were then distributed to the school, and the research team met with interested participants to inform them about the study and their role as a study participant, in order for them to make an informed decision regarding their participation. Once informed consent was obtained from participating school personnel and parents of children with autism, research personnel distributed consent forms to all children in the classroom for participation in the Friendship Survey. Subsequently, a member of the research team reviewed the study with all consented children in lay language and administered an assent comprehension test to ensure that the children understood the study and their rights as research

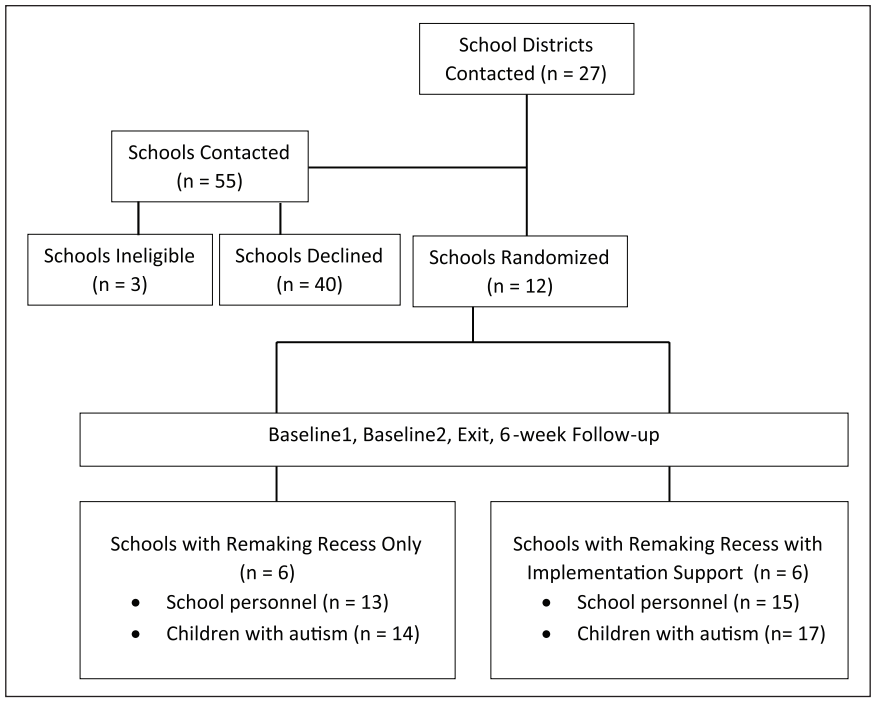

Figure 2. Consort chart.

participants. Consented and assented children completed the Friendship Survey, while blind observers recorded children's peer engagement on the playground at each measurement point during one school year. Observers were contracted from a different university to ensure that they were blinded to study randomization and procedures and did not have access to study information or data. Observers and coaches were separated throughout the study duration. In addition, research personnel separately consented school personnel. In both conditions, school personnel were trained in RR during the child's lunch recess (approximately 30–45 min) for 12 sessions over 6 weeks (two sessions per week) in the presence of children. Schools randomized to training in RR with implementation support received three additional consultation and implementation support sessions with school administrators over the six weeks. Study activities were embedded into children's daily activities (school personnel received $50 for their participation).

RR

RR is a school-based social engagement intervention for children with autism (Kretzmann et al., 2012). The intervention trains school personnel so that schools will be equipped to continue to support children with autism after external support is withdrawn. All school personnel were individually paired with a coach from the research team. The training modules included the following topic areas: (1) scan and circulate the cafeteria/playground for children who may need additional support; (2) identify children's engagement states with peers; (3) follow children's lead, strengths, and interests; (4) provide developmentally and age-appropriate activities and games to scaffold children's engagement with peers; (5) support children's social communicative behaviors (i.e. initiations and responses) and conversations with peers; (6) create opportunities to facilitate reciprocal social interaction; (7) sustain children's engagement within an activity or game; (8) coach children through difficult situations with peers should they arise; (9) provide direct instruction on specific social engagement skills; (10) individualize the intervention to specific children in order to generalize the intervention to other students in their care; (11) work with typically developing peers to engage children with autism; and (12) fade out of an activity/game so children learn independence (Kretzmann et al., 2015; Locke et al., 2015).

Each session took place during the child's lunch period and targeted one didactic skill. Coaches first explained the skill, how it applies to children with autism, and its importance in relation to the development of children's social functioning. Coaches then modeled how to use the targeted skill with children with autism and their peers. Subsequently, school personnel were asked to try the skill, so coaches could provide immediate feedback. At the end of each session, school personnel were given "homework" to practice the skills during the days when coaches were not present. Homework was reviewed at the next session.

Table 1. Child and school personnel characteristics by implementation condition.

Child characteristics	Remaking Recess (total N = 14)		Remaking Recess with implementation support (total N = 17)	
	N	Mean or %	N	Mean or %
Age (years)	14	9	17	8.6
Gender				
Male	13	92.86	14	82.35
Female	1	7.14	3	17.65
Race/ethnicity				
White	5	35.7	10	58.8
African American	5	35.7	5	29.4
Latino	2	14.3	0	0.0
Asian	1	7.1	2	11.8
Other	1	7.1	0	0.0
Grade level, n (%)				
Kindergarten	0	0.0	4	23.5
First grade	2	14.3	0	0.0
Second grade	3	21.4	2	11.8
Third grade	1	7.1	2	11.8
Fourth grade	1	7.1	4	23.5
Fifth grade	7	50.0	5	29.4
School personnel characteristics	Total N = 13		Total N = 15	
Age	13	38.3	15	40.5
Gender*				
Male	0	0.0	4	26.7
Female	13	100.0	11	73.3
Race/ethnicity				
White	8	61.5	9	60
African American	5	38.5	5	33.3
Latino	0	0.0	1	6.7
Highest education				
High school	3	23.1	4	26.7
Bachelors	6	46.2	5	33.3
Graduate degree	2	15.4	6	40
Associate degree	2	15.4	0	0.0
ASD experience	13	4.6	15	7.5
Role				
Teacher	3	23.1	8	50
Other school personnel	10	76.9	7	50

ASD: autism spectrum disorder.
*$p < 0.05$.

Implementation support (RR with implementation support)

Implementation support was individualized to each school (n = 6) to address its specific implementation needs. A PhD-level or MA-level coach (see below) met with school administrators and counselors, psychologists, teachers, and support staff, who were designated by the principals as the school champions of RR. School champions did not deliver RR. Over 6 weeks, implementation needs were identified. Common barriers to implementation were identified and included: (1) schedule staffing during recess; (2) build internal capacity (e.g. a team of trained RR implementers); (3) amend school-wide recess policies (i.e. detention, the removal of recess, and alternatives for indoor recess during inclement weather); (4) provide tangible support and resources (i.e. materials and space); (5) improve implementation climate (e.g. recognizing, supporting, and rewarding implementer efforts, being visible on the playground/in the cafeteria, etc.); (6) adapt and modify the intervention to fit the needs of the school (e.g. identifying the core vs peripheral components); and (7) embed RR within the school culture.

At the initial meeting, school champions were presented with this list of potential implementation supports. As a team, school champions selected 1–3 topics that they thought were most important to the successful use of RR. Schools most frequently selected embed RR in the school culture (n = 4), followed by build internal capacity (n = 3), improve implementation climate (n = 2), provide tangible support and resources (n = 2), adapt and modify the intervention to fit the needs of the school (n = 1), amend school-wide recess policies (n = 1), and schedule staffing during recess (n = 1). Together, the team prioritized the topics and discussed implementation logistics and strategies to address those barriers. In three subsequent meetings, the research team and school champions developed a plan for implementation based on the selected topic. For example, schools that wanted to embed RR within the school culture focused on raising awareness among the faculty, staff, and students that RR is an important new program. Typically, schools notify personnel on an "as-needed" basis; therefore, only a limited number of personnel (e.g. teachers, aides) who were directly involved in implementation were aware of RR. In schools that wanted to raise awareness of RR, the research team attended a school faculty/staff meeting or school assembly to present RR to the faculty and students. Once faculty and students were made aware of RR, school champions were tasked with making morning announcements reminding faculty, staff, and students that RR would occur at lunch, posting posters or a bulletin/marker board of RR activities for the week, and engaging other interested school members to facilitate RR for different grades.

Coaches

Two coaches were used throughout the study for both conditions. Coaches were randomly assigned to provide training in RR and RR with implementation support. Coaches were MA- or PhD-level members of the research team who were trained in education and/or psychology. All coaches were trained in (1) the school

consultation process, working collaboratively with school personnel to deliver an intervention to improve children's outcomes (Erchul and Martens, 2012) and (2) RR. Coaches were trained by one of the RR developers and were evaluated on a fidelity checklist that captured the core components of the intervention. Coaches were considered to be at fidelity with a criterion $\alpha > 0.80$ prior to providing consultation to school personnel.

Primary outcome: implementation outcomes

Fidelity (i.e. use and quality of intervention delivery) was measured in two ways: (1) coach-rated and (2) self-rated. Coaches rated fidelity on seven components of RR (i.e. attended to child engagement on the playground, transitioned child to an activity, facilitated activity, participated in activity, fostered communication, employed peer models, and provided direct instruction of social skills). Use was scored "0" for "no" and "1" for "yes" to measure whether school personnel used each RR component. The proportion of completed steps (completed steps/total number of RR components) was used for analyses. Quality of intervention delivery or how well school personnel used each component of the intervention was coded on a Likert-type scale from "1" (not well) to "5" (very well) for each component of RR that was used. The average quality rating across all intervention components was used for analysis. In addition, *accuracy, frequency of implementation, and understanding* of RR was measured on the same five-point scale. Reliability was collected on 20% of sessions during the study (mean percent agreement = 85; range 80%–90%). School personnel used a separate but parallel fidelity measure to self-rate their use and quality of intervention delivery at each time point.

Secondary outcomes: child outcomes

Playground Observation of Peer Engagement. The Playground Observation of Peer Engagement (POPE) is a timed-interval behavior coding system (Kasari et al., 2005). Independent observers, blinded to randomization, rated children on the playground for 40 consecutive seconds and then coded for 20 s during the lunch recess play period. Four observers were trained and considered reliable with percent agreement ⩾ 80%; reliability was collected on 20% of sessions during the study (mean percent agreement = 82% ranging from 80% to 87%; mean kappa = 90% ranging from 87% to 93%). Playground engagement states were expressed as the percentage (range 0%–100%) of intervals children spent in solitary play (i.e. unengaged with others) and joint engagement with peers (i.e. turn-taking in a game, reciprocal engagement in conversations or joint activities).

Friendship survey. Members of the research team blinded to study randomization administered the Friendship Survey to all consented children at a convenient time (approximately 10–15 min) for the classroom. Children were asked to identify classmates whom they liked to hang out with. This free recall list of friends determined the child's number of received friendship nominations (minimum = 0; no maximum). In addition, sociometric data were gathered within each participating class to gain a robust picture of children's peer groups. Participating students were asked: "Are there kids in your class who like to hang out together? Who are they?" as a method of identifying specific children within each classroom social network. Children listed the names of all children within their classroom who hung out together in a group using free call without additional prompting, class lists, or pictures. Children were reminded to include themselves in groups as well as students of both genders. Young children in Kindergarten and first grade with reading and writing difficulties were interviewed individually.

Coding social network inclusion. Social network inclusion refers to the prominence of each individual in the overall classroom social structure (Cairns and Cairns, 1994; Kasari et al., 2012). Traditional social network classifications were designed to be cross-sectional measures of children's classroom social network inclusion at one time point. A series of social network analyses were conducted to obtain each student's social network inclusion score (Cairns and Cairns, 1994). Two related scores were calculated: (1) the student's "individual centrality" or the total number of nominations to any peer group within the classroom and (2) the student's "cluster centrality" or the average individual centrality of the two highest group members in which the child is connected. Together, these two scores comprise the student's combined social network inclusion score (Cairns and Cairns, 1994). We note that children's social network inclusion score can only be as high as the lowest centrality score, which means that if a child's individual centrality is a "1", which suggests that the child is peripheral, and the cluster centrality is a "2", which suggests that the group is secondary, then the child's social network inclusion score is "1." Children's social network inclusion scores were normalized on the most nominated subject in the classroom during baseline and were calculated before and after each intervention using children's total received nominations divided by the highest total received nomination score within their classroom to examine the *change* of children's social network inclusion within the classroom (range 0–1).

Power analysis. Using a two-sample t-test, we have 80% power to detect a difference of 0.3 in social network inclusion with a Type I error of 0.05 with a sample size of 28 children. Our study has a sample size of 31 children. Hence, we are sufficiently powered to detect treatment effects.

Data analysis

Study data were managed using Research Electronic Data Capture (REDCap), a secure, web-based application designed to support data capture for research studies (Harris et al., 2009). Descriptive analyses for child and school personnel demographic characteristics were conducted for baseline comparisons between conditions to assess the success of randomization. Separate multilevel models were conducted to evaluate the effect of each condition (RR only or RR with implementation support) on implementation fidelity and child outcomes (playground engagement, social network inclusion, and received friendship nominations) across time to account for multiple time points nested within children who were nested within schools. Child outcomes were measured using continuous scales and were modeled using linear mixed models with random intercepts. All analyses were completed on an intent-to-treat basis. We included nesting by school (i.e. children within the same school) as a random effect in the mixed models. Model assumptions were checked for all final models (i.e. normality of residuals, homoscedastic variances, and appropriate covariance structures). Differences in child and aide characteristics were not detected between conditions. Implementation fidelity was measured using Likert-type scales and was modeled using ordinal logistic mixed models or logistic mixed models depending on the distribution of the measures. Each model evaluated the marginal effect of time and the interaction between time and condition along with the individual-level intercepts. Effect sizes (ES) were reported using Cohen's f in the section "Results" where the ES of 0.10, 0.25, and 0.40 are generally regarded as small, moderate, and large, respectively. Self-rated fidelity items were missing for seven children (22.7% of all children) at the follow-up period; therefore, these ratings were excluded from the analysis. In addition, self-rated fidelity items were missing for one child at baseline, and coach-rated fidelity items were missing for one child at exit. The missing data were minimal and treated through the PROC GLIMMIX analysis. PROC GLIMMIX used for the mixed models provided predicted values for observations using both the fixed and random components and produced approximately unbiased estimates even if outcomes were excluded due to missing data (Allison, 2012; Tao et al., 2015).

Results

Child and school personnel characteristics

Table 1 presents the demographic information of children with autism and school personnel in each condition. There were no significant differences between the conditions in children's gender, age, grade, or ethnicity. All school personnel in the RR-only condition were female and 68.7% in the RR with implementation support were female. The two groups were not statistically significantly different in age, race/ethnicity, roles in school, or years of work experience with children with autism.

Self-rated implementation fidelity of RR

On average, school personnel's self-rated fidelity increased from entry to exit in both conditions ($F(1,28)=39.98$, $p<0.001$, $ES=1.2$). The rate of improvement of self-rated fidelity was not significantly different between the two conditions ($F(1,28)=0.74$, $p=0.39$, $ES=0.16$). See Table 2.

Coach-rated implementation fidelity of RR

School personnel in both conditions improved in overall coach-rated fidelity ($F(1,60)=18.78$, $p<0.001$, $ES=0.56$). There were no statistically significant between-group differences in the rate of improvement for coach-rated fidelity of RR ($F(1,60)=0.12$, $p=0.733$, $ES=0.04$). Both conditions improved in each component (accuracy: $F(1,56)=20.83$, $p<0.001$, $ES=0.61$; implementation: $F(1,56)=14.17$, $p<0.001$, $ES=0.50$; and understanding: $F(1,56)=22.35$, $p<0.001$, $ES=0.63$) of the intervention, except for perceptions of feasibility (Table 2).

Child outcomes

Figure 3 presents changes in children's joint and solitary engagement over time. Children with autism in both conditions spent significantly more time in joint engagement ($F(1,29)=18.44$, $p<0.001$, $ES=0.80$) with peers and less time in solitary engagement ($F(1,29)=18.24$, $p<0.001$, $ES=0.80$) on the playground from baseline to exit. The significant increase in joint engagement and decrease in solitary engagement were maintained through follow-up ($F(1,60)=18.88$, $p<0.001$, $ES=0.56$ and $F(1,60)=16.64$, $p<0.001$, $ES=0.53$, respectively) where children in both conditions spent significantly more time in joint engagement and less time in solitary engagement at follow-up compared to baseline. There were no significant condition differences in their rate of change from baseline to exit in both joint and solitary engagement ($F(1,29)=0.13$, $p=0.73$, $ES=0.07$ and $F(1,29)=0.91$, $p=0.35$, $ES=0.18$, respectively).

There was a significant condition difference in the rate of improvement in social network inclusion ($F(1,29)=4.99$, $p=0.033$, $ES=0.41$) from baseline to exit; children in the RR with implementation support condition improved significantly more than children in the RR-only condition. Children's social network inclusion in the RR with implementation support condition increased from 0.38 at baseline to 0.69 at exit and decreased to 0.59 at follow-up; however, children's social network inclusion in the RR with implementation support condition remained significantly higher at follow-up compared to their baseline social network inclusion ($F(1,58)=7.80$, $p=0.007$, $ES=0.37$). Children's

Table 2. Mean aide-rated and coach-rated fidelity scores by condition.

	Remaking Recess		Remaking Recess with implementation support	
	Baseline	Exit	Baseline	Exit
Aide self-rated fidelity (%)	54.4 (30.1)	94 (6.7)	43 (35.8)	72.9 (33.1)
Coach-rated fidelity (%)	38.8 (28.3)	79.6 (24.4)	24.4 (28)	70.6 (32.4)
Accuracy: mean (SD)	2.2 (1.2)	4 (1)	1.6 (1)	3.6 (1.3)
Implementation: mean (SD)	1.6 (1.4)	2.6 (1.3)	1.1 (0.2)	3 (1.6)
Understanding: mean (SD)	2.2 (1.1)	4.4 (0.7)	2.5 (1.3)	4.2 (1.2)

SD: standard deviation.

Table 3. Child outcomes by condition.

	Remaking Recess	Remaking Recess with implementation support
	Mean (SD)	Mean (SD)
Joint engagement		
Baseline	25.24 (21.27)	29.33 (24.19)
Exit	53.87 (39.61)	53.59 (36.51)
Follow-up	57.66 (39.39)	53.55 (32.92)
Solitary		
Baseline	29.51 (26.7)	22.72 (17.7)
Exit	8.31 (13.04)	9.26 (14.78)
Follow-up	11.19 (17.67)	8.55 (12.15)
Social network inclusion		
Baseline	0.27 (0.25)	0.38 (0.22)
Exit	0.37 (0.29)	0.69 (0.36)
Follow-up	0.37 (0.27)	0.59 (0.38)
Friendship nominations		
Baseline	1.43 (1.28)	1.29 (1.16)
Exit	1.57 (1.16)	2.35 (1.46)
Follow-up	1.36 (1.08)	1.88 (1.54)

SD: standard deviation.

social network inclusion in the RR condition increased from 0.27 at baseline to 0.37 at exit and remained stable at follow-up. On average, children in both conditions remained stable in their friendship nominations from baseline to exit ($F(1,29)=2.90$, $p=0.10$, $ES=0.32$) and there were no significant condition differences from baseline to exit ($F(1,29)=1.52$, $p=0.23$, $ES=0.23$). See Figure 4.

The association between children's engagement on the playground and children's social network inclusion in the classroom was explored using a linear mixed model with children's engagement as a time-varying covariate. See Table 2. There were no statistically significant associations between children's engagement on the playground and social network inclusion.

Discussion

This study compared the effects of training school personnel in RR with and without implementation support for elementary-aged children with autism in public schools. Consistent with previous studies of RR, children with autism improved their playground peer engagement (Kretzmann et al., 2015); however, providing implementation support was associated with greater improvements in social network inclusion and friendship nominations than the condition with no implementation support. These results are surprising given that fidelity remained low throughout the study period.

Contrary to our hypotheses, the RR-only condition had significantly higher fidelity than the implementation support condition, albeit both conditions had low fidelity. One explanation for this difference is that the RR with implementation support condition had more teachers ($n=8$) as intervention agents than the RR-only condition ($n=3$). Teachers are not as frequently present on the playground in comparison to other staff (one-to-one or classroom assistants) who are present on the playground every day. As such, teachers may not have had as many opportunities to use or practice RR, resulting in slightly poorer fidelity. However, teachers may have more opportunity for skills generalization, which may explain the difference in the observed social outcomes discussed below. In addition, low implementation fidelity may suggest that it is too difficult to implement RR in under-resourced public schools (Schoenwald et al., 2011). Schools may need to adapt RR to create a better fit within their contexts. For example, schools may need to consider how the application of RR complements existing programs for children with autism and works within their current staff infrastructure. However, it also is likely that low fidelity is a product of implementation in real-world, community-based settings, which often have less-than-optimal delivery of standardized intervention protocols (Southam-Gerow et al., 2010). The results of this study are consistent with the literature on real-world randomized effectiveness trials with elementary-aged children with autism in under-resourced school settings (Locke et al., 2015; Mandell et al., 2013; Pellecchia et al., 2015; Stahmer et al., 2015). The standard of care during recess in these schools did not include active facilitation to foster opportunities for social development for children with autism prior to receiving training. Schools in this study went from "non-use" to "low use" instead of "good use"

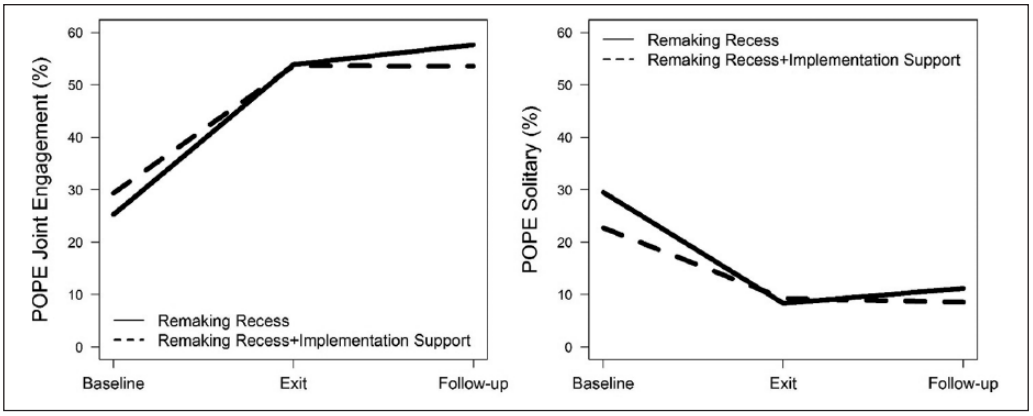

Figure 3. Percentage of time children spent in joint and solitary engagement by condition.

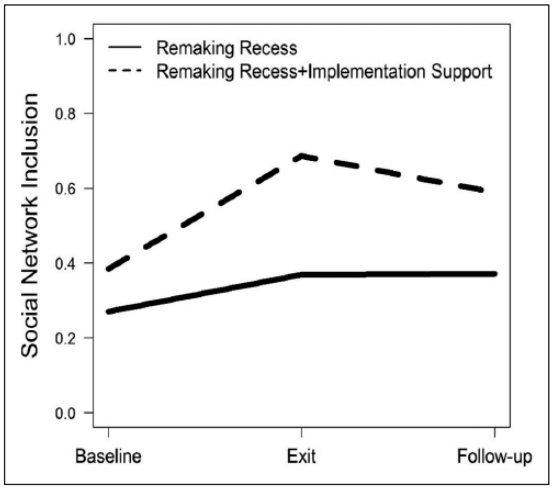

Figure 4. Social network inclusion by condition.

to "great use". This distinction is important as the schools in which this study was conducted were situated in under-resourced districts, and a slight increase in the use of RR may have been sufficient to show improvements in outcomes. The expectation of achieving the same fidelity that experts in research-based settings obtain may be unrealistic given the barriers to implementation (e.g. resources, staffing, time, etc.) of psychosocial interventions that schools face (Forman et al., 2009; Massey et al., 2005). Kendall and Beidas (2007) recommend that there be flexibility within fidelity to ease the transition and dissemination of evidence-based interventions from research to practice. It also is possible that we have not yet isolated the active ingredients of RR or other social engagement interventions, which means that the fidelity measures were not capturing the most important intervention components. Additional research surrounding the development and testing of effective and efficient fidelity measurement instruments in community-based settings such as schools is warranted (Schoenwald, 2011; Schoenwald et al., 2011).

Consistent with our hypotheses, social network inclusion and received friendship nominations improved more in the RR with school-level implementation support condition. While children with autism were able to engage with their peers on the playground during recess particularly when a structured activity was facilitated, they may not have established relationships with those peers in the same ways as when school-level implementation support was provided. There is some evidence that sociometric measures (social network ratings) are more resistant to change (Mikami et al., 2013a) as well as some evidence that changing the context in which children with autism are included (i.e. peer-mediated interventions) is effective in changing peer relationships (Kasari et al., 2012). Data from this study suggest that changing complex, dynamic, and relational social outcomes (social network inclusion and friendship nominations) in the classroom may require broad system-wide support in addition to specific training in the intervention to individual providers. Providing implementation support may have resulted in a subtle but more pervasive use of RR throughout the day for children with autism as well as positive collateral effects for the other faculty, staff, and children in the school (Mikami et al., 2013b). In the RR with implementation support condition, more school personnel (on average three school personnel), specifically teachers compared to other staff, were involved as opposed to the RR-only condition (on average two school personnel). This coupled with the regular presence of the research staff and principal buy-in, who spearheaded the implementation support teams, may have kept participants focused on facilitating peer engagement and social network inclusion.

School-level factors (e.g. implementation climate, culture, and support) also may contribute to and enhance the successful use of RR. Embedding and raising awareness of RR among the faculty, staff, and students may emphasize the importance of socialization throughout the school culture. The RR with implementation support condition incorporated and involved the administrators (leadership

team), faculty, staff, and students, which permeated the school culture and may explain the observed improvements in social network inclusion and friendship nominations. Addressing school-level factors to facilitate the successful implementation of RR may change the context in which relationships are developed and maintained in the classroom and complement the ways in which RR changes the playground context. It also may be that implementation support around the content (e.g. RR) may have had subtle but important effects on a larger group of staff's behavior that resulted in much more positive child outcomes similar to other organizational interventions such as Availability, Responsiveness, and Continuity on the use of evidence-based practices in community-based mental health services (Glisson et al., 2010). Future research should measure implementation fidelity among all school personnel and not just the designated implementer.

Limitations

Despite a number of study strengths, including its community relevance, inclusion of a relatively racially/ethnically diverse sample and its grounding in a practical setting (i.e. public schools; Jonsson et al., 2016), several limitations should be noted. First, the relatively small sample size and wide age range of participating children with autism and schools affect the generalizability of the findings and do not allow for adequate statistical power to explore potential moderators (e.g. age, sex, and ethnicity). Second, these data were gathered in the Northeastern United States, which limits the geographic generalizability of the results to areas with similar racial/ethnic diversity and socioeconomic status as well as school characteristics. Third, the sample was predominantly male. Additional research on females with autism is needed to understand potential sex differences as females with autism have shown differences in socialization patterns in comparison to males with autism in school (Dean et al., 2014). Fourth, due to the logistical and practical constraints of the research study, independent assessments to confirm an autism classification and IQ scores for research eligibility were not conducted. A conscious decision was made to use school documentation to ensure that children had an autism diagnosis and an IQ above 65. Baseline assessments on the POPE were carefully examined and children with autism showed significant social impairments well below the cut point that differentiates children with and without autism (Locke et al., 2016). While the POPE is not a diagnostic assessment or screening tool, the results suggest that the children with autism in this study had significant social impairments. We also ensured that all participants were performing at grade level, which we used as a proxy for cognitive functioning. Despite this, it is important to note that due to the unconfirmed autism diagnosis and IQ scores and reliance on school records, generalizability to the larger population of children with autism, especially those with an IQ below 65, may not be valid as social difficulties may arise for a variety of reasons beyond autism. We also recognize that RR may not be appropriate for many children with autism who do not have the verbal language abilities to independently engage with others without assistive technologies and communication devices, which further limits the generalizability of our study findings. Fifth, many school personnel did not return their self-rated fidelity at the 6-week follow-up; thus, school personnel's self-rated follow-up data were not included in the analyses. Furthermore, the follow-up period was limited to 6 weeks to fit within the school calendar year and did not allow the examination of sustainment of implementation on child outcomes into subsequent school years. In addition, the observed improvements in this study may be due to maturation as a true control group was not included in this study. Our school partners deemed it would be unethical to have a true control group and withhold RR from children with autism due to its potential benefits. Previous research suggests that a true control group, where no intervention is delivered, yields little change in outcomes (Kasari et al., 2012), so a conscious decision was made to provide RR to all study participants while manipulating implementation support to appease our school partners. Finally, although school-level implementation support was provided within the same prescribed and systematic approach and qualitatively documented across all schools that allowed room for flexibility and individualization to each school context (Kendall and Beidas, 2007), fidelity of school-level implementation support including a record of homework completion or practice and its impact on observed outcomes was not quantitatively measured.

Conclusion

Despite these limitations, there are important implications related to these findings. The results suggest that (1) RR may improve child outcomes even in under-resourced settings with low implementation fidelity and (2) implementation strategies may improve child outcomes but not implementation fidelity. Together, these data suggest the importance of implementation strategies and a need for further research on understanding the mechanisms behind how they work.

Acknowledgements

We thank the children, staff, and schools who participated in this study as well as Mark Kretzmann and Connie Kasari for their support.

Declaration of conflicting interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by the Autism Science Foundation (Grants #13-ECA-01L) and FARFund Early Career Award, as well as NIMH K01MH100199 (Locke).

ORCID iD

Jill Locke https://orcid.org/0000-0003-1445-8509

References

Allison PD (2012) Handling missing data by maximum likelihood. SAS Global Forum Paper 312-2012. Available at: http://www.statisticalhorizons.com/wp-content/uploads/MissingDataByML.pdf

Bauminger N, Solomon M and Rogers SJ (2010) Predicting friendship quality in autism spectrum disorders and typical development. *Journal of Autism and Developmental Disorders* 40: 751–761.

Cairns R and Cairns B (1994) *Lifelines and Risks: Pathways of Youth in Our Time*. New York: Cambridge University Press.

Carter EW and Kennedy CH (2006) Promoting access to the general curriculum using peer support strategies. *Research & Practice for Persons with Severe Disabilities* 31: 284–292.

Carter EW, Sisco LG, Chung Y, et al. (2010) Peer interactions of students with intellectual disabilities and/or autism: a map of the intervention literature. *Research & Practice for Persons with Severe Disabilities* 35: 63–79.

Dean M, Kasari C, Shih W, et al. (2014) The peer relationships of girls with ASD at school: comparison to boys and girls with and without ASD. *Journal of Child Psychology and Psychiatry* 55: 1218–1225.

Dingfelder HE and Mandell DS (2011) Bridging the research-to-practice gap in autism intervention: an application of diffusion of innovation theory. *Journal of Autism and Developmental Disorders* 41: 597–609.

Durlak JA and DuPre EP (2008) Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology* 41(3–4): 327–350.

Erchul WP and Martens BK (2012) *School Consultation: Conceptual and Empirical Bases of Practice*. 3rd ed. New York: Springer Science + Business Media LLC.

Fixsen DL, Naoom SF, Blasé KA, et al. (2005) *Implementation Research: A Synthesis of the Literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

Forman SG and Barakat NM (2011) Cognitive-behavioral therapy in the schools: bringing research to practice through effective implementation. *Psychology in the Schools* 48: 283–296.

Forman SG, Olin SS, Hoagwood KE, et al. (2009) Evidence-based interventions in schools: developers' views of implementation barriers and facilitators. *School Mental Health* 1: 26–36.

Forman SG, Shapiro ES, Codding RS, et al. (2013) Implementation science and school psychology. *School Psychology Quarterly* 28: 77–100.

Frankel FD, Gorospe CM, Chang Y, et al. (2011) Mothers' reports of play dates and observation of school playground behavior of children having high-functioning autism spectrum disorders. *Journal of Child Psychology and Psychiatry* 52: 571–579.

Glisson C, Schoenwald SK, Hemmelgren A, et al. (2010) Randomized trial of MST and ARC in a two-level EBT implementation strategy. *Journal of Consulting and Clinical Psychology* 78: 537–550.

Harris PA, Taylor R, Thielke R, et al. (2009) Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics* 42: 377–381.

Hoagwood K and Johnson J (2003) School psychology: a public health framework I. From evidence-based practices to evidence based policies. *Journal of School Psychology* 41: 3–21.

Howlin P, Gordon RK, Pasco G, et al. (2007) The effectiveness of picture exchange communication system (PECS) training for teachers of children with autism: a pragmatic, group randomised controlled trial. *Journal of Child Psychology and Psychiatry* 48: 473–481.

Jonsson U, Olsson NC and Bölte S (2016) Can findings from randomized controlled trials of social skills training in autism spectrum disorder be generalized? The neglected dimension of external validity. *Autism* 20: 295–305.

Kamps D, Thiemann-Bourque K, Heitzman-Powell L, et al. (2014) A comprehensive peer network intervention to improve social communication of children with autism spectrum disorders: a randomized trial in kindergarten and first grade. *Journal of Autism and Developmental Disorders* 45(6): 1809–1824.

Kasari C, Dean M, Kretzmann M, et al. (2016) Children with autism spectrum disorder and social skills groups at school: a randomized trial comparing intervention approach and peer composition. *Journal of Child Psychology and Psychiatry* 57: 171–179.

Kasari C, Rotheram-Fuller E and Locke J (2005) *The Development of the Playground Observation of Peer Engagement (POPE) Measure*. Los Angeles, CA: University of California, Los Angeles.

Kasari C, Rotheram-Fuller E, Locke J, et al. (2012) Making the connection: randomized controlled trial of social skills at school for children with autism spectrum disorders. *Journal of Child Psychology and Psychiatry* 53: 431–439.

Kendall PC and Beidas RS (2007) Smoothing the trail for dissemination of evidence-based practices for youth: flexibility within fidelity. *Professional Psychology: Research and Practice* 38: 13–20.

Kretzmann M, Locke J and Kasari C (2012) Remaking Recess: the manual. Unpublished manuscript funded by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) under grant number UA3 MC 11055 (AIR-B). Available at: http://www.remakingrecess.org

Kretzmann M, Shih W and Kasari C (2015) Improving peer engagement of children with autism on the school playground: a randomized controlled trial. *Behavior Therapy* 46: 20–28.

Lane KL, Oakes WP, Carter EW, et al. (2015) Examining behavior risk and academic performance for students transitioning from

elementary to middle school. *Journal of Positive Behavior Interventions* 17: 39–49.

Langley AK, Nadeem E, Kataoka SH, et al. (2010) Evidence-based mental health programs in schools: barriers and facilitators of successful implementation. *School Mental Health* 2: 105–113.

Licciardello CC, Harchik AE and Luiselli JK (2008) Social skills intervention for children with autism during interactive play at a public elementary school. *Education and Treatment of Children* 31: 27–37.

Locke J, Olsen A, Wideman R, et al. (2015) A tangled web: the challenges of implementing an evidence-based social engagement intervention for children with autism in urban public school settings. *Behavior Therapy* 46: 54–67.

Locke J, Shih W, Kretzmann M, et al. (2016) Examining playground engagement between elementary school children with and without autism spectrum disorder. *Autism* 20: 653–662.

Lyon AR, Ludwig K, Romano E, et al. (2014) Using modular psychotherapy in school mental health: provider perspectives on intervention-setting fit. *Journal of Clinical Child and Adolescent Psychology* 43: 890–901.

McConachie H and Fletcher-Watson S (2015) Building capacity for rigorous controlled trials in autism: the importance of measuring treatment adherence. *Child Care, Health and Development* 41: 169–177.

McLeod BD, Southam-Gerow MA, Bair CE, et al. (2013) Making a case for treatment integrity as a psychological treatment quality indicator. *Clinical Psychology: Science and Practice* 20: 14–32.

Mandell DS, Stahmer AC, Shin S, et al. (2013) The role of treatment fidelity on outcomes during a randomized field trial of an autism intervention. *Autism* 17: 281–295.

Massey OT, Armstrong K, Boroughs M, et al. (2005) Mental health services in schools: a qualitative analysis of challenges to implementation, operation, and sustainability. *Psychology in the Schools* 42: 361–372.

Mendelson JL, Gates JA and Lerner MD (2016) Friendship in school-age boys with autism spectrum disorders: a meta-analytic summary and developmental, process-based model. *Psychological Bulletin* 142: 601–622.

Mikami AY, Griggs MS, Lerner MD, et al. (2013a) A randomized trial of a classroom intervention to increase peers' social inclusion of children with attention-deficit/hyperactivity disorder. *Journal of Consulting and Clinical Psychology* 81: 100–112.

Mikami AY, Reuland MM, Griggs MS, et al. (2013b) Collateral effects of a peer relationship intervention for children with attention deficit hyperactivity disorder on typically developing classmates. *School Psychology Review* 42: 458–476.

Owens JS, Lyon AR, Brandt NE, et al. (2014) Implementation science in school mental health: key constructs in a developing research agenda. *School Mental Health* 6: 99–111.

Pellecchia M, Connell JE, Beidas RS, et al. (2015) Dismantling the active ingredients of an intervention for children with autism. *Journal of Autism and Developmental Disorders* 45: 2917–2927.

Powell BP, Waltz TJ, Chinman MJ, et al. (2015) A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implementation Science* 10: 1–15.

Proctor EK, Landsverk J, Aarons GA, et al. (2009) Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. *Administration and Policy in Mental Health and Mental Health Services Research* 36: 24–34.

Proctor E, Silmere H, Raghavan R, et al. (2011) Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health* 38(2): 65–76.

Proctor EK, Powell BJ and McMillen JC (2013) Implementation strategies: recommendations for specifying and reporting. *Implementation Science* 8: 1–11.

Robertson K, Chamberlain B and Kasari C (2003) General education teachers' relationships with included students with autism. *Journal of Autism and Developmental Disorders* 33: 123–130.

Rotheram-Fuller E, Kasari C, Chamberlain B, et al. (2010) Grade related changes in the social inclusion of children with autism in general education classrooms. *Journal of Child Psychology and Psychiatry* 51: 1227–1234.

Schoenwald SK (2011) It's a bird, it's a plane, it's … fidelity measurement in the real world. *Clinical Psychology Research and Practice* 18: 142–147.

Schoenwald SK, Garland AF, Chapman JE, et al. (2011) Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health* 38: 32–43.

Southam-Gerow MA, Weisz JR, Chu BC, et al. (2010) Does cognitive behavioral therapy for youth anxiety outperform usual care in community clinics? An initial effectiveness test. *Journal of the American Academy of Child & Adolescent Psychiatry* 49: 1043–1052.

Stahmer AC, Reed S, Lee E, et al. (2015) Training teachers to use evidence-based practices for autism: examining procedural implementation fidelity. *Psychology in the Schools* 52: 181–195.

Suhrheinrich J, Stahmer AC and Schreibman L (2007) A preliminary assessment of teachers' implementation of pivotal response training. *Journal of Speech and Language Pathology—Applied Behavior Analysis* 2: 1–13.

Tao J, Kiernan K and Gibbs P (2015) Advanced techniques for fitting mixed methods models using SAS/STAT® software. SAS Paper 1919-2015. Available at: https://support.sas.com/resources/papers/proceedings15/SAS1919-2015.pdf