



# Discrepancies and Agreement in Perceptions of Implementation Leadership: Associations with Dosage of School-Based Evidence-Based Practices for Children with Autism

Rosemary D. Meza<sup>1,5</sup> · Rinad S. Beidas<sup>2</sup> · Mark G. Ehrhart<sup>3</sup> · David S. Mandell<sup>2</sup> · Shannon Dorsey<sup>1</sup> · Lindsay Frederick<sup>1</sup> · Cristine Oh<sup>4</sup> · Jill Locke<sup>1</sup>

Published online: 5 April 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

Studies demonstrate that discrepancies among leader–follower perceptions of leadership are related to organizational processes that may impact evidence-based practice (EBP) implementation. However, it is unknown whether discrepancies in leadership perceptions also predict EBP use. This study examined the association of principal–staff alignment and discrepancy in perceptions of implementation leadership with the dosage of three Autism-focused school-based EBPs, Discrete Trial Training, Pivotal Response Training (PRT), and Visual Schedules. PRT dosage was higher when principals under-rated their implementation leadership and when it was agreed upon that implementation leadership was low. Findings have implications for leaders support of EBPs implemented in school-based settings.

**Keywords** Implementation leadership · Discrepancy · Evidence-based practice dosage · School mental health · Autism

---

✉ Rosemary D. Meza  
rdmeza@uw.edu

Rinad S. Beidas  
rbeidas@upenn.edu

Mark G. Ehrhart  
Mark.Ehrhart@ucf.edu

David S. Mandell  
MandellD@upenn.edu

Shannon Dorsey  
dorsey2@uw.edu

Lindsay Frederick  
Lkf3@uw.edu

Cristine Oh  
cristine.oh@gmail.com

Jill Locke  
jjlocke@uw.edu

<sup>1</sup> University of Washington, Seattle, USA

<sup>2</sup> University of Pennsylvania, Philadelphia, USA

<sup>3</sup> University of Central Florida, Orlando, USA

<sup>4</sup> University of Pittsburgh, Pittsburgh, USA

<sup>5</sup> Department of Psychology, University of Washington, 335 Guthrie Hall, Box 351525, Seattle, WA 98195, USA

Schools represent the largest public sector for behavioral health services, providing 70–80% of behavioral health care for youth (Burns et al. 1995; Farmer et al. 2003). When behavioral health programs are properly implemented in schools, there are demonstrated improvements in children’s behavioral and emotional functioning (Mihalic et al. 2008; Ozer 2006; Wilson et al. 2003). Nonetheless, these improvements are often stifled by the limited intensity or dosage (e.g., frequency and duration) at which behavioral health services are delivered in schools (Gottfredson and Gottfredson 2002). This treatment gap is exemplified in the limited intensity of school-based services for children with Autism Spectrum Disorder (ASD), a pervasive developmental disorder characterized by deficits in social communication and the presence of restricted or repetitive behaviors (American Psychological Association 2013). Treatment intensity or dosage of ASD evidence-based practices (EBPs) delivered by teachers has been low, regardless of considerable training and support provided (Cook and Odom 2013; Locke et al. 2015; Mandell et al. 2013; Pellecchia et al. 2015; Stahmer et al. 2015; Suhrheinrich et al. 2013).

Characteristics of the context in which EBPs are implemented influence EBP implementation outcomes (Cummings et al. 2007; Henggeler et al. 2008). Despite schools being a primary source for child behavioral health services,

efforts to better measure and characterize the school-based implementation context are just emerging (Lyon et al. 2018; Whitaker et al. 2018). Our understanding of the organizational context, resources, and support available for the implementation of EBPs is lacking in schools in general, but this gap is particularly pronounced in special education classrooms which serve clients with diverse behavioral health needs. In order to better understand the factors that influence ASD EBP delivery, research must situate findings within the unique context of special education settings.

Despite our limited understanding of the special education context, implementation frameworks provide theoretical guidance on individual and organizational factors that can affect providers' use EBPs. The Exploration, Preparation, Implementation, and Sustainment (EPIS) implementation framework (Aarons et al. 2011) identifies a number of characteristics within an organization that is theorized to impact indicators of implementation success (i.e., implementation outcomes; Proctor et al. 2011). Leadership is one factor with growing empirical support for its direct and indirect impact on EBP implementation (Aarons and Sommerfeld 2012; Aarons et al. 2016). Principal support has been identified as a critical factor associated with effective implementation in schools (Forman et al. 2009). Moreover, the degree of principal support for an innovation is positively associated with the quality of teachers' innovation delivery (Gottfredson and Gottfredson 2002).

While there is growing support for the importance of principal leadership in supporting EBP implementation, leader and follower perceptions of leadership often are misaligned (Ostroff et al. 2004). Self-enhancement bias often occurs when individuals rate themselves, which may partially explain the tendency for leaders to rate their own leadership abilities higher than their followers (Atwater and Yammarino 1992; Fleenor et al. 1996). This tendency to overrate oneself is problematic because it may leave leaders unaware of gaps in the support needed by their followers and therefore unable to provide support to facilitate EBP use. While leaders have a tendency to overrate their leadership ability, studies have found a smaller proportion of leaders (i.e., "humble leaders") also underrate their leadership relative to their followers (Aarons et al. 2015, 2017b). Theory suggests that individual and organizational outcomes are optimal when both leader and follower perceptions of leadership are high and that outcomes are less optimal when there is agreement that leadership is poor (Yammarino and Atwater 1993). When leaders perceive their leadership more positively than followers (i.e., over-raters), outcomes are expected to be lowest because leaders lack awareness of their leadership weaknesses, limiting their ability to address them. Outcomes for under-raters (i.e., leaders who rate their leadership lower than followers; humble leaders) are mixed, as these leaders are likely to attempt to improve their leadership

but may be limited in their assessment of their own ability (Yammarino and Atwater 1993).

Discrepancies in perceptions of leadership have been linked to organizational factors that, in turn, predict the quality of implementation; yet their direct association to indicators of implementation success, such as EBP dosage, has not been explored. Specifically, leader and follower agreement on ratings of general and strategic leadership are associated with organizational culture and climate (Aarons et al. 2015, 2017a, b). Although examinations of leader–follower agreement have not been linked to the dosage of EBP use, in business settings, there is evidence that leader–follower alignment directly impacts follower job performance (Atwater et al. 1995). Extending from this work to EBP implementation, it is possible that leader–follower alignment may impact the intensity at which providers deliver EBPs (i.e., EBP dosage), one component of their performance in delivering EBPs.

When examining leadership most likely to impact implementation success, such as EBP dosage, strategic implementation leadership may hold more promise than general leadership constructs (e.g., transformational and transactional leadership) since it measures specific leader behaviors enacted to support EBP implementation (Aarons et al. 2014). Strategic implementation leadership behaviors intended to support EBP dosage likely require targeted behaviors such as being knowledgeable about EBPs and implementation, proactive with regard to implementation, supportive toward implementation, and perseverant in the implementation process (Aarons et al. 2014). Additional research that attends to discrepancies between leader and follower perceptions of strategic implementation leadership behaviors to support EBP implementation and the impact on follower delivery of EBPs is needed.

The current study extends from a larger mixed-methods study aimed at understanding individual and organizational factors that impact EBP implementation in schools (Locke et al. 2016). Our aims are twofold. First, we use qualitative methods to characterize the special education context in which teachers and principals are implementing ASD EBPs. The characteristics and scope of work within schools differs from traditional mental and behavioral health facilities, which can pose novel challenges to EBP implementation (Forman et al. 2009; Langley et al. 2010). Special education classrooms may experience unique challenges in addition to those of traditional classroom settings. Second, we use quantitative methods to assess whether discrepancies exist among principal self-ratings and staff (i.e., teachers and classroom support staff) ratings of principals' implementation leadership behaviors, and whether discrepancies predict EBP dosage. This is the first study to examine alignment of perceptions of leadership between principal and classroom staff dyads. We then examined the association among

principal-staff discrepancies on ratings of implementation leadership behaviors with the frequency of staffs' use of EBPs (i.e., EBP dosage) for students with ASD. We examined the use of three interventions, based on the principles of applied behavior analysis (ABA), to address academic, behavioral, and social outcomes for children with ASD (Arick et al. 2003): discrete trial training (DTT), pivotal response training (PRT), and the use of visual schedules (VS) (see descriptions of each in the Method section).

## Method

### Participants

Table 1 depicts the demographic characteristics of the sample. Participants included 57 principals, 73 Kindergarten-through-third-grade autism support teachers, and 122 classroom support staff from 72 special education classrooms. Participants were from 57 schools in the Northeastern (Philadelphia;  $n = 40$  schools) and Northwestern (Seattle;  $n = 17$  schools) United States. All participants were from schools that were implementing the three EBPs. Most schools had a single special education classroom (73.24%), but 19.72% had two and 7.04% had three special education classrooms. Differences in principal and staff characteristics were examined

for each city. Principals and staff from Philadelphia and Seattle did not differ on any characteristics related to their education or professional development experience. The only statistically significant differences were the ethnicity and race of principals and staff across the cities. Specifically, there were differences in the proportion of principals endorsing being White (Philadelphia  $N = 34.0\%$ ; Seattle  $N = 94.1\%$ ,  $t(53.9) = -5.23$ ,  $p < .01$ ), and Black (Philadelphia  $N = 50.0\%$ ; Seattle  $N = 05.9\%$ ,  $t(54.07) = 4.44$ ,  $p < .01$ ). There were also statistically significant differences in teachers and support staff endorsing being Hispanic (Philadelphia  $N = 4.5\%$ , Seattle  $N = 0\%$ ,  $t(109) = 2.17$ ,  $p < .05$ ), White (Philadelphia  $N = 49.6\%$ ; Seattle  $N = 89.7$ ,  $t(189.01) = -6.90$ ,  $p < .01$ ), and Black (Philadelphia  $N = 42.6\%$ ; Seattle  $N = 2.6\%$ ,  $t(146.15) = 8.06$ ,  $p < .01$ ). Participants were from schools that were implementing three EBPs, DTT, PRT, and VS.

### Intervention

Teachers and classroom staff implemented three ABA-based interventions, DTT, PRT, and VS. The goal of these interventions is to develop children's academic, behavioral, and social skills in a structured setting and generalize those skills in a naturalistic environment (Arick et al. 2003). DTT is a highly structured one-on-one method of teaching new behaviors (e.g., speech, motor movements, requests) and

**Table 1** Participant descriptives

	Principals (n = 57)	Autism support teachers (n = 73)	Classroom support staff (n = 122)
	M (SD)/n (%)	M (SD)/n (%)	M (SD)/n (%)
Demographic characteristics			
Age (years)	47.62 (7.59)	37.16 (10.41)	42.46 (12.71)
Gender (female)	36 (63%)	69 (97%)	110 (92%)
Ethnicity (Hispanic/Latino)	4 (7%)	0 (0%)	7 (4%)
Race (not mutually exclusive)			
White	33 (58%)	63 (86%)	66 (54%)
African American	21 (37%)	8 (11%)	44 (36%)
American Indian/Alaska Native	1 (2%)	1 (1%)	1 (1%)
Asian/Pacific Islander	2 (4%)	2 (3%)	9 (7%)
Multiracial and other	1 (2%)	0 (0%)	1 (1%)
Education			
High school	0 (0%)	0 (0%)	22 (18%)
Some college and vocational	0 (0%)	0 (0%)	40 (33%)
College degree	3 (5%)	10 (14%)	42 (35%)
Graduate/professional	52 (91%)	62 (85%)	15 (13%)
Other	2 (4%)	1 (1%)	1 (1%)
Professional experience			
Tenure in current position (years)	5.29 (4.32)	3.60 (4.66)	4.56 (5.42)
Tenure in position as principal (years)	9.67 (7.53)	–	–
Specialized ASD training	10 (18%)	57 (79%)	56 (48%)
Special education certification	–	72 (99%)	20 (16%)

discriminations (Smith et al. 2000). DTT is used to break skills into subunits and shape new behaviors. During interactions, staff initiate a stimulus to elicit a response from the child and reinforcement is used to reward appropriate responses. Trials often involve repeated practice of a successful response during several teaching episodes. PRT, which can occur one-on-one or in a group format, is a naturalistic behavioral intervention focused on improving language and social skills (Koegel et al. 1988). PRT is child-initiated by indicating interest in an item or activity and, upon completing a particular behavior (e.g., labeling the item, imitating the teacher), the student is reinforced by receiving the item. PRT often utilizes novel toys or items to engage the student and increase their motivation. Visual Schedules involve visual supports to transition students from activity to activity (Dettmer et al. 2000). They provide the learner with prompts to independently move through a series of activities and have been used to support daily living, academic, vocational and leisure skills. These visual prompts come in a variety of formats, including pictures or written words. The use of visual schedules involves providing prompts and cues to teach the child to independently engage in school and self-care routines.

## Training

Coaches provided all teachers with 2 days of intensive didactic training covering the EBPs (DTT, PRT, and VS) during district professional development days (Locke et al. 2016; Pellecchia et al. 2015). Coaches worked with teachers in their classrooms to set up the classrooms and plan for the implementation of the EBPs with students. Following training, all teachers were provided ongoing coaching on the EBPs, approximately once per month for 2 h each visit for the duration of the school year. Coaches were Master's and PhD-level clinicians with expertise in the EBPs. Coaches communicated the content of consultation to all principals during an initial meeting; however, coaches did not regularly meet with principals.

It was recommended that each student receive at least two sessions of DTT and at least one session of PRT daily. Visual Schedules are used more fluidly throughout the day, and it was recommended that each student receive targeted instruction on one-to-two visual schedules daily.

## Procedure

The research team met with the school district officials in both cities to obtain a list of Kindergarten through third grade autism support classrooms and schools. Meetings were arranged with the principal at each prospective school to discuss the research activities and obtain a letter of agreement to conduct research on their campus. Subsequently,

recruitment materials (e.g., informational handouts, flyers) were distributed to the schools, and the research team met with interested participants to provide additional information about the study and their role as a study participant, to enable them to make an informed decision regarding their participation. Upon receiving informed consent, principals, teachers, and classroom support staff were asked to complete study measures. A subset of participants ( $n=26$  teachers) from high and low performing classrooms (based on their dosage data) also participated in a semi-structured interview to understand the implementation process of each of the EBPs and experiences of the teachers. Participants were compensated \$50 for their survey completion time and another \$50 for their semi-structured interview. Procedures were reviewed and approved by the [University of Washington and University of Pennsylvania] Institutional Review Boards and each participating school district.

## Measures

### Demographic Questionnaire

Participants completed a short demographic questionnaire that asked about their age, sex, racial/ethnic background, educational attainment (highest degree), teaching experience (in years), autism experience (yes/no), and special education certification (teachers and classroom support staff only).

### Leadership

Principals' implementation behaviors were rated using the Implementation Leadership Scale (ILS), a 12-item measure rated on a scale of 0 (not at all) to 4 (to a very great extent) (Aarons et al. 2014). The ILS has four subscales that assess the degree to which a leader is knowledgeable (deep understanding of EBP and implementation issues), supportive (support for EBP adoption/use), proactive (anticipating and addressing implementation challenges), and perseverant (consistent and responsive to challenges) in implementing EBPs. The ILS is psychometrically validated and has demonstrated strong internal consistency previously and in the current sample ( $\alpha=.97$ ). The ILS was adapted in collaboration with the developers for use in schools, and the factor structure and psychometric properties were upheld (Lyon et al. 2018). Principals completed self-ratings of their implementation leadership and teachers and classroom support staff completed ratings of their principal's implementation leadership.

### Dosage

Dosage was measured using a teacher-report of how frequently the three intervention strategies were used in the

classroom (Locke et al. 2016; Pellecchia et al. 2015). Since teachers and classroom support staff within a single classroom were both trained in and implemented the interventions together, dosage was measured at the classroom-level. Teachers reported on the intensity with which each of the three interventions was delivered for every student with ASD in their classroom during the past week. The intensity of DTT and PRT delivery were rated as follows: “0” (less than one time per week), “1” (one time per week), “2” (two to four times per week), “3” (one time per day), and “4” (two times per day). Since Visual Schedules are used during transitions, teachers used the following scale to rate intensity: “0” (never), “1” (few transitions), “2” (some transitions), “3” (most transitions), “4” (every transition). Dosage was measured on two occasions during the middle of the school year to allow time for teachers to learn the interventions. Dosage measurements occurred approximately 2 months apart and scores were averaged across the two time points.

### Qualitative Measure

The parent study included a systematic and comprehensive interview guide with questions that elicited open-ended responses that asked about teachers’ experiences implementing each of the EBPs in their classroom and their school context. We assessed broad inner context themes related to the school and special education environment, leadership and other supports for implementing ASD EBPs, and classroom staffs’ experience implementing the ASD EBPs. We include major themes from this analysis to provide context for the implementation of these EBPs.

### Data Analysis

#### Qualitative Data Analysis

This study used a mixed-methods approach. The purpose of the qualitative analyses was to characterize the context in which the EBPs were implemented as special education classrooms represent a unique context for EBP delivery. First, quantitative data in the parent study were collected and qualitative interviews were subsequently conducted with a subset of participants to supplement the quantitative findings (“QUAN → qual”; Palinkas et al. 2011). Second, the function of the qualitative analysis was largely complementarity. The qualitative data were used to provide depth of understanding of the context of the quantitative findings. Data were analyzed in this way to better understand the classroom staff’s use of the EBPs in the classroom.

Semi-structured interviews were transcribed and uploaded to NVivo QSR 10 for data management. The coding scheme was developed using an iterative approach. Three raters independently coded four initial transcripts line-by-line to

identify recurring codes. They met as a group to discuss recurring codes to develop a codebook. The codebook was developed using an integrated and iterative approach to coding as certain codes were conceptualized during the interview guide development. This deductive approach centered on codes related to the school and classroom environment, experience implementing the EBPs, and barriers and facilitators to implementation. Other codes were developed through a close reading of the transcripts (i.e., inductive approach; Bradley et al. 2007). The codebook contained operational definitions of each code, examples of the code from the data, and when to use or not use the code. The coding scheme was refined in multiple stages and then applied to the data (Bradley et al. 2007). Five coders coded all data and overlapped on 20% of randomly selected transcripts to determine inter-rater reliability. They met together on a regular basis to discuss, clarify, and compare emerging codes to ensure consensus. Agreement was calculated based on the number of words agreed upon; agreement between raters was excellent (percent agreement = 94.95% on all parent codes).

#### Quantitative Data Analysis

The ILS ratings by teacher and classroom support staff were aggregated to create a school-level rating of implementation leadership for each principal. On average, 3.34 staff ( $SD = 1.56$ , range 1–8) reported on their principal’s implementation leadership. The average within-group agreement ( $a_{wg(j)}$ ) was .63, indicating moderate agreement (Lebreton and Senter 2008) and the intraclass correlation coefficient (ICC(1)) was .23, indicating a strong classroom-level effect for leadership. Thus, although the within-group agreement was only moderate, the strong ICC(1) value provided adequate support for the aggregation of the ILS scores (Bliese 2000; James 1982; Lebreton and Senter 2008).

Prior to conducting polynomial regression and response surface analyses, staff and principal implementation leadership ratings were inspected to ensure adequate frequency of discrepancies between their ratings (Shanock et al. 2010). Implementation leadership ratings were standardized and principal–staff scores that differed by at least half a standard deviation were considered discrepant (Fleenor et al. 1996). Three polynomial regression models were estimated to examine the relation between discrepancies in implementation leadership ratings and classroom-level dosage to each of the EBPs, after controlling for teachers’ age, years teaching special education, and whether or not teachers have specialized training in ASD (Shanock et al. 2010, 2014). Predictor variables were centered on their midpoint (see Atwater et al. 2005). Regression model assumptions were assessed to test the accuracy of the models. Next, three variables were computed: (a) the square of the centered variable for principal implementation leadership ratings; (b) the cross-product of

the principal and staff implementation leadership ratings; and (c) the squared term of the centered variable for staff implementation leadership ratings. The proportion of the outcome variance explained ( $R^2$ ) was examined to determine whether it was significantly greater than zero. For models with a significant  $R^2$ , coefficients (which are not directly interpretable when conducting polynomial regression models) were transformed into four surface test values that are interpretable:  $a_1$ – $a_4$  (Shanock et al. 2010).

To facilitate interpretation of the results, the three-dimensional response surface was plotted and its features were examined (Shanock et al. 2010). The line of congruence is the axis along which leader and follower ratings of implementation leadership are congruent. The line of incongruence is the axis along which leader and follower ratings of implementation leadership are incongruent. For the line of congruence ( $x = y$ ), we evaluated the linear ( $a_1$ ) and non-linear ( $a_2$ ) relation between implementation leadership agreement and EBP dosage. Along the line of incongruence ( $x = -y$ ), we examined the relation between EBP dosage with the direction of discrepancy ( $a_3$ ) and the degree of discrepancy ( $a_4$ ). All data analysis was performed in R version 3.3.1.

## Results

### Qualitative Analyses

Several themes emerged from the principal and teacher interviews. Below, we discuss four important themes (working in under-resourced contexts; working in isolation; teacher motivation; reliance on external support) that characterize the context in which the EBPs were being delivered. The themes highlight the nuances of the implementation context that may impact the teachers' use of the EBPs and principals support of their EBP use.

### Working in Under-Resourced Contexts

Participants described the lack of resources in many schools and the tumultuous conditions in which they work that often impeded EBP implementation. Teachers often reported purchasing their own materials in order to deliver the EBPs. One teacher described that “most of the resources [she] has were purchased by [her]” or donated by family members. Another teacher remarked that she “[goes] to Goodwill a lot” and has applied for external funding to purchase supplies such as “Donors’ Choose projects to get different things in the classroom.” Other teachers “make copies of things [since they often] do not have booklets for each of the kids” just to “get by.”

### Working in Isolation

Compounding the under-resourced settings of many schools is the frequent isolation in which many special education teachers work. Teachers often described the isolating conditions that special education teachers face as they often are removed from general education altogether. One teacher commented that she is “in a bubble [and] there is so much going on outside of this bubble that [they] sort of get lost and left alone until there’s a problem.” Other teachers described physical isolation where their classrooms are situated far away from general education classrooms. One teacher said that they are “on the fourth floor” and “the only classroom up there.” Other teachers were situated in the basement as captured by this teacher’s statement, “Even if there’s another autism support [special education] teacher, I really think I’m alone because I’m in the basement. Half of the time the school doesn’t even know I’m a teacher. I covered a sixth grade classroom 1 day, and the teachers up there thought I was a substitute teacher. That’s how I just feel, I’m by myself, and I just do everything by myself.” Aside from classroom separation, teachers also noted their children with ASD are excluded from communal spaces. One teacher noted that her students “do not get to go out to recess with the general education kids.”

### Teacher Motivation

Motivation to help their students appears to be endemic to the teaching profession. Many teachers described their strong desire to support their students with ASD, even in the absence of resources or other supports in their schools. One teacher expressed her strong belief in her students with ASD and her calling as a teacher, stating that she “will work so hard to do what’s best for [students].” She continued on to say that she will do whatever it takes to make research-based practices available because “having research-based practices in the classroom is what they deserve.” Many teachers also noted that they would “do what is best” for their students with ASD because of their love for what they do and their students.

### Reliance on External Support

ASD is a complex disorder and schools often do not have the breadth and depth of knowledge that may be necessary to support students with ASD. Teachers frequently noted their collaboration and reliance on external supports, in particular, their coach. One teacher commented it is so helpful “to get suggestions from [the coach],” which helps with “feelings of isolation.” Another teacher noted that she often “does not have much interaction with other classrooms, so [she] does not know what other classrooms

are doing,” so the coaches’ support is invaluable since the coach “observes other classrooms to see what they are doing and how they are implementing things.”

### Quantitative Analyses

Means, standard deviations, and correlations for all variables included in the polynomial regression analysis are shown in Table 2.

Overall, 34.18% of staff and principals were in agreement (Principal ILS Mean = 2.85; Staff ILS Mean = 2.50), 32.91% of staff rated their principal higher than their principal rated themselves (Principal ILS Mean = 2.59; Staff ILS Mean = 3.18), and 32.91% of staff rated their principals lower than their principals rated themselves (Principal ILS Mean = 3.31; Staff ILS Mean = 2.03). The distribution of discrepancies in both directions was sufficient to proceed with polynomial regression analyses. Table 3 includes the results of the polynomial regression models and the response surface analyses. Teachers’ age, years teaching special education, and whether or not teachers have specialized training in ASD were included in the models as control covariates; however, their effects were not statistically significant, and did not change the interpretation of the coefficients. They were therefore excluded from the final model for parsimony. Results of the omnibus tests suggest that staff and principal ratings of implementation leadership account for a statistically significant amount of variance in dosage of PRT ( $R^2 = 0.135, p < .01$ ), but did *not* account for a significant amount of variance in dosage of DTT ( $R^2 = 0.126, p = .10$ ) or VS ( $R^2 = 0.056, p = .55$ ). Since DTT and VS did not meet the requirement for a significant proportion of variance to be explained by the model, response surface modeling was only conducted for PRT. Inspection of model assumptions revealed four outliers (i.e., standardized residuals  $\geq 3$ ); however, exclusion of these observations did not impact the significance of model estimates or their interpretation and calculation of Cook’s Distance indicated the observations did not have undue influence on the model (Cook and Weisberg 1982). Therefore, all observations were retained in the model.

### Pivotal Response Training

The line of congruence had a significant negative slope ( $a_1 = -1.04, SE = 0.36, p < .01$ ) and non-significant curvature ( $a_2 = 0.37, SE = 0.22, p = .10$ ), indicating a negative linear relation between staff and principal-rated implementation leadership and PRT dosage. Findings suggest that as agreed-upon implementation leadership scores increased, classroom level PRT dosage decreased. The direction of discrepancy in staff and principal implementation leadership ratings was significantly associated with PRT dosage ( $a_3 = 0.99, SE = 0.44, p < .05$ ), but the curvature of the line of incongruence was not significant ( $a_4 = 0.53, SE = 0.36, p = .15$ ). The significant positive direction coefficient indicated that when staff rated their principal higher on implementation leadership and principals rated themselves lower on implementation leadership (i.e., under-raters) PRT dosage was higher as compared to when staff rated principals lower and principals rated themselves higher on implementation leadership (i.e., over-raters).

### Discussion

The quantitative findings suggest that discrepancies in principal and staff perceptions of principals’ implementation leadership behaviors exist, with a third of staff rating their principal’s implementation leadership higher and a third rating their implementation leadership lower relative to principals’ ratings of themselves. While these discrepancies exist, they only accounted for a significant proportion of variance in one outcome: PRT dosage, which increased as agreed-upon ratings of principals’ implementation leadership decreased. The direction of the discrepancy in ratings of implementation leadership also was associated with PRT dosage. Results showed that PRT dosage was higher when staff rated principals higher in implementation leadership relative to principal’s ratings of themselves as compared to when the staff rated principals lower in implementation leadership relative the principal’s self-ratings. The qualitative findings suggest that special education teachers are often operating in under-resourced settings and express feelings of isolation in their schools that may affect their

**Table 2** Implementation leadership and dosage descriptive statistics

	Mean	SD	1	2	3	4	5
1. ILS—principal	2.91	0.59	–				
2. ILS—staff	2.60	0.77	0.26*	–			
3. Dosage—DTT	1.06	1.15	–0.27*	–0.21	–		
4. Dosage—VS	1.85	1.84	–0.15	–0.09	0.11	–	
5. Dosage—PRT	0.59	0.85	–0.24*	–0.19	0.28**	0.19	–

\*  $p < .05$ ; \*\*  $p < .01$

**Table 3** Polynomial regression and response surface analyses

Discrete trial training dosage		<i>Coeff</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Polynomial regression	Intercept	1.87	0.30	6.17	<0.01
	Implementation leadership: staff	0.07	0.33	0.21	0.83
	Implementation leadership: principal	−0.68	0.53	−1.28	0.20
	Implementation leadership: staff squared	−0.29	0.24	−1.12	0.23
	Implementation leadership: staff × principal	−0.02	0.40	−0.04	0.97
	Implementation leadership: principal squared	0.10	0.30	0.33	0.74
	R <sup>2</sup>	0.13	–	–	0.10
Visual schedules dosage		<i>Coeff</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Polynomial regression	Intercept	2.08	0.40	5.26	< 0.01
	Implementation leadership: staff	−0.23	0.43	−0.55	0.58
	Implementation leadership: principal	−0.94	0.69	−1.36	0.18
	Implementation leadership: staff squared	0.18	0.31	0.59	0.56
	Implementation leadership: staff × principal	0.18	0.53	0.34	0.74
	Implementation leadership: principal squared	0.30	0.39	0.75	0.46
	R <sup>2</sup>	0.06	–	–	0.55
Pivotal response training dosage		<i>Coeff</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Polynomial regression	Intercept	1.03	0.20	5.26	<0.01
	Implementation leadership: staff	−0.02	0.21	−0.10	0.92
	Implementation leadership: principal	−1.02	0.34	−2.96	< 0.01
	Implementation leadership: staff squared	−0.01	0.16	−0.09	0.93
	Implementation leadership: staff × principal	−0.08	0.26	−0.32	0.75
	Implementation leadership: principal squared	0.46	0.20	2.36	0.02
	R <sup>2</sup>	0.15	–	–	0.04
Response surface analyses	<i>a</i> <sub>1</sub>	−1.04	0.36	−2.91	<0.01
	<i>a</i> <sub>2</sub>	0.37	0.22	1.66	0.10
	<i>a</i> <sub>3</sub>	0.99	0.44	2.24	0.03
	<i>a</i> <sub>4</sub>	0.53	0.36	1.47	0.15

implementation of autism EBPs. Despite these challenges, teachers reported strong motivation to serve their students with ASD, including through the use of EBPs for ASD. In order to support their use of the EBPs, teachers noted utilizing their coaches for guidance to mitigate the effects of their isolating context.

The increase in PRT dosage when principals underestimate their implementation leadership relative to staff is consistent with previous research finding positive outcomes when leaders underestimate their own leadership (Aarons et al. 2017b; Atwater et al. 1995). Aarons et al. (2017a, b) found that “humble leaders,” those who under-rate their implementation leadership relative to their staff, reported greater involvement and performance feedback climates. While their findings suggest that underestimating one’s leadership can impact outcomes proximally related to

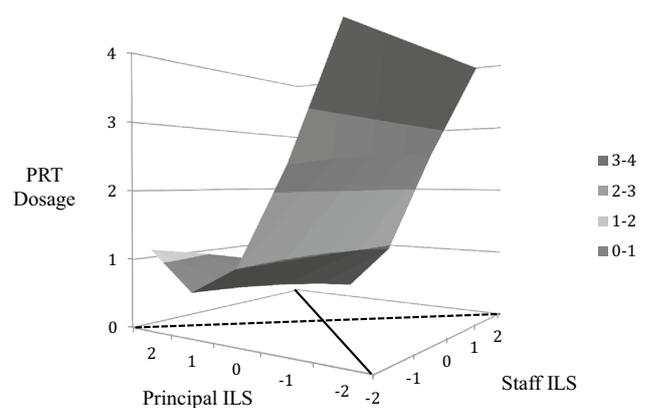
implementation, this is the first study to demonstrate that underestimating one’s implementation leadership relative to staff ratings is associated with staff-reported dosage of EBPs.

The tendency for PRT dosage to be higher when leaders underrate their implementation leadership may be indicative of a general process through which leaders impact the work environment. Empirical research on humble leadership, where leaders recognize their limitations and others’ strengths, suggests that it operates through social contagion (Owens and Hekman 2012, 2016). That is, humble leadership fosters processes of a team’s collective humility (e.g., being teachable, owning limitations and mistakes) and this humility, on behalf of the leader and team, impacts the team’s performance (Owens and Hekman 2016). The impact of humble leadership on team functioning may be

particularly relevant in classroom-based settings where these EBPs were implemented as classroom staff and teachers are working together as a team to implement the EBPs. Alternatively, the tendency for leaders to underrate themselves could be reflective of team psychological safety, the “shared belief that the team is safe for interpersonal risk taking” (Edmondson 1999). Prior research has demonstrated that team psychological safety is positively associated with team performance, and this association is mediated through engaging in learning behaviors (Edmondson 1999). The negative effects of over-ratings one’s own leadership may be amenable to current leadership interventions. The leadership and organizational change for implementation (LOCI) intervention involves a 360° assessment where leaders and their staff rate their leadership (Aarons et al. 2017a). These ratings are used to provide feedback to leaders and develop tailored plans to improve their leadership behaviors to support EBP implementation. The LOCI intervention is currently being applied in school settings to support the implementation of ASD interventions (Brookman-Frazer and Stahmer 2018).

Although the results were consistent that “under-rating” leaders had more positive outcomes than “over-rating” leaders, those results must be interpreted in light of the general pattern of results that PRT dosage was highest when principal’s self-rated implementation leadership was low across all levels of staff ratings of the principals’ implementation leadership (see Fig. 1). In fact, the regression results showed a significant and negative main effect for principal ILS ratings but a non-significant relationship for staff ILS ratings, which is reflected in Fig. 1 as well as the findings for the line of congruence. This negative relationship between implementation leadership and PRT dosage was unexpected in light of theory (Bass and Yammarino 1991) and previous research demonstrating a positive association between agreement that leadership is strong with both organizational climate (Aarons et al. 2015, 2017a, b) and job performance (Atwater et al. 2005).

There are a number of possibilities for these unexpected results. One possibility is that when staff and principals are both aware that principals are not a source of implementation support, this may engender greater intrinsic motivation among staff in implementing the EBPs and potentially improve their treatment use. Empirical studies demonstrate that autonomy is required for intrinsic motivation (Niemic and Ryan 2009) and that intrinsically motivated, versus externally controlled, individuals show enhanced performance (Ryan and Deci 2000). Our qualitative data suggest that the special education teachers tended to be intrinsically motivated to use the practices and use them well, despite limited resources and feeling isolated in their schools. Another explanation is that when both staff and principals recognize that principals’ implementation leadership is lacking, staff may be more likely to seek out the support of the



**Fig. 1** The figure depicts the Implementation Leadership Scale agreement, degree, and direction of disagreement associated with Pivotal Response Training dosage. The solid line depicts the line of congruence, representing how agreement in ILS ratings is related to PRT dosage. This line has a significant negative slope ( $a_1$ ), suggesting that PRT increases as agreed-upon ILS scores decrease. The curvature of the line of congruence ( $a_2$ ) was non-significant, suggesting a linear relationship between ILS agreement and PRT dosage. The dashed line depicts the line of incongruence, representing how discrepancy in ILS ratings relates to PRT dosage. The direction of the discrepancy ( $a_3$ ) matters, such that PRT dosage is higher when staff ILS is high, and principal ILS is low (right corner of graph). The curvature of the line of incongruence ( $a_4$ ) was non-significant, suggesting the degree of discrepancy did not relate to PRT dosage

consultants who provided training and ongoing coaching in the EBPs, which may enhance their use of PRT. The qualitative results suggest that consultants functioned as a form of support for teachers and means to compare the quality of their EBP use to other teachers implementing the EBPs. Given consultants’ expertise in the EBPs, higher reliance on implementation support from consultants, rather than principals, may better support teachers in their use of PRT. A final explanation could be that the causal direction may be reversed, such that the success of the intervention is used as a guide for principals to know how much implementation leadership is needed. If fidelity is high, then principals may not think their emphasis on implementation is necessary, and thus they focus on other strategic goals at their schools. While these explanations may be possible, we cannot make definitive claims about any of these processes given the current data. Further exploration of principals’ role in supporting EBP implementation is needed to clarify this association.

It is important to highlight that the self-reported dosage of all three EBPs was quite low. Teachers’ average reported use of the EBPs most closely maps onto usage of DTT and PRT “once per week” and during “some transitions” for Visual Schedules. These rates are below the recommended frequency of delivery for these EBPs (National Research Council 2001). While having a principal who underrates their implementation leadership predicted the higher use

of PRT relative to principals who overrate their leadership, only two classrooms reported average use of PRT that corresponds to the recommended daily use. This suggests that there are likely additional factors, such as the limited resources in schools, that may be interfering with classroom staff's use of the EBPs.

The influence of agreement and discrepancies in implementation leadership on only PRT dosage may be due to the increased complexity and resources required to implement PRT relative to DTT and VS. Prior research has suggested that organizational level constructs, in this case leadership, may have limited influence on the implementation of innovations that can be implemented by an individual with relatively minimal training and support (Weiner et al. 2011). While the three treatment models implemented are all based on principles of applied behavioral analysis, they differ in their level of structure, complexity, and training required for mastery (Stahmer et al. 2015). DTT and VS are both highly structured, and the steps for implementation are clear and simple to follow. PRT is the least structured of the three strategies and requires additional materials that are not typically provided by schools. For instance, PRT requires staff to have novel toys that students do not regularly have access to in order to enhance motivation and interest. Visual Schedules also require some materials, but these are more readily available in schools (e.g., paper, laminators, printers) and the resources used can be modified based on available materials (e.g., written or drawn schedules can be used instead of printed images). The relative difficulty of PRT and necessity for additional resources to implement may enhance the importance of the degree and direction of implementation leadership in supporting the use of PRT, especially when operating in already under-resourced environments. The relative ease and low resource burden of VS and DT may mean they can be sufficiently supported by consultants, precluding the need for principal implementation support. Additionally, dosage for DT and VS may be better explained by individual-level factors, such as teachers' intrinsic motivation which we speculate may support PRT dosage in the absence of supportive principal leadership.

### Limitations

Several limitations should be noted. First, dosage was measured via teachers' self-report, potentially introducing self-report bias. The necessity for teachers to recall the intervention dosage for each student over the past week may have led to over- or underestimates of the actual dosage provided. While this method of measuring dosage ensures dosage for all students is accounted for and also is more feasible than requiring teachers to track more frequently, it does likely introduce error into our dosage measurement. Second, teachers and classroom support

staff shared the role of delivering the EBPs and therefore dosage was measured at the classroom level, rather than the individual level. Due to the limited number of schools with multiple special education classrooms and very small number of classrooms within schools with more than one special education classroom (i.e., 2–3), we were unable to account for the nesting of classes in schools or disentangle the effects of the school from those of the classroom. Third, the agreement levels for the ILS were not as high as we would have liked. Although the ICC(1) levels were strong and supported studying leadership at the classroom level, future research should investigate the issues that may cause variability in how principals' implementation leadership is perceived. Given that the primary focus of the parent-study was to assess implementation outcomes, we were not able to control for possible student-level covariates that may have influenced intervention dosage such as ASD severity. Finally, the data are cross-sectional and correlational; therefore, we are unable to make conclusions about causality.

### Conclusion

Alignment and discrepancies between staff and principal's assessment of principal's implementation leadership was associated with the dosage of PRT. This further supports the view that leadership is important in the use of EBPs and also suggests that alignment in perceptions of implementation leadership may be more influential in the use EBPs with increased complexity and resource requirements. The findings support previous research demonstrating the utility of "humble leadership" and also raise additional questions about how the use of EBPs is supported when there is consensus that implementation leadership is weak. These findings underscore the need for additional research on EBP characteristics that necessitate strong implementation leadership to support their use and how EBP use is supported in the face of limited leadership support.

**Funding** This study was funded by the National Institute of Mental Health (K01 MH100199).

### Compliance with Ethical Standards

**Conflicts of interest** The authors have no conflicts of interest to declare.

**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.

## 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), 1–10.
- Aarons, G. A., Ehrhart, M. G., Farahnak, L. R., Sklar, M., & Horowitz, J. (2015). Discrepancies in leader and follower ratings of transformational leadership: Relationship with organizational culture in mental health. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*(4), 480–491. <https://doi.org/10.1007/s10488-015-0672-7>.
- Aarons, G. A., Ehrhart, M. G., Moullin, J. C., Torres, E. M., & Green, A. E. (2017a). Testing the leadership and organizational change for implementation (LOCI) intervention in substance abuse treatment: A cluster randomized trial study protocol. *Implementation Science*, *12*(1), 1–11. <https://doi.org/10.1186/s13012-017-0562-3>.
- Aarons, G. A., Ehrhart, M. G., Torres, E. M., Finn, N. K., & Beidas, R. S. (2017b). The humble leader: Association of discrepancies in leader and follower ratings of implementation leadership with organizational climate in mental health. *Psychiatric Services*, *68*(2), 115–122. <https://doi.org/10.1176/appi.ps.201600062>.
- Aarons, G. A., Green, A. E., Trott, E., Willging, C. E., Torres, E. M., Ehrhart, M. G., et al. (2016). The roles of system and organizational leadership in system-wide evidence-based intervention sustainment: A mixed-method study. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(6), 991–1008. <https://doi.org/10.1007/s10488-016-0751-4>.
- 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*, *28*(1), 4–23. <https://doi.org/10.1007/s10488-010-0327-7>.
- Aarons, G., & Sommerfeld, D. (2012). Leadership, innovation climate, and attitudes toward evidence-based practice during a statewide implementation. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(4), 423–431. <https://doi.org/10.1016/j.jaac.2012.01.018>.
- American Psychological Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- Arick, J. R., Young, H. E., Falco, R. A., Loos, L. M., Krug, D. A., Gense, M. H., et al. (2003). Designing an outcome study to monitor the progress of students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *18*(2), 75–87. <https://doi.org/10.1177/108835760301800201>.
- Atwater, L., Roush, P., & Fischthal, A. (1995). The influence of upward feedback on self- and follower ratings of leadership. *Personnel Psychology*, *48*(1), 35–59. <https://doi.org/10.1111/j.1744-6570.1995.tb01745.x>.
- Atwater, L., Waldman, D., Ostroff, C., Robie, C., Waldman, D., & Johnson, K. M. (2005). Self-other agreement: Comparing its relationship with performance in the U.S. and Europe. *International Journal of Selection and Assessment*, *13*(1), 25–40. <https://doi.org/10.1111/j.0965-075X.2005.00297.x>.
- Atwater, L. E., & Yammarino, F. J. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology*, *45*(1), 141–164. <https://doi.org/10.1111/j.1744-6570.1992.tb00848.x>.
- Bass, B. M., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology*, *40*(4), 437–454.
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability. In K. J. Klein & S. W. Kozlowski (Eds.), *Multilevel theory research, and methods in organizations* (pp. 349–381). San Francisco: Jossey-Bass.
- Bradley, E. H., Curry, L. A., & Devers, K. J. (2007). Qualitative data analysis for health services research: Developing taxonomy, themes, and theory. *Health Services Research*, *42*(4), 1758–1772. <https://doi.org/10.1111/j.1475-6773.2006.00684.x>.
- Brookman-Frazee, L., & Stahmer, A. C. (2018). Effectiveness of a multi-level implementation strategy for ASD interventions: Study protocol for two linked cluster randomized trials. *Implementation Science*, *13*(66), 1–14. <https://doi.org/10.1186/s13012-018-0757-2>.
- Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M. Z., et al. (1995). Children's mental health service use across service sectors. *Health Affairs*, *14*(3), 147–159.
- Cook, B. G., & Odom, S. L. (2013). Evidence-based practices and implementation science in special education. *Exceptional Children*, *79*(2), 135–144. <https://doi.org/10.1177/001440291307900201>.
- Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman and Hall.
- Cummings, G. G., Estabrooks, C. A., Midodzi, W. K., Wallin, L., & Hayduk, L. (2007). Influence of organizational characteristics and context on research utilization. *Entrepreneurship and Regional Development*, *55*(4), S24–S39. <https://doi.org/10.1080/08985626.2017.1328905>.
- Dettmer, S., Simpson, R. L., Smith Myles, B., & Ganz, J. G. (2000). The use of visual supports to facilitate transition of students with autism. *Focus on Autism and Other Developmental Disabilities*, *15*(3), 163–169.
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350. <https://doi.org/10.2307/2666999>.
- Farmer, E. M. Z., Burns, B. J., Phillips, S. D., Angold, A., & Costello, E. J. (2003). Pathways into and through mental health services for children and adolescents. *Psychiatric Services*, *54*(1), 60–66. <https://doi.org/10.1176/appi.ps.54.1.60>.
- Fleener, J. W., McCauley, C. D., & Brutus, S. (1996). Self-other rating agreement and leader effectiveness. *Leadership Quarterly*, *7*(4), 487–506. [https://doi.org/10.1016/S1048-9843\(96\)90003-X](https://doi.org/10.1016/S1048-9843(96)90003-X).
- Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Mental Health*, *1*(1), 26–36. <https://doi.org/10.1007/s12310-008-9002-5>.
- Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency*, *39*(1), 3–35.
- Henggeler, S. W., Chapman, J. E., Rowland, M. D., Halliday-Boykins, C. A., Randall, J., Shackelford, J., et al. (2008). Statewide adoption and initial implementation of contingency management for substance-abusing adolescents. *Journal of Consulting and Clinical Psychology*, *76*(4), 556–567. <https://doi.org/10.1037/0022-006X.76.4.556>.
- James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology*, *67*(2), 219–229.
- Koegel, R. L., Schreibman, L., Good, A., Cerniglia, L., Murphy, C., & Koegel, L. K. (1988). *How to teach pivotal behaviors to children with autism: A training manual*. Santa Barbara: University of California.
- Langley, A. K., Nadeem, E., Kataoka, S. H., Stein, B. D., & Jaycox, L. H. (2010). Evidence-based mental health programs in schools: Barriers and facilitators of successful implementation. *School Mental Health*, *2*(3), 105–113. <https://doi.org/10.1007/s12310-010-9038-1>.
- Lebreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815–852.

- Locke, J., Beidas, R. S., Marcus, S., Stahmer, A., Aarons, G. A., Lyon, A. R., et al. (2016). A mixed methods study of individual and organizational factors that affect implementation of interventions for children with autism in public schools. *Implementation Science*, 11(1), 1–9. <https://doi.org/10.1186/s13012-016-0501-8>.
- Locke, J., Olsen, A., Wideman, R., Downey, M. M., Kretzmann, M., Kasari, C., 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(1), 54–67. <https://doi.org/10.1016/j.beth.2014.05.001>.
- Locke, J., Violante, S., Pullmann, M. D., Kerns, S. E. U., Jungbluth, N., & Dorsey, S. (2017). Agreement and discrepancy between supervisor and clinician alliance: Associations with clinicians' perceptions of psychological climate and emotional exhaustion. *Administration and Policy in Mental Health and Mental Health Services Research*, 0(0), 1–13. <https://doi.org/10.1007/s10488-017-0841-y>.
- Lyon, A. R., Cook, C. R., Brown, E. C., Locke, J., Davis, C., Ehrhart, M., et al. (2018). Assessing organizational implementation context in the education sector: Confirmatory factor analysis of measures of implementation leadership, climate, and citizenship. *Implementation Science*, 13(1), 1–14. <https://doi.org/10.1186/s13012-017-0705-6>.
- Mandell, D. S., Stahmer, A. C., Shin, S., Xie, M., Reisinger, E., & Marcus, S. C. (2013). The role of treatment fidelity on outcomes during a randomized field trial of an autism intervention. *Autism*, 17(3), 281–295. <https://doi.org/10.1177/1362361312473666>.
- Mihalic, S. F., Fagan, A. A., & Argamaso, S. (2008). Implementing the LifeSkills training drug prevention program: Factors related to implementation fidelity. *Implementation Science*, 3(1), 1–16. <https://doi.org/10.1186/1748-5908-3-5>.
- National Research Council (U.S.). (2001). *Educating children with autism*. Washington, DC: National Academy Press.
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*, 7(2), 133–144. <https://doi.org/10.1177/1477878509104318>.
- Ostroff, C., Atwater, L. E., & Feinberg, B. J. (2004). Understanding self-other agreement: A look at rater and ratee characteristics, context, and outcomes. *Personnel Psychology*, 57(2), 333–375. <https://doi.org/10.1111/j.1744-6570.2004.tb02494.x>.
- Owens, B. P., & Hekman, D. R. (2012). Modeling how to grow: An inductive examination of humble leader behaviors, contingencies, and outcomes. *The Academy of Management Journal*, 55(4), 787–818.
- Owens, B. P., & Hekman, D. R. (2016). How does leader humility influence team performance? Exploring the mechanisms of contagion and collective promotion focus. *Academy of Management Journal*, 59(3), 1088–1111. <https://doi.org/10.5465/amj.2013.0660>.
- Ozer, E. J. (2006). Contextual effects in school-based violence prevention programs: A conceptual framework and empirical review. *Journal of Primary Prevention*, 27(3), 315–340. <https://doi.org/10.1007/s10935-006-0036-x>.
- Palinkas, L. A., Aarons, G. A., Horwitz, S., Chamberlain, P., Hurlburt, M., & Landsverk, J. (2011). Mixed method designs in implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(1), 44–53. <https://doi.org/10.1007/s10488-010-0314-z>.
- 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. <https://doi.org/10.1007/s10803-015-2455-0>.
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., et al. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(2), 65–76. <https://doi.org/10.1007/s10488-010-0319-7>.
- Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>.
- Shanock, L. R., Baran, B. E., Gentry, W. A., Clever Pattison, S., & Heggestad, E. D. (2014). Erratum to: Polynomial regression with response surface analysis: A powerful approach for examining moderation and overcoming limitations of difference scores. *Journal of Business and Psychology*, 29(1), 161. <https://doi.org/10.1007/s10869-013-9317-6>.
- Shanock, L. R., Baran, B. E., Gentry, W. A., Pattison, S. C., & Heggestad, E. D. (2010). Polynomial regression with response surface analysis: A powerful approach for examining moderation and overcoming limitations of difference scores. *Journal of Business and Psychology*, 25(4), 543–554. <https://doi.org/10.1007/s10869-010-9183-4>.
- Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation*, 105(4), 269–285.
- Stahmer, A. C., Reith, S., Lee, E., Reisinger, E. M., Mandell, D. S., & Connell, J. E. (2015). Training teachers to use evidence-based practices for autism: Examining procedural implementation fidelity. *Psychology in Schools*, 52(2), 181–195. <https://doi.org/10.1002/pits>.
- Suhrheinrich, J., Stahmer, A. C., Reed, S., Schreibman, L., Reisinger, E., & Mandell, D. (2013). Implementation challenges in translating pivotal response training into community settings. *Journal of Autism and Developmental Disorders*, 43(12), 2970–2976. <https://doi.org/10.1007/s10803-013-1826-7>.
- Weiner, B. J., Belden, C. M., Bergmire, D. M., & Johnston, M. (2011). The meaning and measurement of implementation climate. *Implementation Science*, 6(1), 78. <https://doi.org/10.1186/1748-5908-6-78>.
- Whitaker, K., Fortier, A., Bruns, E. J., Nicodimos, S., Ludwig, K., Lyon, A. R., et al. (2018). How do school mental health services vary across contexts? Lessons learned from two efforts to implement a research-based strategy. *School Mental Health*, 10(2), 134–146. <https://doi.org/10.1007/s12310-017-9243-2>.
- Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 71(1), 136–149. <https://doi.org/10.1037/0022-006X.71.1.136>.
- Yammarino, F. J., & Atwater, L. E. (1993). Understanding self-perception accuracy: Implications for human resource management. *Human Resource Management*, 32(2–3), 231–247. <https://doi.org/10.1002/hrm.3930320204>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.