Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-019-04304-6

ORIGINAL PAPER

Trajectories of Evidence Based Treatment for School Children with Autism: What's the Right Level for the Implementation?

Victor Lushin[1,5] · David Mandell[2] · Rinad Beidas[2] · Steven Marcus[2] · Heather Nuske[2] · Victor Kaploun[3] · Max Seidman[2] · Daphney Gaston[2] · Jill Locke[4]

Abstract

Evidence-based practices (EBP) for children with autism are under-used in special-education schools. No research compared child-level versus teacher-level influences on EBP use, which could guide implementation strategies. We derived longitudinal profiles of EBP receipt by children (N = 234) in 69 autism-support classrooms, over an academic year. We compared overall impacts of child-level and teacher-level factors on profile membership. Most children received little EBP throughout the year; however substantial subgroups received increasing, and decreasing, doses of EBP. Child-level and teacher-level factors contributed about equally to profile membership. Children's autism symptoms and verbal ability, teachers' EBP skills, training/experience, classroom support, class size, and implementation leadership climate predicted profile membership. Early identification of treatment profiles could facilitate targeted implementation strategies increasing EBP use.

Keywords Autism · Evidence based practices · Special education

Background

Evidence-based practices (EBP) for treating autism spectrum disorder (ASD) often don't make their way into schools, where most children with ASD receive care (Locke et al. 2017; Stichter et al. 2016). These practices often require intensive, extended implementation support to be effective (Locke et al. 2017). The importance of continuous support and fidelity monitoring for the sustainment of EBPs has been emphasized in implementation science literature, including in the Exploration, Preparation, Implementation, and Sustainment (EPIS) framework (Aarons et al. 2011). EPIS framework recommends public-academic collaborations as a mechanism to support sustaining EBP use in service settings (Aarons et al. 2011). Little empirical evidence, however, suggests how to best organize sustainment support over time, and how implementer and client characteristics shape sustainment needs. For more than a decade, our university based team of coaches has supported special education teachers in EBP for students with ASD. Over that

✉ Victor Lushin
onishulo@upenn.edu; Victor.Lushin@LIU.edu

David Mandell
mandelld@upenn.edu

Rinad Beidas
rbeidas@upenn.edu

Steven Marcus
marcuss@upenn.edu

Heather Nuske
hjnuske@upenn.edu

Victor Kaploun
kaploun@eu.spb.ru

Max Seidman
seidmanmax@gmail.com

Daphney Gaston
seidmanmax@gmail.com

Jill Locke
jjlocke@uw.edu

[1] Long Island University, Brooklyn, NY, USA

[2] University of Pennsylvania, Philadelphia, PA, USA

[3] National Research University Higher School of Economics, St. Petersburg, Russia

[4] University of Washington, Seattle, WA, USA

[5] Long Island University Brooklyn, School of Health Professions, 1 University Plaza, Room 246, Brooklyn, NY 11201, USA

Published online: 03 December 2019

time, we have found that teachers vary considerably in their willingness and ability to implement EBP the way they were designed, and that their use of these practices changes over the course of the year. These anecdotal observations have not been tested in quantitative analysis, however.

Studying these patterns of implementation over the course of the year and teacher and child characteristics associated with different patterns of implementation may have implications for how to best support teachers in using EBP. For example, if certain times during the year are characterized by poor implementation, it may suggest the need for concentrated support at those times. Teachers who decrease their use of EBP over time may need different messages and support than teachers who increase their use over time to maximize implementation. Most studies of EBP use for children with ASD are cross-sectional (Cook and Odom 2013). Studies that account for EBP use over time tend to average EBP use across multiple time points, rather than take advantage of the multiple data points (Howlin et al. 2009).

Additionally, little is known about the relative contributions of child and provider characteristics to the use of evidence-based treatment for children with autism. Some teachers have told us that they don't use an intervention because they think it is the wrong intervention for a particular child, as opposed to considering it a bad fit for their classroom in general; this anecdotal information has not been formally tested yet. Parsing out what portion of variance in implementation is due to each of these two perceptions may have important implications for how administrators and coaches message the importance of EBP use.

The present study addresses these gaps by pursuing three aims: (a) examining changes in the intensity of teachers' use of EBP with individual children over an academic year; (b) estimating the proportion of variance in EBP use that can be attributed to overall child-level or teacher-level characteristics; and (c) examining the association of specific teacher and classroom characteristics, and child characteristics, with trajectories of EBP use. Our outcomes of interest are fidelity to two evidence-based one-to-one interventions for children with autism: discrete trial training (DTT; Smith 2001) and pivotal response training (PRT; Koegel and Koegel 2006).

Method

Procedure and Participants

Data came from a study conducted in 69 K-2 autism support classrooms in Philadelphia with children (n = 234) aged 5 to 8 (Pellecchia et al. 2016). In the year from which data was provided for the current study, all teachers were trained in an intervention package that included DTT and PRT. Most students served by the school district are ethnic minorities (69%); 75% live below the poverty line. Teachers (N = 69) were recruited from 69 autism support classrooms that had an average of 8 students. The procedures were approved by the university Institutional Review Board and the research office of the School District of Philadelphia.

Throughout the year, teachers received coaching in using DTT and PRT. Coaches encouraged teachers to use both methods for each child instead of picking one method over another.

Measures

Dependent Variable

Intensity of DTT and PRT The dependent variable of interest was teachers' report of how often the teacher used DTT and PRT with each child in their classroom during the past week, measured on a 0 to 4 scale, where 0 = less than once a week, 1 = once a week, 2 = from two to four times a week, 3 = once a day, and 4 = twice a day. Data were collected monthly during the academic year (8 time-points total) by coaches who visited participating classrooms. Teachers were coached to conduct each DTT and PRT session for at least 20 min, and to report only sessions of this, or longer, duration. To minimize social desirability bias, coaches were trained to form cohesive, honest and mutual relationships with teachers, devoid of power differentials. Additionally, coaches observed and took notes on each teacher's work with each child. The coach and teacher discussed teachers' intensity reports that did not converge with the coach's case observations until they came to consensus. Previous research (Pellecchia et al. 2015) has found strong convergence ($r = 0.73$) between the teacher-reported measure of intensity (used in the present study), and observation (Pellecchia et al. 2015). We also estimated test–retest reliability of teachers' intensity reports by correlating teacher-reported scores of DTT and PRT intensity between two adjacent time points at which the least external influences on the teachers' intensity of EBP were expected, because of the year's lowest level of coaching activity. We found high correlations ($r = 0.7$–0.8) between these adjacent intensity scores.

Independent Variables

Teacher's Intentions to Provide DTT and PRT Intentions, a mental representation of a commitment to carrying out an action, are an important predictor of behavior (Jaccard and Levitz 2015). A weak intent-behavior link may suggest a need to account for contextual and skill-related factors, among others, in order to increase implementation. Intentions were measured by one item for each (Fishbein and Ajzen 2011). Teachers evaluated the likelihood of their use of DTT, and separately, of PRT at least once a day, in the following 4–6 weeks. The scale ranged from 0 = *very unlikely* to 4 = *very likely.* We did not include a more generic measure of intent to use evidence-based practices, because multiple studies have shown that behavior-specific measures of intentions are uniformly much stronger predictors of behavior (e.g. Jaccard and Levitz 2015). A new psychometric study based on the same data set as the present research (Fishman et al. under review) shows that behavior-specific measures of EBP intentions have considerably stronger predictive power compared to more generic EBP intentions measures. The same study also demonstrates that for some EBPs, a single-item measure of intent to implement has predictive validity on the par with that of a three-item intent scale (Fishman et al. under review).

Teacher's Baseline DTT and PRT Skills Skill is among the modifiable moderators of the translation of intentions into successful behavior (Jaccard and Levitz 2015; Fishbein and Ajzen 2011). The skill level, or accuracy of the teachers' performance of DTT and PRT techniques was measured via hour-long direct observations. Trained research assistants recorded the occurrence or non-occurrence of each specified component of DTT and PRT on a data collection form (Pellecchia et al. 2015). Components were specific to DTT (e.g. prompting and reinforcement strategies) and to PRT (e.g. providing the child with a choice of tasks). The research assistants were trained to reliability with the doctoral-level supervising research staff. Training of research assistants involved didactic instruction on data coding procedures, role-playing, and practice coding. Training continued until research assistants were 80% reliable with the supervising research staff, as recommended by the previously published training protocol (Pellecchia et al. 2015). Teachers' performance of each DTT and PRT component was measured by a 0 to 4 observation scale: 0 = teacher does not implement correct intervention technique during the session; 1 = teacher occasionally implements, but misses the majority of opportunities, 2 = teacher sometimes implements, but misses many opportunities, 3 = teacher often implements, but misses some opportunities, and 4 = teacher always implements technique during appropriate opportunities. The items measured implementation of 16 components of DTT, and 17 components of PRT, and were averaged across components. In prior studies using video recordings of DTT and PRT sessions, this observational scale has demonstrated inter-rater agreement over 80% (Stahmer et al. 2015). Measurement of teacher skills occurred after the initial DTT and PRT coaching sessions.

Teacher Characteristics Teachers' special education experience and prior training in ASD intervention are possible sources of "tacit knowledge" (Nonaka and Von Krogh 2009; Polanyi 1966, 2009) that may influence EBP use, and may be difficult to measure by a structured observational protocol. At baseline, teachers were asked about the number of years they had worked in special education, whether they received training in ASD intervention, as well as their age, gender, race and ethnicity.

Classroom/Teaching Team Characteristics Smaller numbers of children in classes and more support staff may increase the capacity for EBP implementation. Research staff measured, through direct observation: number of children in each classroom, and number of classroom support staff; each variable was averaged across data collection time points.

Implementation Leadership A growing literature has highlighted the relevance of organizational leadership for implementing EBP (Aarons et al. 2014). Teachers completed the Implementation Leadership Scale (ILS; Aarons et al. 2014), which measures various dimensions of implementation leadership in a respondents' organization, with response anchors ranging from 0 = *not at all* to 4 = *to a very great extent.* Examples of items included: [Leadership in my organization] has developed a plan to facilitate implementation of EBP; is knowledgeable about EBP; supports employee efforts to learn more about EBP; and perseveres through the ups and downs of implementing EBP. Items did not reflect leadership specific to DTT or PRT.

Implementation Climate Studies have highlighted the role of implementation climate, usually operationalized as providers' perception of openness to innovation within an organization (Ehrhart et al. 2014). Teachers completed the Implementation Climate Scale (ICS; Ehrhart et al. 2014), which measures several dimensions of implementation climate in a respondents' organization, with response anchors ranging from 0 = *not at all* to 4 = *to a very great extent.* Examples of items included: (to what extent your organization) is open to new interventions; values using EBP; and promotes using EBP as a top priority. Items did not reflect implementation climate specific to DTT or PRT.

Child Characteristics: Autism Symptoms These were measured using the teacher version of the Pervasive Developmental Disorders Behavior Inventory (PDDBI; Cohen and Sudhalter 2005). This questionnaire is designed to measure both adaptive skills and maladaptive behaviors relevant to ASD. It provides a quantitative assessment of the severity of a child's ASD symptomology, as compared to other children with ASD. We used standard scores from the two subscales of ASD symptoms, social approach abilities, and sensory symptoms, which were associated with DTT and PRT use in a prior study (Nuske et al. 2019). For the social approach subscale, higher scores indicate better ability; for the sensory symptoms, higher scores indicate more problems. Teachers completed a Teacher PDDBI form for each consented student in their classroom a minimum of 1 month after the beginning of the school year, therefore they had adequate time to become familiar with the child's ASD symptom presentation prior to filling out this measure.

Child Characteristics: Language Ability The Early Years Battery of the Differential Abilities Scales 2nd Edition (DAS-II) was used as a clinical assessment of a child's cognitive abilities (Beran 2007; Elliott 1990; Marshall et al. 2011), administered by researchers at the beginning of the school year. The present study used the DAS-II Verbal Ability subscale that had been associated with DTT and PRT use in the prior study (Nuske et al. 2019). Other studies have similarly chosen to use the DAS-II as the outcome measure in studies of children with ASD (Anderson et al. 2007; Thurm et al. 2007).

Analytic Strategy

The primary analytic goal of this study was to examine how the frequency of DTT and PRT use for each child changed over the academic year, across eight monthly time points. To identify distinct clusters of children with similar trajectories of DTT and PRT use within clusters, we relied on transitional cluster analysis (Schulenberg et al. 1996) via latent profile analysis (LPA; Asparouhov and Muthén 2014). We examined longitudinal patterns, across eight monthly time points, of the use of DTT for each child, and separately, of the use of PRT for each child, nested in classrooms. Exploratory model building began with a one-class model, continuing until fit indices suggested further classes did not improve fit. Consistent with guidelines (Nylund et al. 2007), final model selection was determined through a combination of fit indices [Bayesian information criterion (BIC), Akaike information criterion (AIC), Vuong-Lo-Mendell-Rubin likelihood ratio test (VLMR LRT), and the Lo-Mendell-Rubin Adjusted LRT test] along with parsimony, interpretability, and generalizability. All analyses were conducted using Mplus with robust (Huber-White) maximum likelihood algorithms (Lai 2018). Each resulting profile represented trajectories of use of DTT (Fig. 1) and PRT (Fig. 2) for each child.

We then estimated the overall contribution of (1) teacher and classroom characteristics, and (2) child characteristics, to children's trajectory-profile membership by deriving intraclass correlations (ICC), an established technique in multilevel data analysis (Castro 2002; Preacher et al. 2011). ICC compares the portions of outcome variation that occur on different levels of data organization (e.g. group level variation vs. individual level variation). In our case, ICC analyzed the variation of children's membership in different trajectories, and compared the magnitude of this variation between classroom groups versus among individual children. The proportion of classroom-level

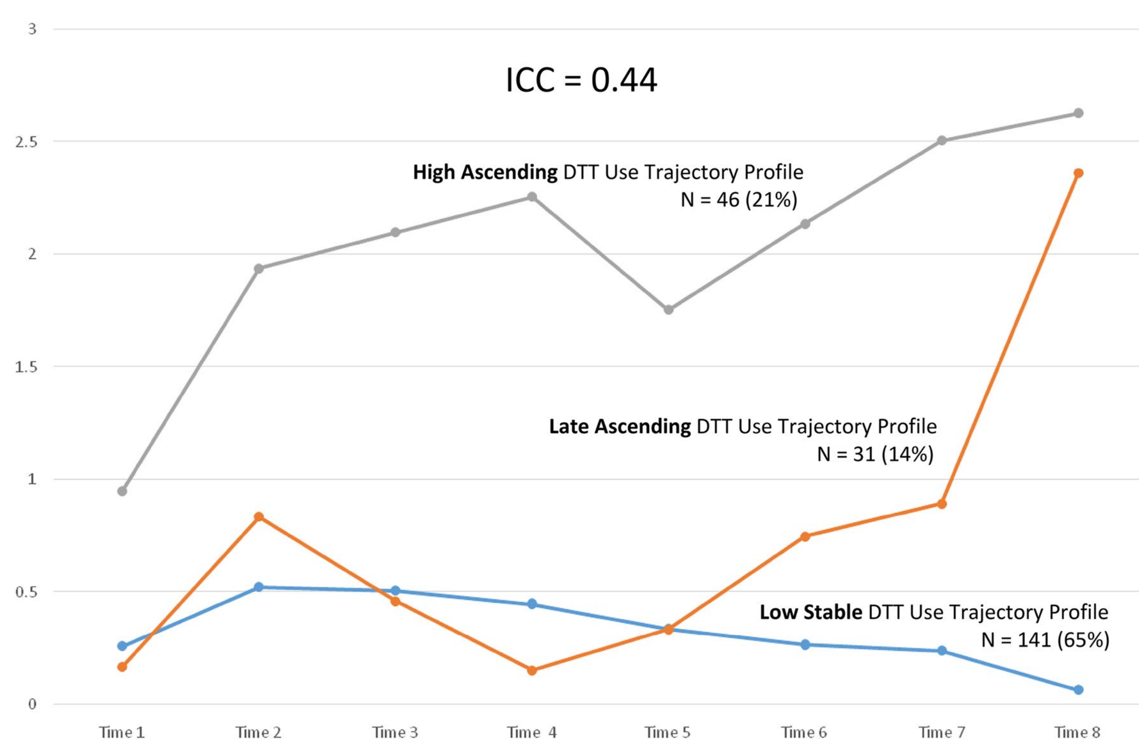

Fig. 1 Trajectory profiles of the average DTT use intensity across the 8 time points

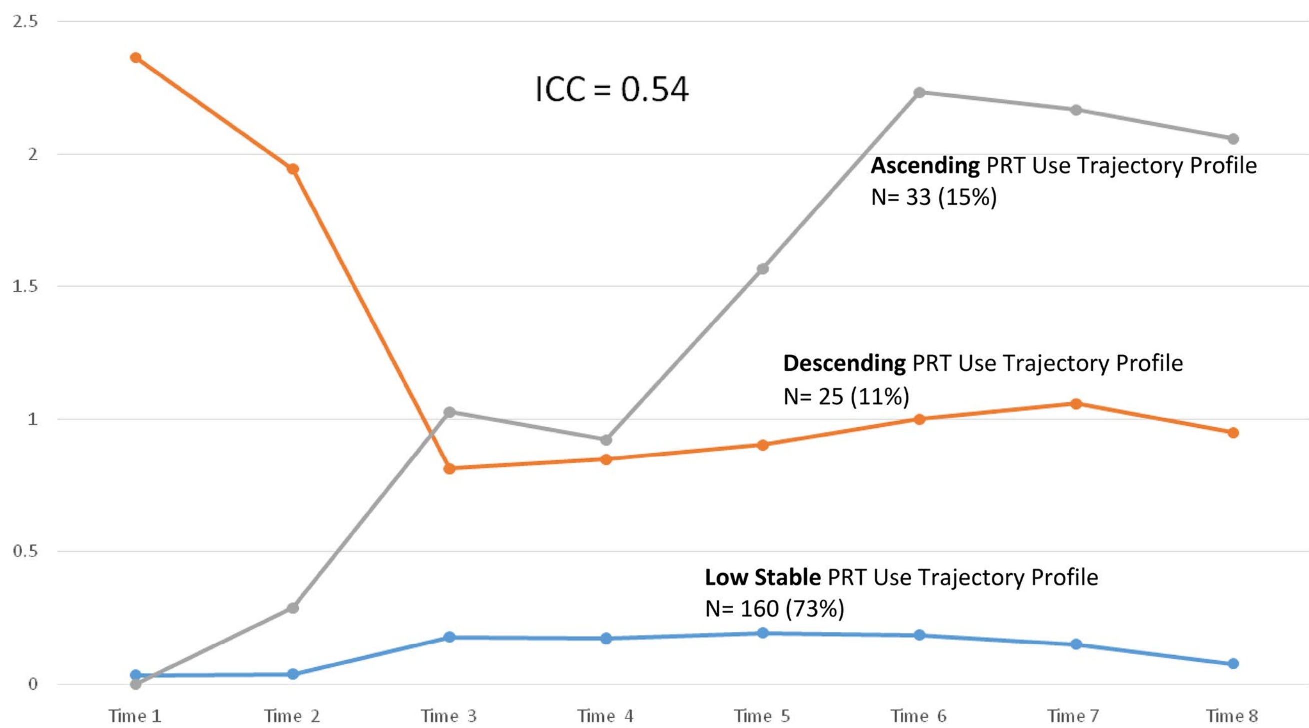

Fig. 2 Trajectory profiles of the average PRT use intensity across the 8 time points

variation in the mix represents the overall impact of teacher characteristics, because each teacher is specific to a classroom. The proportion of child-level variation represents the overall impact of child characteristics. ICC method does not compare the contributions to the outcome variance from specific teacher/classroom variables or child variables per se.

As a secondary analysis, we examined the associations between classroom and child characteristics and trajectory profile membership, using multi-level, multinomial logistic regression (MLR) to account for the nested nature of the data. Regression models controlled for teacher age and ethnicity/race, and provided results in the form of odds ratios. In our MLR analysis, continuous predictor variables, e.g., number of years spent by teachers in Special Education, or the child's DAS Verbal Subscale score, were not dichotomized and were used in their original, continuous form, following recommendations by Jaccard (2001). In this approach, an odds ratio indicates by how much, on average, each step of the continuous predictor variable (e.g. each year of special education experience) changes the likelihood of the child entering a "higher" EBP use trajectory, instead of the "Low Stable" trajectory (e.g., "For each additional step of the scale, the likelihood of entering High Ascending trajectory decreases by…"). The potential multiple statistical test error was addressed by Holm-modified Bonferroni sequential rejective multiple test procedure (Holm 1979).

Results

Table 1 displays the sample characteristics. Teachers' mean age was 37; they averaged 8.4 years of special education experience, and 68% had received some training in ASD interventions; 97% was female; therefore teachers' gender was not used as a covariate. Most (85%) teachers identified as white. Teachers' race/ethnicity was dichotomized as White versus Non-White.

Trajectories of DTT and PRT Use

Profiles

Latent profile analyses for trajectories of DTT and PRT use was conducted for 1–5 profile models. Table 2 displays fit indices for 2–4 profile solutions. For DTT trajectory profiles, a 3-profile model yielded better model fit than 2-profile model (AIC = 3661.324, BIC = 3776.396, Entropy = 0.868). Although fit indices for 4-profile model were slightly better than for 3-profile model (Table 2), adjusted likelihood ratio test (LRT) indicated that a 4 profile model did not improve fit beyond a 3 profile model ($p > .05$). For PRT trajectory profiles, a 3-profile model also yielded better model fit than 2-profile model (AIC = 2757.664, BIC = 2872.737, Entropy = 0.942). For the PRT trajectory profiles, fit indices for 4-profile model were slightly better than for 3-profile model (Table 2); however, adjusted LRT indicated that a 4

Table 1 Descriptive statistics and means of key variables

	Mean (SD)	Percent
Average intensity of DTT use, across time points[a]	0.76 (0.80)	
Average intensity of PRT use, across time points[a]	0.42 (0.58)	
Teacher's age	37.3 (10.6)	
Teachers' gender female		97.4
Teachers' race: White		85.0
Teachers' race: Black		10.7
Teachers' race: Asian + others		2.1
Teachers' Latino identity		3
Average class size	7.4 (1.2)	
Average number of classroom support staff	3.7 (1.2)	
Teacher's intentions to provide DTT	2.2 (1.2)	
Teacher's intentions to provide PRT	2.1 (1.3)	
Teacher's DTT skill (DTT fidelity measured at baseline)[a]	1.7 (1.6)	
Teacher's PRT skill (PRT fidelity measured at baseline)[a]	1.4 (1.4)	
Teacher received training in ASD intervention		68.0
Teacher's number of years spent in special education	8.4 (7.1)	
Implementation leadership[a]	2.7 (1.1)	
Implementation climate[a]	3.1 (2.4)	

[a]Intensity of DTT and PRT use, fidelity of DTT and PRT performance, implementation leadership and implementation climate are measured on the 0 to 4 scale

Table 2 Comparative model fit of latent profile solutions

Fit indices	Number of DTT profiles			Number of PRT profiles		
	DTT 2 profiles	DTT 3 profiles	DTT 4 profiles	PRT 2 profiles	PRT 3 profiles	PRT 4 profiles
Loglikelihood	− 1862.090	− 1796.662	− 1717.596	− 1498.259	− 1344.832	− 1195.974
AIC	3774.180	3661.324	3521.192	3046.518	2757.664	2477.948
BIC	3858.792	3776.396	3666.725	3131.131	2872.737	2623.481
Entropy	0.916	0.868	0.777	0.964	0.942	0.950
Adjusted LRT	$p=.10$	$p=.09$	$p=.18$	$p<.01$	$p=.09$	$p=.69$
Bootstrapped LRT	$p<.01$	$p<.01$	$p<.01$	$p<.01$	$p<.01$	$p<.01$

profile model did not improve fit beyond a 3 profile model ($p>.05$). Additionally, 4-profile solutions yielded some profiles with very small sample sizes, making interpretation difficult. We therefore chose 3-profile solutions over 4-profile solutions for both DTT and PRT trajectory profiles.

Figure 1 illustrates the three-profile solution for DTT implementation. The graph represents average DTT intensity at each of the 8 time points for each of the three profiles. Profile 1, termed "Low Stable" comprised the largest proportion of the sample ($n=141$, 64%) and was characterized by low mean DTT implementation at all 8 time points, from 0.26 ± 0.15 (on the metric of 0–4), at Time 1, to 0.52 ± 0.16 at Time 2, and to 0.07 ± 0.05 at Time 8. Profile 2, termed "High Ascending" ($n=46$, 21%) was characterized by an ascending trajectory from low mean DTT implementation at Time 1 (0.95 ± 0.58), with a steep increase by Time 2 (1.94 ± 0.38), ending with a high average DTT use at Time 8 (2.6 ± 0.34). Average DDT use in this profile was significantly higher than in both other profiles at all time points. Profile 3, termed "Late Ascending" ($n=31$, 14%), was characterized by low mean levels of DTT use from Times 1 (0.17 ± 0.32) through 6 (0.75 ± 0.50), followed by a spike at Time 8 (2.4 ± 0.34).

Figure 2 displays the three PRT implementation profiles over the 8 time points. Profile 1, termed "Low Stable", as with DTT model, represented the largest proportion of the sample ($n=160$, 73%); it was characterized by low use of PRT implementation at all 8 time points, from 0.03 ± 0.04 at Time 1, to 0.19 ± 0.13 at Time 5, and to 0.08 ± 0.05 at Time 8. Profile 2, termed "Ascending" ($n=33$, 15%), was characterized by low mean PRT implementation at Time 1 (0 ± 0), with a relatively consistent increase though Time 6 (2.23 ± 0.50), continuing through Time 8 (2.06 ± 0.66). Profile 3 termed, "Descending" ($n=25$, 11%), represented a trajectory starting with high mean intensity of PRT use (2.36 ± 0.4, at Time 1), with a steep decrease to 0.82 ± 0.46, at Time 3, then plateauing with mean PRT use of 0.95 ± 0.52, at Time 8.

ICC Analysis

To estimate the overall contribution of teacher and classroom team characteristics to predicting children's membership in one of the three trajectory profiles, we derived intraclass correlations (ICC) that reflect the portion of overall outcome variance accounted for by group-level variation. ICCs were 0.44 for DTT trajectory profiles and 0.54, for PRT trajectory profiles, indicating that teacher and classroom characteristics account for about a half of the variance in predicting children's profile membership.

Associations of Covariates with Profile Membership

Because most children were in the Low Stable profiles for both DTT and PRT implementation, this group was used as the reference for both DTT and PRT analyses. The upper portion of Table 3 presents odd ratios reflecting differences in teacher/classroom level predictors, and child level predictors, between the High Ascending DTT Trajectory versus the Low Stable DTT Trajectory profiles (left column), and between the Late Ascending DTT Trajectory profile versus the Low Stable DTT Trajectory profiles (right column). The lower portion of Table 3 displays odd ratios representing differences in teacher/classroom level predictors, and child level predictors, between the Ascending PRT Trajectory versus the Low Stable PRT Trajectory profiles (left column) and between the Descending PRT Trajectory profile versus the Low Stable PRT Trajectory profiles (right column).

Teacher/Classroom Level Predictors of DTT Use

Teachers' training in ASD intervention significantly predicted students' membership in the High Ascending DTT trajectory versus the Low Stable DTT trajectory (*OR* 3.70, [95% CI 1.25, 10.97]), as did the number of years teachers spent in special education (*OR* 1.20, [95% CI 1.06, 1.35]). Teachers' DTT skill significantly predicted the High Ascending DTT trajectory vs. the Low Stable DTT trajectory (*OR* 1.45, [95% CI 1.04, 2.04]). Higher ILS (Aarons et al. 2014) predicted fewer children being in the High Ascending DTT trajectory, versus the Low Stable DTT trajectory (*OR* 0.58, [95% CI 0.40, 0.81]). Teachers' assessment of implementation climate via ICS (Ehrhart et al. 2014) was not significantly associated with DTT trajectories. Number of classroom support staff significantly predicted the High Ascending DTT trajectory versus the Low Stable DTT trajectory (*OR* 1.45, [95% CI 1.04, 2.01]); but average class size difference was not statistically significantly associated with DTT trajectories. Teachers' baseline intentions to provide DTT did not significantly predict DTT trajectories. Teachers' age significantly predicted fewer children being on the High Ascending DTT trajectory (*OR* 0.92, [95% CI 0.86, 0.97]). Teachers' racial identity did not predict DTT trajectory. The only statistically significant difference between the Late Ascending DTT use trajectory and the Low Stable DTT use trajectory was teacher's age (*OR* 1.04 [95% CI 1.01, 1.08]) (Table 3).

Child Level Predictors of DTT Use

Children's greater scores on DAS Verbal Subscale (greater verbal ability) and on PDDBI Social Approach Subscale (greater ability for social approach) significantly decreased children's chances of being on the High Ascending versus the Low Stable DTT trajectory (*OR* 0.97 [95% CI 0.95, 0.99] and *OR* 0.99 [95% CI 0.95, 0.99], respectively). It is worth noting that, while these ORs are modest, they represent the average change of trajectory membership likelihood mapped onto only one point of a DAS and PDDBI metric, not onto the whole continuum of the 4-point metric; therefore these ORs should not be interpreted as marginal. Children's greater scores on PDDBI Sensory Symptoms Subscale (greater sensory processing problems) non-significantly predicted the High Ascending DTT trajectory versus the Low Stable DTT trajectory. Child-level characteristics' prediction of the Late Ascending versus the Low Stable DTT trajectory were statistically non-significant, in the same directions as for High Ascending DTT trajectory (Table 3).

Teacher/Classroom Level Predictors of PRT Use

Teachers' training in ASD intervention and the years they spent in special education did not significantly predict differences among PRT use trajectories (Table 3). Teachers' PRT skill significantly predicted the Ascending PRT trajectory vs. Low Stable PRT trajectory (*OR* 1.68 [95% CI 1.06, 2.70]). As with DTT, higher ILS predicted fewer children being in the Ascending PRT trajectory, versus the Low Stable DTT Trajectory (*OR* 0.60, [95% CI 0.38, 0.94]). Teachers' assessment of implementation climate was not significantly associated with PRT trajectories. Average class size significantly predicted less children being on the Ascending PRT trajectory (*OR* 0.56, [95% CI 0.36, 0.86]); but difference in the number of classroom support staff was not significant for PRT trajectories. Teachers' baseline intentions to provide PRT did not significantly predict the Ascending PRT Trajectory versus the Low Stable PRT Trajectory. Teachers of white racial identity had, on average, less children in their care on

Table 3 Multinomial logistic regression predicting DTT and PRT trajectory class membership

\	High ascending DTT use (referent group: low stable DTT use) OR [95% CI]	Late ascending DTT use (referent group: low stable DTT use) OR [95% CI]
Teacher/classroom/organizational characteristics		
Training in ASD intervention	3.70* [1.25–10.91]	1.13 [0.41–3.09]
Years teacher spent in special education	1.20* [1.06–1.35]	1.03 [0.97–1.09]
DTT skill (accuracy measured at baseline)	1.45* [1.04–2.04]	1.01 [0.66–1.52]
Implementation leadership	0.58* [0.40–0.81]	0.62 [0.38–1.01]
Implementation climate	1.02 [0.71–1.33]	0.92 [0.60–1.25]
Average class size	0.88 [0.6–1.27]	1.27 [0.80–2.01]
Number of classroom support staff	1.45* [1.04–2.01]	0.89 [0.50–1.57]
Teacher's intentions to provide DTT	0.76 [0.39–1.50]	0.76 [0.39–1.50]
Teacher's age	0.91* [0.86–0.97]	1.04* [1.01–1.08]
Teachers' race: White	0.36 [0.12–1.10]	0.36 [0.12–1.10]
Child characteristics		
DAS verbal subscale	0.97* [0.95–0.99]	0.98 [0.97–1.01]
PDDBI social approach subscale	0.98* [0.95–0.99]	0.98 [0.95–1.01]
PDDBI sensory symptoms subscale	1.02 [0.99–1.04]	1.01 [0.99–1.04]
	Ascending PRT use (referent group: low stable PRT use) OR [95% CI]	Descending PRT use (referent group: low stable PRT use) OR [95% CI]
Teacher/classroom/organizational characteristics		
Training in ASD intervention	1.00 [0.34–2.89]	0.34 [0.10–1.20]
Years teacher spent in special education	1.03 [0.90–1.18]	0.98 [0.90–1.08]
PRT skill (accuracy measured at baseline)	1.68* [1.06–2.70]	1.21 [0.79–1.87]
Implementation leadership	0.40* [0.23–0.68]	1.24 [0.67–2.30]
Implementation climate	1.05 [0.53–1.57]	0.94 [0.60–1.28]
Average class size	0.56* [0.36–0.86]	0.63* [0.46–0.87]
Number of classroom support staff	1.35 [0.87–2.08]	1.42* [1.05–1.92]
Teacher's intentions to provide PRT	0.91 [0.60–1.39]	0.75* [0.57–0.99]
Teacher's age	0.92* [0.85–0.99]	0.97 [0.96–1.01]
Teachers' race: White	0.16* [0.04–0.69]	0.51 [0.35–3.97]
Child characteristics		
DAS verbal subscale	0.99 [0.97–1.02]	1.00 [0.99–1.02]
PDDBI social approach subscale	0.99 [0.98–1.01]	1.00 [0.98–1.02]
PDDBI sensory symptoms subscale	1.01 [0.99–1.03]	1.03* [1.00–1.05]

PDDBI pervasive developmental disorder behavioral inventory; *DAS* differential ability scales for autism

$^{*}p < .05$

Ascending PRT trajectory vs. Low Stable PRT trajectory (*OR* 0.16, [95% CI 0.04, 0.69]); the difference in teacher age was not significant for PRT trajectories.

Comparisons between Descending PRT use trajectory and Low Stable PRT use trajectory yielded several statistically significant differences. Children in the Descending PRT use trajectory (which started at a relatively high average intensity of PRT use) had, on average, fewer students in the class (*OR* 0.63 [95% CI 0.46, 0.87]) and more classroom staff (*OR* 1.42 [95% CI 1.05, 1.92]). Teachers of children in the Descending PRT Use trajectory profile had, on average, lower intentions to use PRT (*OR* 0.75 [95% CI 0.57, 0.99]) (Table 3).

Child Level Predictors of PRT Use

Children's greater scores on PDDBI Sensory Symptoms Subscale significantly predicted the Descending PRT use trajectory versus the Low Stable PRT trajectory (*OR* 1.03 [95% CI 1.00, 1.05]). Other child-level characteristics' prediction of the Ascending and Descending PRT Trajectories versus the Low Stable PRT trajectory were statistically non-significant, in the same directions as for High Ascending and Late Ascending DTT trajectories (Table 3).

Based on the Holm-modified Bonferroni sequential rejective procedure (Holm 1979), multinomial logistic regression coefficients of DTT skill and DAS Verbal Subscale on High Ascending DTT trajectory vs. Low Stable DTT trajectory were no longer statistically significant; other reported regression coefficients remained statistically significant.

Multinomial logistic regression analyses comparing the High Ascending DTT Trajectory profile to the Late Ascending DTT Trajectory profile, and comparing the Ascending PRT Trajectory profile to the Descending PRT Trajectory profile, did not yield statistically significant regression coefficients. Therefore, these comparisons are not reported.

Discussion

Our study examined latent profiles of children's receipt of DTT and PRT in autism support classrooms during the course of an academic year. The results revealed considerable variability in children's receipt of evidence-based treatment over the course of the year. Throughout the year, more than 30% of children received increasing doses of DTT, some at faster and others at slower pace; 15% of children received increasing doses of PRT. Most children received consistently low doses of DTT and PRT throughout the year. Our findings emphasize that children with ASD and their teachers cluster by differential needs for EBP sustainment support patterned over time.

One can conceptualize the factors that shape teachers' use of EBP for children with autism as falling into two categories. First, the teacher decides whether the particular EBP is appropriate for each child based on the child's characteristics. Second, the teacher and classroom characteristics facilitate or hinder the use of that EBP. Our ICC analyses suggest that child-level and teacher-level factors contribute roughly in equal proportion to determining the child's year-long trajectory of treatment. One major implication is that "half the battle" in increasing implementation of evidence-based treatment for children with ASD may be won by understanding and modifying the teacher's perception of the child's needs and barriers in working with a particular child, such as behavior management issues, or teacher's perception of each child's malleability to evidence-based treatment. For example, the higher severity of children's autism symptoms was linked to their greater likelihood of being on higher trajectories of DTT and PRT use. More specifically, greater scores on adaptive subscales, DAS Verbal Subscale and PDDBI Social Approach Subscale, were associated with lower likelihood of being on the High Ascending DTT trajectory; while greater scores on a maladaptive subscale, PDDBI Sensory Symptoms Subscale, were associated with a greater likelihood of being on the Descending (vs. the Low Stable) PRT trajectory (Table 3). This finding is similar to that in Nuske et al. (2019), where teachers gave more treatment to children with more obvious impairments. While promising in some sense, this also means that the children with less severe symptoms may be inadvertently neglected, although they also are likely to benefit from DTT and PRT (Nuske et al. 2019).

The other half of the "battle" to increase EBP use may be won by implementation strategies that target teacher and classroom characteristics. For example, we found that teachers' prior training in ASD interventions, their EBP skills, as well as the number of children in the classroom and the number of support staff were associated with their use of EBP. These findings suggest that implementation strategies targeting the sustainment of EBP for children with ASD should focus on multiple components of the EPIS framework (Aarons et al. 2011), including (a) inner context modifications such as increasing staff-student ratios, (b) fidelity support by teacher training in ASD treatment, and (c) ongoing coaching to support teachers' understanding of each individual child's need for evidence-based instruction.

Neither half of the "battle" has been won yet. The results of our ICC analysis should be interpreted in the context of the limited uptake of both DTT and PRT in the classrooms where the data were collected. Teacher and classroom variables, and particularly child variables, used in this study, contribute modestly to the trajectory membership. Our study encourages further search for stronger implementation targets, both on teacher-classroom level and among individual child characteristics. More generally, further research is needed to better understand teachers' decision-making about particular interventions being appropriate for either the classroom or the individual child.

Finally, our finding that patterns of treatment intensity throughout the year differ across children, highlights a need for early identification of children and teachers likely to be on ascending, high descending, or consistently low trajectories of treatment. Different trajectory groups may require different supportive coaching approaches (Aarons et al. 2011). For example, for teachers in ascending trajectory clusters, who apparently respond well to the skill-based coaching our consultants provided, more intense coaching earlier in the academic year may be sufficient. For teachers in the descending PRT profile, reassessments of psychological and environmental barriers to PRT use throughout the year, and matching coaching efforts accordingly, may be warranted. In the low stable profiles, traditional training and coaching don't seem to work. In this group, a potentially promising approach may involve individually tailored strategies. For example, some teachers in stable low profiles may have unfavorable opinion about DTT and PRT; they might potentially respond to a strategy strengthening their behavioral and normative beliefs in reference to these treatment methods. Other teachers may experience long-term environmental barriers to DTT and PRT use such as difficulty organizing classroom time; they may respond to a strategy focused on better classroom time organization. Additionally, teachers may be persuaded about each particular child's potential developmental malleability in response to DTT and PRT.

Somewhat puzzling is the finding that teachers' positive beliefs about organization leaders' EBP support and expertise measured by ILS was *inversely* associated with ascending DTT and PRT profiles. One possible interpretation relates to the fact that the ILS asks questions about "evidence-based practice" in general rather than specific practices such as DTT and PRT. Teachers may perceive that principals support for EBP as support for whatever interventions the teacher is using. This may lead to teachers' complacency, resulting in their reduced motivation to increase the use of EBP over time. Literature on organizational behavior suggests that studies focusing on specific behaviors such as particular treatment methods, should measure highly behavior-specific constructs (Judge and Kammeyer-Mueller 2012). Implementation science may benefit from this approach to using measures more specific to the practice of interest. Teachers' assessment of the implementation climate in their organizations was not associated with their trajectory profile. It may be that teachers in ASD support classrooms are isolated from the broader fabric of their schools, potentially leading to the low relevance of organizational implementation climate for teachers' classroom decisions.

Several study limitations should be mentioned. First, our analyses examined EBP use at the child rather than teacher level; to the extent that the choice of EBP is the teacher's decision, this approach may inflate the statistical power because there are fewer teachers than children in the analysis. We used this approach because emerging evidence emphasizes the importance of individual child characteristics for EBP use (Nuske et al. 2019). Our ICC analysis confirmed that some of teachers' treatment decisions may be based on children's individual characteristics, while some other decisions are more teacher-driven. Not accounting for child-level variation in deriving the trajectories of EBP use would effectively wipe out individual variability among the children and would assume that teachers approach to intervention selection is the same for all children in their classrooms. Our ICC analysis, and the analyses of predictors of EBP use would not be possible without accounting for child-level variability of longitudinal EBP patterns. Our analyses predicting child EBP trajectories from teacher and child characteristics statistically accounted for the multilevel data, increasing our confidence in our results, which still should be interpreted with caution.

Another limitation is that the measures of DTT and PRT intensity with each child relied on teachers' reports which were not verified by an independent source, e.g. by observation of intensity. This may potentially introduce social desirability bias and recall bias, among other problems. Unfortunately, in the real classroom context, the lack of process control and teachers' constant attention to the emergent needs of multiple students with ASD makes it unfeasible to reliably quantify the intensity of their DTT and PRT use with each child by any means except self-report. Appropriate steps were taken to assure honest responding: teachers were assured of the confidentiality of their responses and that their responses would not be shared with their employers; and research staff were trained to build trustworthy relationships with the teachers. For example, teachers did not perceive coaches as having an evaluative function or authority; therefore teacher reports were unlikely to be distorted by social desirability bias (Pellecchia et al. 2016). Additionally, coaches' case observations of teachers' work with each individual child served as an informal verification procedure: teachers' intensity reports that did not converge with the coach's case observations were discussed between the coach and the teacher until agreement was reached. To minimize recall bias, teachers reported intensity of their work with each child in a short time span (previous week only). Such reporting tends to tap into the episodic memory of the reported events versus the semantic memory of how the events tend to/should happen. Episodic memory is known to provide more accurate reconstruction of events; and recall bias has been linked to semantic memory more than to episodic memory (Houtveen and Oei 2007; Jaccard and Wan 1995). A third limitation is that DTT and PRT intentions were measured by only one item each. This is not ideal; however studies have demonstrated high predictive validity of these measures in autism support classrooms (Fishman et al. 2018); and prior research has indicated high correspondence of single-item intentions indicators with more complex multiple-item measures of the same constructs (Guilamo-Ramos et al. 2011). Additionally, a new psychometric study (Fishman et al. under review) shows that, at least for some EBPs, a single-item measure of intent to implement has predictive validity on the par with that of a three-item intent scale.

One more potential concern is teachers' skills in PRT and DTT are likely to grow throughout the year, given the coaching program they receive; while our data on teacher skills represent only the baseline measure (or "intercept" value) of this presumably growing characteristic. A similar consideration may also be relevant for some other (but not all) predictor variables used in our MLR analysis, including teachers' intentions to use PRT and DTT. Within the scope of this study, we did not pursue formal analysis of the co-evolution between variables, e.g. growth curve analysis; we rather explored whether the data would reveal meaningfully clustered trajectories of EBP use. Further studies should follow our novel discovery of the above trajectories; and examine how they temporally co-evolve with the development of teachers' skills. An additional limitation is that, for an individual child, a teacher may choose not to use DDT because they are using PRT, or vice versa. This could potentially lead to a lack of independence of the two dependent variables. It is worth mentioning that coaching procedures encouraged the use of DTT and PRT as distinct treatment methods, and the use of both methods for each child instead of picking DTT over PRT for a particular child, or vice versa. Additionally, the majority of children received little of either DTT or PRT throughout the year (the Stable Low Trajectories). There is, therefore, little likelihood for a child to receive DTT at stable low intensity, but receive PRT at a higher intensity, and vice versa. This somewhat reduces the concern about a lack of independence of the two dependent variables. Finally, the modest sample size apparently contributed to the absence of statistically significant comparisons between the High Ascending DTT Trajectory profile and the Late Ascending DTT Trajectory profile, as well as between the Ascending PRT Trajectory profile and the Descending PRT Trajectory profile. Further studies in schools for children with autism may benefit from larger samples of teachers and children to allow more precise comparisons.

Despite these limitations, there are important implications related to these findings. There is considerable variability in children's receipt of evidence-based treatment over the course of the year, with some children receiving increasing doses of treatment and most receiving no or consistently low treatment throughout the year. Child-level and teacher-level factors contribute roughly in equal proportion to determining the child's year-long trajectory of treatment. About half of the potential success in increasing the implementation of EBP for children with autism may be achieved by modifying factors on the level of child-teacher dyads. Another half may be obtained by addressing factors at the teacher/classroom level such as teacher training in ASD interventions, EBP skills, as well as the number of children in the classroom and the number of support staffers. There is a need to identify, early in the academic year, groups of children and teachers likely to be on ascending, or descending, or consistently low trajectories of treatment, because different groups may require tailored implementation strategies.

Funding The project providing data for this study was funded by NIMH (Grant Number R01 MH106175 04). The PI is author, David Mandell, ScD. Victor Lushin, Ph.D., worked on this manuscript while supported by NIH T32 Postdoctoral Fellowship, T32-MH109433-01A1.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Aarons, G. A., Ehrhart, M. G., & Farahnak, L. R. (2014). The implementation leadership scale (ILS): Development of a brief measure of unit level implementation leadership. *Implementation Science, 9*(1), 45.

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research, 38*(1), 4–23.

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., …, Pickles, A. (2007). Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology*, *75*(4), 594.

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(3), 329–341.

Beran, T. N. (2007). Elliott, CD (2007). Differential ability scales. San Antonio, TX: Harcourt Assessment. *Canadian Journal of School Psychology, 22*(1), 128–132.

Castro, S. L. (2002). Data analytic methods for the analysis of multilevel questions: A comparison of intraclass correlation coefficients, $r_{wg(j)}$, hierarchical linear modeling, within-and between-analysis, and random group resampling. *The Leadership Quarterly, 13*(1), 69–93.

Cohen, I. L., & Sudhalter, V. (2005). *PDD behavior inventory (PDDBI)*. Lutz, FL: Psychological Assessment Resources.

Cook, B. G., & Odom, S. L. (2013). Evidence-based practices and implementation science in special education. *Exceptional Children, 79*(2), 135–144.

Ehrhart, M. G., Aarons, G. A., & Farahnak, L. R. (2014). Assessing the organizational context for EBP implementation: The development and validity testing of the implementation climate scale (ICS). *Implementation Science, 9*(1), 157.

Elliott, S. N. (1990). The nature and structure of the DAS: Questioning the test's organizing model and use. *Journal of Psychoeducational Assessment, 8*(3), 406–411.

Fishbein, M., & Ajzen, I. (2011). *Predicting and changing behavior: The reasoned action approach*. Abingdon: Taylor & Francis.

Fishman, J., Beidas, R., Reisinger, E., & Mandell, D. S. (2018). The utility of measuring intentions to use best practices: A longitudinal study among teachers supporting students with autism. *Journal of School Health, 88*(5), 388–395.

Fishman, J., Lushin, V., Pellecchia, M., & Mandell, D. S. (Under review). Intentions to implement evidence based practices in autism services: Measurement issues and predictive validity.

Guilamo-Ramos, V., Jaccard, J., Lushin, V., Martinez, R., Gonzalez, B., & McCarthy, K. (2011). HIV risk behavior among youth in the Dominican Republic: The role of alcohol and other drugs. *Journal of the International Association of Physicians in AIDS Care, 10*(6), 388–395.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

Houtveen, J. H., & Oei, N. Y. (2007). Recall bias in reporting medically unexplained symptoms comes from semantic memory. *Journal of Psychosomatic Research, 62*(3), 277–282.

Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American Journal on Intellectual and Developmental Disabilities, 114*(1), 23–41.

Jaccard, J. (2001). *Interaction effects in logistic regression* (Vol. 135). Thousand Oaks, CA: Sage.

Jaccard, J., & Levitz, N. (2015). Parent-based interventions to reduce adolescent problem behaviors: New directions for self-regulation approaches. in *Self-regulation in adolescence*. New York: Cambridge University Press.

Jaccard, J., & Wan, C. K. (1995). A paradigm for studying the accuracy of self-reports of risk behavior relevant to AIDS: Empirical perspectives on stability, recall bias, and transitory influences. *Journal of Applied Social Psychology, 25*(20), 1831–1858.

Judge, T. A., & Kammeyer-Mueller, J. D. (2012). General and specific measures in organizational behavior research: Considerations, examples, and recommendations for researchers. *Journal of Organizational Behavior, 33*(2), 161–174.

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

Lai, K. (2018). Estimating standardized SEM parameters given nonnormal data and incorrect model: Methods and comparison. *Structural Equation Modeling: A Multidisciplinary Journal, 25*(4), 600–620.

Locke, J., Wolk, C. B., Harker, C., Olsen, A., Shingledecker, T., Barg, F., et al. (2017). Pebbles, rocks, and boulders: The implementation of a school-based social engagement intervention for children with autism. *Autism, 21*(8), 985–994.

Marshall, S., McGoey, K. E., & Moschos, S. (2011). Test review: CD Elliott differential ability scales. San Antonio, TX: Harcourt Assessment, 2007. *Journal of Psychoeducational Assessment, 29*(1), 89–93.

Nonaka, I., & Von Krogh, G. (2009). Perspective—Tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory. *Organization Science, 20*(3), 635–652.

Nuske, H. J., Pellecchia, M., Lushin, V., Rump, K., Seidman, M., Ouellette, R. R., …, Reisinger, E. M. (2019). Do student characteristics affect teachers' decisions to use 1: 1 instruction? *Journal of Autism and Developmental Disorders*, 1–9.

Nylund, K., Bellmore, A., Nishina, A., & Graham, S. (2007). Subtypes, severity, and structural stability of peer victimization: What does latent class analysis say? *Child Development, 78*(6), 1706–1722.

Pellecchia, M., Beidas, R. S., Marcus, S. C., Fishman, J., Kimberly, J. R., Cannuscio, C. C., …, Mandell, D. S. (2016). Study protocol: implementation of a computer-assisted intervention for autism in schools: A hybrid type II cluster randomized effectiveness-implementation trial. *Implementation Science*, *11*(1), 154.

Pellecchia, M., Connell, J. E., Beidas, R. S., Xie, M., Marcus, S. C., & Mandell, D. S. (2015). Dismantling the active ingredients of an intervention for children with autism. *Journal of Autism and Developmental Disorders, 45*(9), 2917–2927.

Polanyi, M. (1966). The logic of tacit inference. *Philosophy, 41*(155), 1–18.

Polanyi, M. (2009). *The tacit dimension*. Chicago: University of Chicago press.

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011). Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM. *Structural Equation Modeling, 18*(2), 161–182.

Schulenberg, J., O'Malley, P. M., Bachman, J. G., Wadsworth, K. N., & Johnston, L. D. (1996). Getting drunk and growing up: trajectories of frequent binge drinking during the transition to young adulthood. *Journal of Studies on Alcohol, 57*(3), 289–304.

Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities, 16*(2), 86–92.

Stahmer, A. C., Rieth, S., Lee, E., Reisinger, E. M., Connell, J. E., & Mandell, D. S. (2015). Training teachers to use evidence-based practices for autism: Examining procedural implementation fidelity. *Psychology in the Schools, 52*(2), 181–195.

Stichter, J. P., Riley-Tillman, T. C., & Jimerson, S. R. (2016). Assessing, understanding, and supporting students with autism at school: Contemporary science, practice, and policy. *School Psychology Quarterly, 31*(4), 443.

Thurm, A., Lord, C., Lee, L. C., & Newschaffer, C. (2007). Predictors of language acquisition in preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*(9), 1721–1734.