

Increasing Teachers' Use of Behavior-Specific Praise Using a Multitiered System for Professional Development

Journal of Positive Behavior Interventions
2017, Vol. 19(4) 239–251
© Hammill Institute on Disabilities 2017
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1098300717693568
jpbj.sagepub.com



Nicholas A. Gage, PhD¹, Ashley S. MacSuga-Gage, PhD¹,
and Emily Crews, MA¹

Abstract

Successful instruction is contingent upon effective classroom management. Unfortunately, not all teachers are effective classroom managers and many require in-service professional development (PD) to increase their use of evidence-based classroom management skills. Although PD models have been developed and evaluated, many are resource-intensive. To address this, we developed a multitiered system for PD (MTS-PD) to provide increased levels of support based on need as measured by teacher behavior. We conducted two studies to examine the effect of Tiers 1 and 2 of the MTS-PD model on elementary school teachers' use of behavior-specific praise (BSP). Study 1 examined the effect of Tier 1, a single 25-min didactic training, and Study 2 examined the effect of Tier 2, which included regular visual performance feedback delivered via email. The results of Study 1 indicated that teachers did not increase their rates of BSP to a priori defined levels of success. Study 2 found that the Tier 2 PD approach increased teachers' BSP rates, a functional relationship was established, and the rates maintained 3 months after the Tier 2 PD. Study limitations and future directions are discussed.

Keywords

professional development, behavior-specific praise, classroom-based studies, single-case designs, data analysis, studies

Successful classroom instruction is contingent upon effective classroom management to maintain appropriate student behavior, increase academic engagement, and subsequently, academic achievement (Evertson & Weinstein, 2006; Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016; Stronge, Ward, & Grant, 2011). Unfortunately, not all teachers are effective classroom managers (Reinke, Stormont, Herman, Puri, & Goel, 2011; Scott, Alter, & Hirn, 2011) and many require additional in-service professional development (PD) to increase their use of classroom management skills (Allen, Mikami, Hafen, & Pianta, 2014). Research has identified a number of classroom management skills teachers can easily implement in their classrooms, including increased opportunities to respond (OTR) and behavior-specific praise (BSP; Conroy, Alter, & Sutherland, 2014; Marzano, Marzano, & Pickering, 2003; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). Yet research has not identified the most salient approach to increase in-service teachers' use of those skills (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Therefore, we conducted two PD studies examining the effects of universal and targeted PD delivered as part of a multitiered system for PD (MTS-PD; Simonsen et al., 2014) to increase teachers' use of BSP.

Classroom Management PD

Research suggests that PD is ineffective without sustained and significant support and follow-up (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005), defined as 30 or more direct contact hours with each teacher (Yoon et al., 2007). A number of successful classroom and behavior management PD programs that include approximately 30 or more hours of direct contact have been developed and evaluated, demonstrating successful adult behavior change. For example, Hemmeter, Snyder, Fox, and Algina (2016) evaluated the implementation of the *Pyramid Model for Promoting Young Children's Social-Emotional Competence*, a behavior management program for young children. Teachers received almost 20 hr of workshops and an average of 13 practice-based coaching sessions, each including almost 2 hr of observation and 45 min of debriefing. Teachers who received

¹University of Florida, Gainesville, USA

Corresponding Author:

Nicholas A. Gage, University of Florida, 1403 Norman Hall, P.O. Box 117050, Gainesville, FL 32611, USA.
Email: gagenicholas@coe.ufl.edu

Action Editor: Mack Burke

the intensive PD (52 hr) implemented the intervention components 1.5 standard deviation units more than the control group.

The *BEST in CLASS* intervention also relies on practice-based coaching to increase teachers' implementation of classroom management. Participating teachers receive 6 hr of workshops and 14 weekly 2-hr coaching sessions (34 hr total). *BEST in CLASS* focuses on implementation of six classroom management practices, including BSP. Following the practice-based coaching, teachers implemented BSP 1.0 standard deviation units more than the control group.

Reinke et al. (2014) used the *Incredible Years Teacher Classroom Management Program* coaching model to increase elementary school teachers' implementation of classroom management. Teachers' received six 6-hr workshops across the school year and an average of 6 hr of one-on-one coaching (42 hr total). Teachers significantly increased their classroom management practices, including a 0.6 standard deviation unit increase for teachers' use of praise.

Another PD model used to increase teachers' classroom management is behavioral consultation. For example, Motoca et al. (2014) used a consultation model that involved observations, interviews, workshops, online training modules, and team and individual implementation meetings to increase teachers' use of classroom management strategies. Overall, teachers received a total of 18.5 hr of PD, and those in the treatment group significantly improved their classroom management skills. Specifically, teachers in the treatment group used positive feedback 0.4 standard deviation units more than the control group.

Although successful PD models have been developed and implemented, the number of PD hours required in each of the models is large. Furthermore, not all teachers necessarily require the same amount of PD support. For example, some teachers may be able to effectively implement classroom management skills following a workshop alone, whereas others may require additional coaching and feedback. Yet the models described above are manualized and generally a one-size-fits-all approach, with a minimum number of hours required for the PD model to have been implemented with fidelity. Therefore, to address these limitations, an MTS-PD model was developed based on the three-tiered prevention logic to increase teachers' classroom management skills.

MTS-PD

An alternative, potentially more efficient framework for delivering PD to teachers, MTS-PD, was recently developed by Simonsen and colleagues (2014) to differentiate levels of PD support for teachers based on observed need. The goal of the MTS-PD is to use PD support in schools that is both time- and resource-efficient. The model consists

of three tiers of PD, each with increasing levels of intensity based on need. In Tier 1, or universal support, all teachers participate in a short (30 min) didactic instruction-based PD session on one classroom management skill with opportunities for questions and practice. The teachers are also provided with tools for self-monitoring their use of the skill following the training.

In Tier 2, or targeted support, a coach uses data to identify teachers who are not responsive to the universal PD, defined as not delivering a skill at an a priori level. The coach meets with the teacher for approximately 20 min to reteach the skill, often shows the teacher her data (i.e., visual performance feedback; Fallon, Collier-Meek, Maggin, Sanetti, & Johnson, 2015), and provides suggestions for increasing skill use. Furthermore, the teacher can also receive repeated visual performance feedback in person or via email for at least 1 week. If necessary (i.e., teacher nonresponse), intensive individualized support can be implemented, which involves more intensive coaching, including action planning, goal setting, visual and verbal performance feedback, and modeling.

A handful of studies have examined the effectiveness of the MTS-PD framework to increase teachers' use of classroom management skills. Myers, Simonsen, and Sugai (2011) implemented an MTS-PD to increase four teachers' use of BSP. Unlike the PD described above, Tier 1 consisted of a handout and verbal recommendations about optimal rates of BSP and Tier 3 was daily performance feedback. Results suggested that most teachers increased their delivery of BSP; however, a functional relationship was not clear.

Thompson, Marchant, Anderson, Pater, and Gibb (2012) also used an MTS-PD approach to increase teachers' use of BSP for three elementary school teachers. Unlike the previous study, the authors used video self-monitoring as the Tier 2 PD, while Tier 1 and Tier 3 remained aligned with the approach described above. Again, teachers increased their rate of BSP generally, but no functional relationship was demonstrated. Simonsen et al. (2014) conducted a series of pilot case studies to further examine the potential impact of MTS-PD on BSP. Overall, results were promising, with teachers demonstrating an increase in BSP, but no functional relationship was established. Of note was that teachers in the case studies who did not respond to Tier 1 training also did not fully participate in self-monitoring.

MacSuga-Gage (2013) used the MTS-PD to increase elementary teachers' OTRs using Tier 1 and Tier 2 of MTS-PD. Four of the five teachers responded to Tier 1, while only one teacher required Tier 2 PD. All teachers increased their rates of OTR, but a functional relationship was not established. Recently, Simonsen and colleagues (2016) examined the effects of a modified Tier 2 PD approach in isolation on elementary teachers' use of BSP. Using email prompts without performance feedback to

Table 1. Teacher Characteristics for Study 1 and 2.

Characteristic	Bree	Deborah	Kathleen	Robin
Gender	Female	Female	Female	Female
Ethnicity	White	White	White	Black
Grade level	1st/2nd	1st/2nd	3rd/2nd	3rd
Years teaching	1 year/2 years	1 year/2 years	1 year/2 years	3 years
Highest degree	Master's degree	Bachelor's degree	Master's degree	Bachelor's degree
Certification	Dual certification (Elementary K-6, Special Education K-12)	Single certification (Elementary Education K-6)	Dual certification (Elementary K-6, Special Education K-12)	Single certification (Elementary K-6)
Previous classroom management experience	College coursework	1-week Kagan workshop/College coursework	College coursework	None noted

promote self-monitoring, the authors found that the Tier 2 PD increased teachers' use of BSP and a functional relationship was established. Overall, additional research is needed to determine whether the MTS-PD is an effective and efficient approach to increase teachers' use of classroom management practices.

BSP

As noted, the MTS-PD model focuses on a single, discrete classroom management skill. Research suggests that there are a number of high leverage classroom management skills that should be widely implemented (see Conroy et al., 2014, for a review), but it is unclear which skill(s) are the most salient. Gage and MacSuga-Gage (2017) conducted a multilevel analysis of teachers' discrete classroom management skills and found that BSP, defined as a positive statement delivered by a teacher contingent upon a desired behavior to inform students specifically what they did well (Simonsen et al., 2008), was the only significant predictor of positive student behavioral outcomes. Across 12 teachers and 195 observations, the average BSP was 0.40 per minute, or six per 15-min observation, congruent with prior recommended levels (Myers et al., 2011; Sutherland, Wehby, & Copeland, 2000). Therefore, research suggests that (a) BSP may be among the most salient classroom management skills and (b) teachers should be providing approximately six BSP statements per 15-min observations during large group instruction.

Purpose

To date, research on the MTS-PD framework has produced increases in classroom management skills, including BSP, but a functional relationship has not been established for the continuum (i.e., examining response across the tiers). Based on prior research, we conducted two separate studies, one for Tier 1 and one for Tier 2, to isolate the effects of the PD on teachers' use of BSP. In addition, we examined the distal

relationship between teachers' increased BSP and student behavior. Furthermore, we augmented the Tier 2 PD by including visual performance feedback during the reteaching session and in the regular (two to three times per week) email feedback to teachers. Specific research questions were as follows:

Research Question 1: Is there an increase in teachers' use of BSP following Tier 1 of the MTS-PD model?

Research Question 2: Is there a functional relationship between Tier 2 of MTS-PD and teachers' use of BSP?

Research Question 3: Is there a functional relationship between teachers' increased use of BSP and the classroom behavior of students at high risk for emotional and/or behavioral disorders?

Study 1

Method

Setting and participants. Four general education elementary school teachers (pseudonyms are used for confidentiality) from a Title I elementary school (K-5) in a large city (population > 100,000) in the Southeastern United States requested classroom management PD and volunteered to participate in the research study. Approximately, 84% of the students in the school received free or reduced-price lunch, 70% of the students were Black, and less than 40% performed at or above state benchmarks for reading and math. Teacher characteristics are presented in Table 1.

Measures

Teacher behaviors. Teachers' use of BSP was the primary dependent variable collected via direct observation. BSP was operationally defined as follows: Teacher gives an individual student or whole class a BSP, defined as a contingent verbal statement that communicates positive feedback to a student and explicitly tells student(s) what he or she did right. Examples include "I really like the way you raised

your hand” and “Thanks for being such as great listener!” Nonexamples include “Good job,” “Thanks!” and “We are going to have a really great day!”

Student behaviors. Students’ disruptive behavior and engagement with instruction were distal effect variables measured via direct observation. Operational definitions of both behaviors were as follows: (a) a student is academically engaged if he or she is actively or passively participating in the classroom activity (i.e., writing, raising hand, answering a question, talking about a lesson, listening to the teacher, reading silently, or looking at instructional materials), and (b) a student is disruptive if he or she displays behavior that does or potentially could interrupt the lesson in such a way that it distracts the teacher and/or other students (i.e., out of seat, fidgeting, playing with objects, acting aggressively, talking/yelling about things that are unrelated to classroom instruction). Both definitions were based on those used in previous research studies (Christ, Riley-Tillman, Chafouleas, & Jaffery, 2011; MacSuga-Gage, 2013).

Procedures

Recruitment. Following University Internal Review Board (IRB) approval, we recruited teachers at an informational session held after school and organized by the school’s assistant principal, who requested additional classroom management PD from the second author. Five teachers participated in the informational session and all five volunteered and consented to participate. Unfortunately, after four baseline observations, one teacher dropped out of the study due to scheduling conflicts. We asked the teachers to send home a parental notification/passive consent form to parents/guardians of all students in their classrooms informing them that a study was being conducted in their child’s classroom that would involve direct observations of teacher and student behavior. No parent denied observation consent.

Baseline. Baseline data (i.e., typical BSP delivery for each teacher) were collected for at least five observations prior to universal intervention to ensure an accurate estimate of baseline performance and to meet the recommended single-subject design standards (Kratowich et al., 2010). The average number of baseline observations across the four teachers was 12.8 (range: 5–16 observations). We collected daily observations of teachers and students during baseline; however, due to changing schedules and teacher absences, data were not collected every consecutive day for each teacher.

Tier 1: Universal PD. Once five baseline data points were collected for all four teachers, the universal training was conducted with all of the teachers simultaneously in a group

setting. The first author provided a 30-min direct instructional session with (a) an overview and definition of BSP, (b) the evidence supporting the value of increasing BSP, and (c) guidance on how to implement BSP in the classroom during large group instruction. In addition, each teacher was provided with a golf counter and instructions on how to self-monitor during the same observation period data collectors were observing. Self-monitoring was defined as teachers using the golf counter while a data collector observed and entered the data into an Excel spreadsheet located on the Microsoft OneDrive cloud with individual accounts only accessible by the teachers and the authors. Each teacher was taught (a) how to use the golf counter, (b) how to access and use the Excel spreadsheet, and (c) how to access their OneDrive accounts. After the universal PD, direct observation data continued to be collected along with teacher self-monitoring data. The a priori criterion of success for all teachers was at least six BSP per 15-min observations, or an average of 0.40 BSP per minute.

Fidelity of implementation

Fidelity of intervention training. To ensure fidelity of the PD session, the authors created and used a script for the universal (Tier 1) training. The lead author delivered the group training, while the second author evaluated the training using a fidelity checklist. The fidelity checklist contained the specific components of the PD, including whether the trainer operationally defined the skill (BSP), provided a research overview, described what the skill looks like in classroom, modeled at least three examples of the skill, allowed opportunities for all participants to model the skill, provided suggestions and advice for each teacher, and allowed the opportunity for questions. After the training, the total number of components checked as completed was divided by the total number of possible components delivered, yielding a fidelity score (i.e., the percentage of training steps implemented). Across all trainings, 100% of the intervention components were delivered.

Adherence to self-management. Fidelity of the teacher’s self-monitoring was assessed using the Excel spreadsheets on the Microsoft OneDrive University of Florida cloud. Teachers received a score of “1” if a data point was entered and “0” if a data point was not entered. The total number of data points entered by the teacher was divided by the total possible number of data points that could have been entered, yielding a fidelity score (i.e., the percentage of self-monitoring data points entered).

Accuracy of self-monitoring. We compared the researcher-collected data with the teacher self-monitoring data to assess accuracy of teacher-reported self-monitoring data. Accuracy was calculated using the percentage of agreement formula.

Table 2. Study 1: Pre–Post Universal Professional Development Training Impact on Teacher and Student Outcomes.

Teacher	Condition	Teacher behavior			Student behavior					
		BSP			Disruptions			Engagement		
		M	SD	IOA (%)	M	SD	IOA (%)	M (%)	SD (%)	IOA (%)
Bree	Pre	0.08	0.06	92	0.01	0.01	90	69.6	29.6	89
	Post	0.18	0.19	89	0.01	0.01	94	53.2	23.6	85
Deborah	Pre	0.23	0.28	88	0.00	0.00	89	92.3	8.8	91
	Post	0.33	0.24	97	0.01	0.01	85	77.0	16.8	90
Kathleen	Pre	0.43	0.42	94	0.01	0.01	90	61.0	28.6	85
	Post	0.32	0.23	90	0.02	0.02	82	47.8	24.2	90
Robin	Pre	0.03	0.04	92	0.00	0.01	89	66.7	34.9	90
	Post	0.00	0.00	94	0.01	0.01	85	51.3	34.4	94

Note. BSP = behavior-specific praise; IOA = interobserver agreement.

Direct observation procedures. We collected 15-min direct observations of each teacher during large group instruction. Each teacher was asked to identify a 20-min time period in which she consistently provided large group instruction in either reading or mathematics. We also observed three different students, chosen at random, during each observation to capture an estimate of overall class-wide performance. Data collectors were instructed to randomly choose three students at the beginning of each observation and to not include students who were observed the prior observation. The data collectors observed the teacher and the first student for the first 5 min of the observation, followed by the second student the next 5 min, and the third student the last 5 min. No data on student-level characteristics were collected.

A trained graduate research assistant or hired hourly data collector (undergraduate or graduate student) stood near the rear of the classroom and quietly observed the teacher without distracting instruction. Data collectors used Dragon Touch 18 8" Quad Core Windows Tablet PCs loaded with the Lily data collection application, part of the Multi-Option Observation System for Experimental Studies (MOOSES; Tapp, Wehby, & Ellis, 1995) system. All data collectors received a 2-hr group didactic training and conducted periodic observer drift checks to ensure the accuracy of the observations.

Interobserver agreement (IOA) procedures. We collected IOA data for 52% of all observations. Two observers stood near each other, but did not talk or interact during the observation except to start the observation at the exact same time. IOA was calculated in MOOSES using the point-by-point method with a 3-s window (Tapp et al., 1995). Across all measured behaviors, the average IOA was 90.1% (see Table 2 for IOA by teacher).

Design and analysis. We used a pre–post case study design to identify a relationship between teachers' implementation of

BSP and the universal PD. We calculated teachers' average performance before and after the universal PD. In addition, we considered the average per minute rate of student disruptive behavior, calculated by dividing the frequency of disruptions by the number of minutes for each observation (i.e., 15 min), and the percentage of time students were engaged with instruction. A prior generalizability study confirmed that at least three 15-min observations are necessary for a generalized estimate of each outcome (Gage, Prykanowski, & Hirn, 2014). All baseline means and standard deviations are based on five observations, and all post-universal PD means and standard deviations are based on at least three observations.

Results

Across the four teachers, the average per minute rate of BSP was 0.19 during baseline or less than 3 BSP statements per 15-min observation. Kathleen's baseline BSP rate was at the a priori criterion BSP rate of 0.40 per minute, suggesting that she may not need the PD, whereas two teachers delivered almost no BSP. Following the universal PD, the average per minute rate of BSP was approximately 0.20, suggesting no overall change. Two teachers demonstrated increased rates of BSP, but the rates did not approach the criterion rate.

Student outcomes also remained stable from baseline to post-PD. Overall, very few disruptions were recorded in the classrooms during either condition. The students were engaged with instruction an average of 72% of the time during baseline and an average of 57% of the time following the PD. Teacher and student outcomes by condition are presented in Table 2.

The self-monitoring data indicate that only one of the teachers collected and reported data in the Excel sheet on the OneDrive account. Bree collected data for all three observations after the training; however, the accuracy was 0% when compared with the direct observation data.

Discussion

Overall, Study 1 results indicate that universal PD, which included a 30-min direct instructional session and tools for self-monitoring, had no meaningful effect on teachers' rates of BSP. This supports prior research that indicates a single "train and hope" PD session does not result in significant behavior change (Gage, MacSuga-Gage, & Evanovich, 2015). Furthermore, the results suggest that simply providing self-monitoring tools to teachers, including self-graphing Excel sheets and golf counters, does not necessarily mean that they will use them. Therefore, based on Study 1's results, the teachers were eligible for the targeted (Tier 2) PD. With regard to the student outcomes, it is worth noting that the post-PD observations were conducted near the end of the school year, which may have confounded student engagement levels.

Study 2

Method

Setting and participants. We conducted Study 2 at the same Title I elementary school with three second-grade teachers who participated in Study 1 (two were reassigned to second grade in fall). One teacher from Study 1 (Robin) was no longer at the school. All three teachers volunteered to participate in Study 2.

For Study 2, each teacher identified one student in her classroom at risk for emotional and/or behavioral disorders using a standardized screening process. We used this approach to examine the distal effect of increasing BSP on students most at risk instead of an overall estimate of student performance. All three students were in second grade, African American, and reading below grade level based on teacher report. No additional student-level data were available.

Measures

Student screening measure. The *Student Risk Screening Scale* (SRSS) is a seven-item, student-screening tool used to identify elementary students at risk for problem behavior patterns. Research indicates that the SRSS has strong internal consistency ($\alpha > .80$) and test-retest stability ($r = .68-.74$; Lane et al., 2012). Students with summed scores greater than nine are considered at high risk.

Teacher and student behaviors. The same definitions for teacher BSP and student disruptive behavior and engagement used in Study 1 were used in Study 2.

Social validity. The *Intervention Rating Profile-15* (IRP-15) was used to collect data on the social validity of both the PD intervention and the use of BSP from the teachers' perspective. The IRP-15 is a 15-item rating scale that asks teachers to rate each item (e.g., "This would be an

acceptable intervention for the child's problem behavior.") on a 6-point Likert-type scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). Total scores range from 15 to 90, with higher scores suggesting higher acceptability. Reported internal consistency reliabilities range from $\alpha = .88$ to $.98$.

Procedures

Recruitment. Based on the same University IRB protocol from Study 1, we recruited teachers for continued PD support via an informational session. All three teachers participated in the informational session and all three volunteered and consented to participate in Study 2. Participation was not mandated by the school administration. We then asked teachers to complete the SRSS for all students in their classroom. We summed all scores and asked the teachers whether the student with the highest scores (a) had acceptable attendance (i.e., present for 80% of school days or more) and (b) was considered "at risk" by the teacher due to classroom disruptive behavior. All three teachers confirmed that the highest risk student according to the SRSS was regularly present and regularly disruptive. All three students' parents were then invited by the teacher to have their child participate and all three students' parents consented.

PD

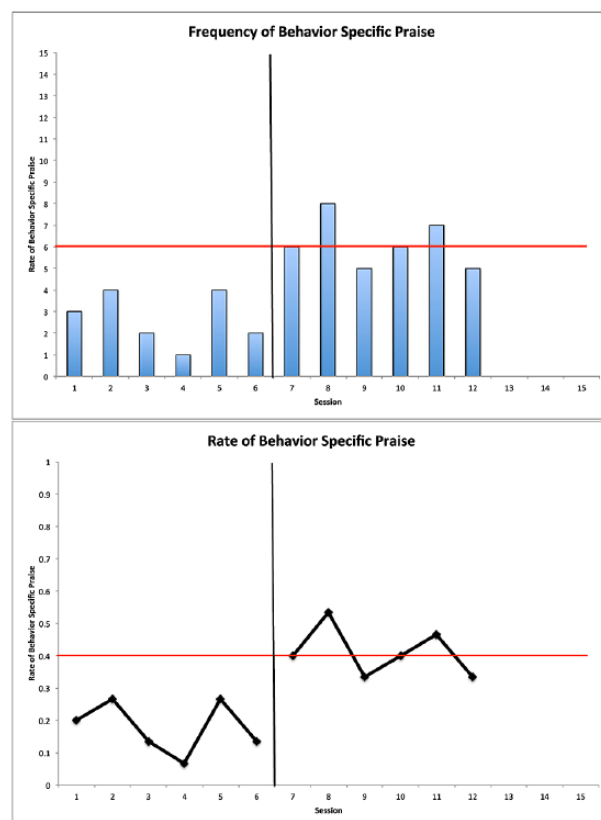
Baseline. We collected at least five baseline observations starting in October 2015 to confirm that the teachers remained at BSP rates below the a priori goal (i.e., six per 15-min observation).

Tier 2: Targeted PD. After five observations with stable BSP rates, we implemented the targeted PD with the first teacher (chosen at random among the three). During the targeted PD, we asked teachers to continue self-monitoring using the golf counters and the Excel sheets on OneDrive. We then met with each teacher in a staggered order. The decision rule for starting intervention was as follows: The second teacher began Phase 2 intervention if the teacher who received the intervention first was implementing BSP at the criterion level for at least three observations. Teacher 3 received the Phase 2 intervention if the first two teachers who received the intervention were implementing BSP at the criterion level for at least three observations.

The targeted PD included (a) a review of the BSP definition and guidance for increasing BSP; (b) a printed bar graph, line graph, and raw data from the fall baseline observations (we chose not to use the spring data to focus on fall performance only); and (c) a set goal that each teacher would deliver at least six BSP statements during reading instruction. Following the one-on-one coaching (average session lasted 20 min), we provided each teacher with weekly visual performance feedback via email, which

Hi Ashley,

You are doing great increasing your behavior specific praise! However, the other day, you were just slightly below the goal. Consider increasing your behavior specific praise by continuing to tell the students sitting correctly on the carpet that they are doing a great job sitting quietly. I also would like to see you provide more behavior specific praise to Nick, as he appears to respond very well to your positive attention. Like before, the graphs below present your frequency and per minute rate of behavior specific praise. The information to the left of the black vertical line is before training and the information to the right is after our short training. Also, don't forget to collect your own self-monitoring data!



Thanks!!!

Figure 1. Example performance feedback email.

included a bar graph and a line graph with a red goal line, positive reinforcement for their performance, and suggestions for increasing their BSP. Figure 1 is an example email message. Once all qualifying teachers entered the targeted intervention, a minimum of five data points were collected for the final participant, and once at least three of the five data points were at or above the criterion of success, the teachers participated in a final consultation meeting to review progress and plan for skill maintenance.

Follow-up probes. We conducted three follow-up probes 3 months after the last targeted PD phase data point was collected. Data collection procedures were the same as those used throughout the study. The teachers were no longer asked to self-monitor.

Social validity. At the conclusion of the study, we emailed a Qualtrics link to all three teachers and asked them to complete an electronic version of the IRP-15.

Table 3. Study 2: Descriptive Statistics and Effect Sizes for Targeted Professional Development Training Impact on Teacher Outcomes.

Teacher	Condition	Disruptions		Engagement		BSP		Effect sizes for BSP		
		M (SD)	IOA (%)	M (SD)	IOA (%)	M (SD)	IOA (%)	SMD _{all}	PND (%)	Tau-U
Bree	Baseline	0.37 (0.13)	100	62.7% (16.3%)	89	0.11 (0.06)	89			
	Intervention	0.10 (0.21)	96	91.5% (14.9%)	99	0.75 (0.28)	98	10.8	100	1.00**
	Maintenance	0.22 (0.04)	97	67.4% (24.0%)	89	0.87 (0.07)	90	12.7	100	1.00*
Kathleen	Baseline	0.73 (0.51)	97	30.8% (27.4%)	92	0.40 (0.21)	87			
	Intervention	0.64 (0.28)	86	20.6% (20.5%)	90	1.46 (0.50)	97	5.1	90	0.97***
	Maintenance	0.53 (0.20)	100	49.7% (22.1%)	99	1.02 (0.10)	89	2.9	100	1.00*
Deborah	Baseline	0.24 (0.33)	90	92.0% (12.7%)	97	0.19 (0.23)	87			
	Intervention	0.21 (0.21)	94	86.4% (13.5%)	88	0.80 (0.23)	93	2.7	100	1.00**
	Maintenance	0.22 (0.04)	100	98.4% (2.7%)	97	0.87 (0.07)	100	3.0	100	1.00**

Note. Disruptions are rate per minute. Effect sizes contrast baseline and intervention, and baseline and maintenance phases. BSP = behavior-specific praise; IOA = interobserver agreement; SMD_{all} = standardized mean difference for all data points; PND = percentage of nonoverlapping data points. * $p < .05$. ** $p < .01$. *** $p < .001$.

Fidelity of implementation of intervention

Fidelity of intervention training. To ensure consistency across the PD sessions, the authors created and used a script for all targeted trainings for all teachers. The first author delivered all three trainings, while the second author observed each training and monitored delivery of content using a fidelity checklist. Across all trainings, 100% of the intervention components were delivered.

Adherence to self-management. Fidelity of the teacher-completed self-monitoring component of the intervention was again assessed using the Excel spreadsheets on the Microsoft OneDrive. Teachers received a score of “1” if a data point was entered and “0” if a data point was not entered.

Accuracy of self-monitoring. We compared the researcher-collected data with the teacher self-monitoring data to assess the accuracy of teacher-reported self-monitoring data.

Direct observation procedures. We collected 15-min direct observations of each teacher during reading instruction following the same procedures outlined in Study 1, except that we focused on an individual student instead of three random students. The lead author and the third author collected all data using two Microsoft Surface Pro 3 Windows Tablet PCs loaded with the Lily data collection application and the classroom management code file.

IOA procedures. We collected IOA data for 53% of direct observations following the same procedures as Study 1. The average IOA across all phases was 93.7% (see Table 3 for IOA by teacher).

Research design and data analysis. We used a multiple-baseline across subjects design to examine the effect of targeted

PD on teachers' implementation of BSP. All data arrays were assessed using visual analysis. Visual analysis focused on six criteria as defined by *What Works Clearinghouse* (WWC): level, trend, variability, immediacy of the effect, overlap, and consistency of data patterns across similar phases (Kratochwill et al., 2010). The visual analysis included four steps: (a) identifying a predictable baseline; (b) assessing within-phase data pattern (level, trend, and variability); (c) comparing level, trend, and variability of adjacent phases (within participant); and (d) comparing level, trend, and variability across subjects to confirm replication of effect. The first author, a WWC certified single-subject reviewer, conducted the visual analysis and the second author confirmed the results. We also calculated a series of effect sizes (ESs) to compliment the visual analysis to improve the credibility, reliability, and defensibility of our findings (Vannest & Ninci, 2015). To compare results across different types of ES (Gage & Lewis, 2013), we calculated percentage of nonoverlapping data points (PND; Scruggs, Mastropieri, & Casto, 1987), the standardized mean difference of all data points (SMD_{all}; Busk & Serlin, 1992), the single-subject Hedges's g , based on a hierarchical linear modeling approach (Hedges, Pustejovsky, & Shadish, 2012), and the nonparametric Tau- U (Parker, Vannest, Davis, & Sauber, 2011).

Results

Both Bree and Deborah delivered an average of two BSP statements per observation during baseline, whereas Kathleen again delivered BSP at the criterion level of six per 15-min observation. After five observations, Bree received one-on-one targeted PD and performance feedback. The performance feedback was emailed after two observations so that data could be presented. A graphic display is presented in Figure 2. Following the first observation after the targeted PD, there

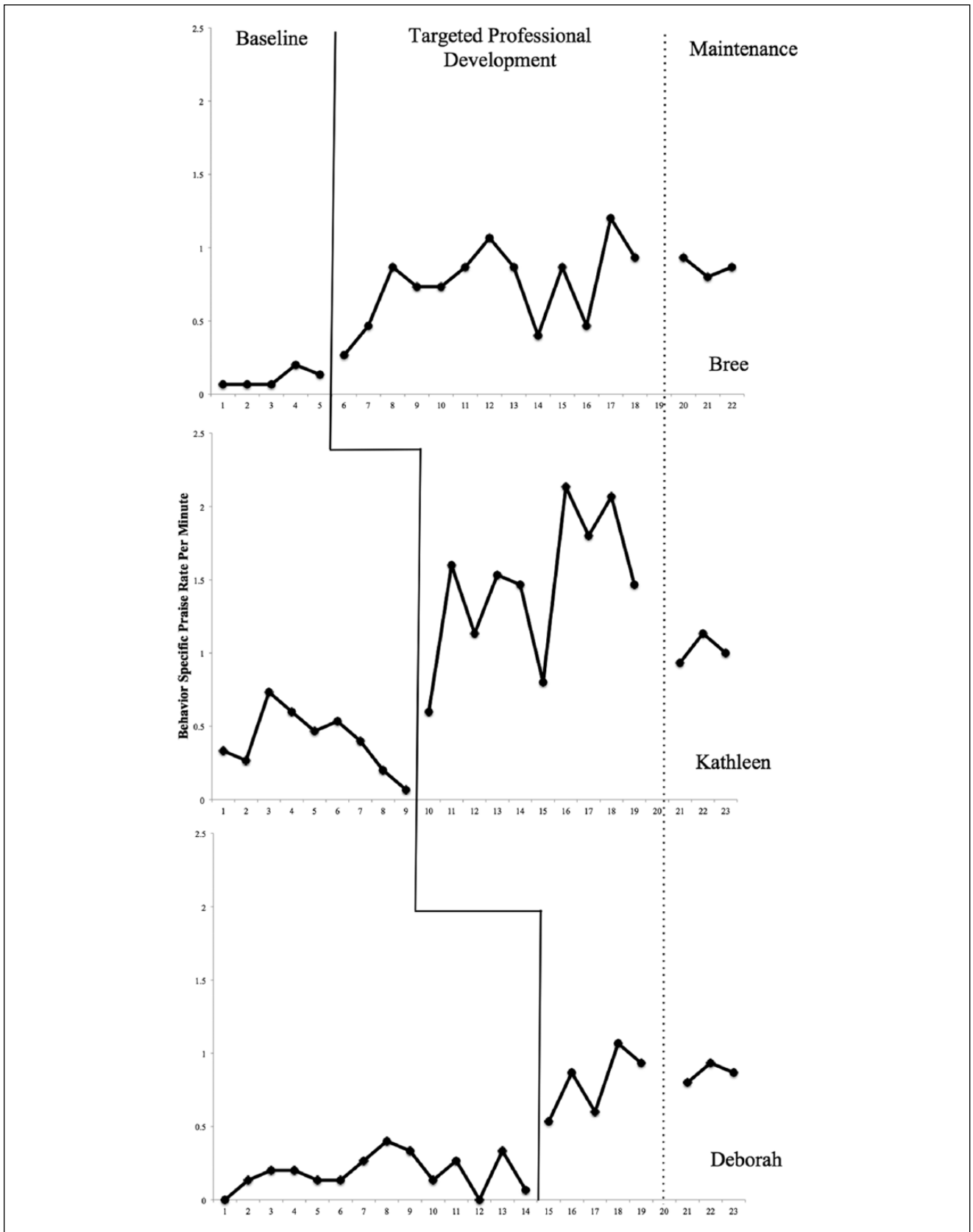


Figure 2. Teachers' behavior-specific praise following targeted professional development.

was a steep increase in slope and a sustained average per minute rate of 0.75 BSP, or just over 11 BSP statements per 15-min observation. Concurrent to the first four post-PD observations for Bree, Kathleen's per minute rate decreased to almost no BSP statements. Following the targeted PD, Kathleen's per minute rate of BSP immediately increased and had a slightly increasing slope during the rest of the targeted PD phase. Although there was variability and one overlapping data point, all rates were well above the 0.40 per minute criterion for success. Overall, Kathleen's average per minute rate of BSP was 1.46 following the targeted PD, or 21.9 BSP statements per 15-min observation. We collected six concurrent observations for Deborah following starting Kathleen's targeted PD due to the variability to ensure stability prior to introducing the targeted PD. Deborah's average rate of BSP was 0.18 during baseline. Following the introduction of the targeted PD, her per minute rate demonstrated an immediate increase in level and in slope. Deborah's average per minute rate of BSP following the targeted PD was 0.80, well above the a priori criterion. Maintenance data indicated that even after 3 months, all three teachers' per minute rates were at an average of 0.92 per observation, or almost 14 BSP statements per 15-min observation. Although teachers were asked to continue self-monitoring, none of the teachers collected any self-monitoring data.

We calculated a series of single-subject design ESs to support the visual analysis findings (see Table 3 for ESs per baseline and intervention, and baseline and maintenance phases). The average PND for the impact of targeted PD was 97%, with the only overlapping data point occurring in Kathleen's data, which was predominantly above the criterion for success. The single-case Hedges's *g* ES, calculated using the *scdhl*m package in R (Pustejovsky, 2015), was 1.74 with a variance of 0.38, indicating that teachers' BSP rates increased by almost 1.75 standard deviation units following the targeted PD. Last, we calculated Tau-*U* as it provides a nonparametric significance level. Overall, we found a combined Tau-*U* of 0.99 ($p < .000$), indicating a statistically significant increase in BSP following the targeted PD.

Student-level results are reported as distal effects of the targeted PD. As noted, all students observed were the most at risk for emotional and/or behavioral disorders in each teacher's classroom. It is important to note that the BSP statements were not targeted at the individual student, but to either the whole class or any individual student in the classroom. Therefore, the student-level results indicate the impact of increased BSP in the classroom on a target student's behavior. The target student in Bree's classroom had an average rate of 0.37 disruptions per minute ($SD = 0.13$), or approximately five per 15-min observation, and was engaged only 63% percent of time ($SD = 16%$). Following the targeted PD, the student's rate of disruptions decreased to only 0.13 disruptions per minute ($SD = 0.20$), or two disruptions per 15-min observation. His percentage of time engaged also increased to 87% of the time ($SD = 18%$).

Results were not as promising for Kathleen and Deborah. The average rate of disruptions for the student in Kathleen's classroom was 0.73 per minute ($SD = 0.51$) during baseline and 0.62 per minute ($SD = 0.25$) following the targeted PD. His engagement was only 31% during baseline ($SD = 27%$) and 27% following the targeted PD ($SD = 22%$). The rate of disruptions was less for the student in Deborah's classroom, 0.24 per minute during baseline ($SD = 0.33$) and 0.22 following the targeted PD ($SD = 0.16$). Results were similar for engagement, with an average of 92% of the time engaged during baseline ($SD = 13%$) and 91% following the targeted PD ($SD = 12%$; graphic display available from first author).

Social validity results indicate that the three teachers generally agreed that the MTS-PD approach would be acceptable to address the PD needs of teachers and that they would suggest the use of the approach to other teachers. The only item that all teachers rated as "strongly agree" was that the behavior problems they experienced in their classroom were severe enough to warrant the use of PD. Overall, the teachers indicated that they "agreed" that the MTS-PD approach would be beneficial for teachers.

Discussion

It was clear that there was a functional relationship between targeted PD and teachers' implementation of BSP in their classrooms. All three teachers increased their rates of BSP well above the recommended rate during large group instruction. Perhaps most encouraging was the sustained increase as measured by the maintenance data, which were collected 3 months after the last email performance feedback was sent. Furthermore, it is worth noting that teachers increased their BSP rates during the targeted PD even though they did not conduct any self-management data collection. Unfortunately, the increase in BSP only had a direct impact on one of the three target students. Although this finding is discouraging, the relationship between increased universal classroom management practices, including BSP, and student performance are not well established. Based on the results of this and previous research, it is clear that students exhibiting the most intensive classroom-based problems probably need individualized interventions to decrease disruptions and increase student engagement.

General Discussion

The goal of this study was to evaluate the effectiveness of MTS-PD, a resource-efficient PD model, to increase elementary teachers' use of BSP. The two studies were developed to provide the most comprehensive analysis of the MTS-PD to date that experimentally evaluated the effects of Tier 2 PD following Tier 1 PD. The teachers did not respond to the Tier 1 PD, defined as increasing their use of BSP to at least six statements per 15-min observation during large group instruction, providing additional evidence

that short, didactic trainings without follow-up may not change behavior for some teachers (cf. Gage, MacSuga-Gage, & Evanovich, 2015). However, teachers did respond to the Tier 2 PD, which included regular visual performance feedback via email, and the increase in BSP statements was still evident 3 months after the performance feedback stopped. Furthermore, the model required only 30 min for the universal PD and an average of 6 hr for the targeted PD (30-min training and an average of twenty-two 15-min observations).

We hypothesize that the sustained increase of BSP as a result of the Tier 2 PD may have been due to teachers “buying-in” because they saw their frequency of BSP increase via performance feedback and experienced an impact on student behavior and the general classroom climate. We anecdotally note that the teachers began to use BSP instead of negative feedback following a student disruption even though we did not directly teach this approach. Unfortunately, we did not collect classwide data during Study 2 and instead focused on a high-risk student. Anecdotally, we noted fewer disruptions and more engagement classwide; therefore, future research should examine both the relationship between BSP and negative feedback as well as the effect of increased frequency of BSP and classwide student performance.

Although we had hoped to find a functional relationship between increases in BSP and positive student behavior for the student identified as at highest risk, the lack of experimental evidence further highlights that classroom management is the first step in building a continuum of evidence-based behavior interventions in the classroom. High-quality classroom management should prevent significant or recurring behavior problems for approximately 80% of the students in the classroom (Simonsen et al., 2008). Put another way, targeting the most at-risk students and increasing BSP did not match the needs of the student, by intensity and perhaps by function. Yet in one classroom we did find a reduction of problem behavior for a student at high risk. Therefore, future research should examine the effect of the continuum of behavioral interventions in the classroom on the most at-risk students’ behavior, starting with high-quality classroom management such as increased use of BSP.

Limitations

A number of limitations necessitate discussion. First, the full model could not be tested, which would include a larger group or all teachers in a school, as Tier 1 would indicate who needs additional PD support. The four participating teachers in this study may have been more typical of teachers in need of “Tier 2” support given that they struggled with classroom management and requested PD. Second, consistent with prior MTS-PD research, the PD was delivered by university-based researchers and not natural

implementers. Future research is needed to determine whether or not effects would be different if school-based personnel implemented the approach. Similarly, direct observations were conducted by the research team, which may have also influenced teacher behavior. Last, single-subject design may not be appropriate for evaluating the full MTS-PD model. Future research should examine the effectiveness of the MTS-PD model using group experimental designs to evaluate the model in its entirety (i.e., randomly assigning teachers to receive the MTS-PD).

Conclusion

Evidence continues to mount that teachers are not ubiquitously implementing evidence-based classroom management (Reinke et al., 2011; Scott et al., 2011) and may be in need of in-service PD. The results of this study, as well as prior research, suggest that (a) didactic training (Tier 1 in an MTS-PD model) alone may not be an intense enough intervention for some teachers and that (b) emailed visual performance feedback (Tier 2 in an MTS-PD) can increase teachers’ use of evidence-based classroom management skills, such as BSP. There is no doubt that changing adult behavior can be as difficult, and may be even more difficult than changing student behavior. Yet if we take what we know works for changing student behavior, such as increasing the intensity of intervention efforts based on accurate and reliable data, and apply it to PD, the same positive effects could be evident. Future research will need to find creative ways to scale-up this approach schoolwide due to the data collection requirements. Regardless of the scale-up challenges ahead, we believe in, and evidence is beginning to support, the potential value of an MTS-PD model.

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Allen, J. P., Mikami, A. Y., Hafen, C. A., & Pianta, R. C. (2014). Effects of a professional development program on behavioral engagement of students in middle and high school. *Psychology in the Schools, 51*, 143–163.
- Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single case research design and analysis: New directions for psychology and education* (pp. 187–212). Hillsdale, NJ: Lawrence Erlbaum.
- Christ, T. J., Riley-Tillman, T. C., Chafouleas, S., & Jaffery, R. (2011). Direct behavior rating: An evaluation of alternate

- definitions to assess classroom behaviors. *School Psychology Review*, 40, 181–199.
- Conroy, M. A., Alter, P. J., & Sutherland, K. S. (2014). Classroom-based research in the field of EBD: Current and future research directions. In P. Garner, J. Kauffman, & J. Elliot (Eds.), *Handbook of emotional and behavioral difficulties* (2nd ed., pp. 465–477). London, England: SAGE.
- Evertson, C. M., & Weinstein, C. S. (2006). *Handbook of classroom management: Research, practice, and contemporary issues*. New York, NY: Lawrence Erlbaum.
- Fallon, L. M., Collier-Meek, M. A., Maggin, D. M., Sanetti, L. M. H., & Johnson, A. H. (2015). Is performance feedback for educators an evidence-based practice? A systematic review and evaluation based on single-case research. *Exceptional Children*, 81, 227–246.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication #231). Tampa: The National Implementation Research Network, Louis de la Parte Florida Mental Health Institute, University of South Florida.
- Gage, N. A., & Lewis, T. J. (2013). Analysis of effect for single-case design research. *Journal of Applied Sport Psychology*, 25, 46–60. doi:10.1080/10413200.2012.660673
- Gage, N. A., & MacSuga-Gage, A. S. (2017). Salient classroom management skills: Finding the most effective skills to increase student engagement and decrease disruptions. *Report on Emotional & Behavioral Disorders in Youth*, 17, 19–24.
- Gage, N. A., MacSuga-Gage, A. S., & Evanovich, L. L. (2015). Training teachers to use antecedent-based classroom management strategies to support inclusion of students with intellectual disabilities: A feasibility study. *Journal of Global Research in Education and Social Science*, 6, 168–176.
- Gage, N. A., Prykanowski, D., & Hirn, R. (2014). Increasing reliability of direct observation measurement approaches in emotional and/or behavioral disorders research using generalizability theory. *Behavioral Disorders*, 39, 228–244.
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods*, 3, 224–239. doi:10.1080/09602011.2013.819021
- Hemmeter, M. L., Snyder, P. A., Fox, L., & Algina, J. (2016). Evaluating the implementation of the Pyramid Model for promoting social-emotional competence in early childhood classrooms. *Topics in Early Childhood Special Education*, 36, 133–146. doi:10.1177/0271121416653386
- Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, 86, 643–680. doi:10.3102/0034654315626799
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from <https://ies.ed.gov/ncee/wwc/Document/229>
- Lane, K. L., Oakes, W. P., Harris, P. J., Menzies, H. M., Cox, M., & Lambert, W. (2012). Initial evidence for the reliability and validity of the Student Risk Screening Scale for internalizing and externalizing behaviors at the elementary level. *Behavioral Disorders*, 37, 99–122.
- MacSuga-Gage, A. S. (2013). *Supporting teachers' professional development: Investigating the impact of a targeted intervention on teachers' presentation of opportunities to respond*. Retrieved from <http://digitalcommons.uconn.edu/dissertations/103/>
- Marzano, R. J., Marzano, J. S., & Pickering, D. J. (2003). *Classroom management that works: Research-based strategies for every teacher*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Motoca, L. M., Farmer, T. W., Hamm, J. V., Byun, S. Y., Lee, D. L., Brooks, D. S., . . . Moohr, M. M. (2014). Directed consultation, the SEALS model, and teachers' classroom management. *Journal of Emotional and Behavioral Disorders*, 22, 119–129. doi:10.1177/1063426614521299
- Myers, D. M., Simonsen, B., & Sugai, G. (2011). Increasing teachers' use of praise with a response-to-intervention approach. *Education & Treatment of Children*, 34, 35–59.
- Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, 42, 284–299.
- Pustejovsky, J. E. (2015). *scdhlmm: Estimating hierarchical linear models for single-case designs (Version 0.2.1)* [R package version]. Retrieved from <http://blogs.edb.utexas.edu/pusto/software/>
- Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly*, 26, 1–13.
- Reinke, W. M., Stormont, M., Herman, K. C., Wang, Z., Newcomer, L., & King, K. (2014). Use of coaching and behavior support planning for students with disruptive behavior within a universal classroom management program. *Journal of Emotional and Behavioral Disorders*, 22, 74–82. doi:10.1177/1063426613519820
- Scott, T. M., Alter, P. J., & Hirn, R. (2011). An examination of typical classroom context and instruction for students with and without behavioral disorders. *Education & Treatment of Children*, 34, 619–642.
- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research methodology and validation. *Remedial and Special Education*, 8, 24–33.
- Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education & Treatment of Children*, 31, 351–380.
- Simonsen, B., Freeman, J., Dooley, K., Maddock, E., Kern, L., & Meyers, D. (2016). Effects of targeted professional development on teachers' specific praise rates. *Journal of Positive Behavior Interventions*, 19, 37–47. doi:10.1177/1098300716637192
- Simonsen, B., MacSuga-Gage, A. S., Briere, D. E., Freeman, J., Myers, D., Scott, T. M., & Sugai, G. (2014). Multitiered support framework for teachers' classroom-management practices: Overview and case study of building the triangle

- for teachers. *Journal of Positive Behavior Interventions*, 16, 179–190. doi:10.1177/1098300713484062
- Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, 62, 339–355. doi:10.1177/0022487111404241
- Sutherland, K., Wehby, J., & Copeland, S. (2000). Effects of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders*, 8, 2–8.
- Tapp, J., Wehby, J. H., & Ellis, D. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavioral Research Methods, Instruments, and Computers*, 27(1), 25–31.
- Thompson, M. T., Marchant, M., Anderson, D., Prater, M. A., & Gibb, G. (2012). Effects of tiered training on general educators' use of specific praise. *Education & Treatment of Children*, 35, 521–546.
- Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling & Development*, 93, 403–411. doi:10.1002/jcad.12038
- Yoon, K. S., Duncan, T., Lee, S. W., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007–No. 033). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee/edlabs>