Hammill Institute on Disabilities

Journal of Positive Behavior Interventions
2017, Vol. 19(4) 228–238

DOI: 10.1177/1098300716683634
jpbi.sagepub.com
SAGE

Improving Implementation of Function-Based Interventions: Self-Monitoring, Data Collection, and Data Review

Sarah E. Pinkelman, PhD, BCBA-D[1] and Robert H. Horner, PhD[2]

Abstract

The success of function-based interventions depends not just on the quality of procedures but also on the extent to which procedures are implemented as planned. Too often in schools, effort is committed to functional assessment and behavior support plan design, only to be followed by weak implementation. This study used a multiple baseline across participants design to examine whether a functional relation exists between a treatment package consisting of (a) self-monitoring treatment fidelity, (b) collecting data on student behavior, (c) entering fidelity and student behavior data into an online data management application, and (d) reviewing those data weekly using graphs generated by the application. Results indicate improved treatment fidelity and student behavior were associated with the treatment package. Potential contributions are discussed in terms of establishing efficient systems for schools that provide the structure for effective implementation of function-based interventions.

Keywords

positive behavior support, implementation, function-based interventions and supports, treatment fidelity

Introduction

Substantial empirical evidence indicates the effectiveness of function-based interventions for students who engage in problem behavior (see Dunlap & Carr, 2007; Dunlap et al., 2003; Dunlap & Fox, 2012; Ingram, Lewis-Palmer, & Sugai, 2005; Iwata, Dorsey, Slifer, Bauman, & Richman, 1994; Marquis et al., 2000; Newcomer & Lewis, 2004). Function-based interventions are guided by comprehensive assessment information (i.e., functional analysis [FA] or functional behavior assessment [FBA]) and organized around a written plan of support (i.e., behavior support plan [BSP]) that outlines antecedent, consequent, and teaching strategies that address the function of problem behavior. Once a BSP is developed, the next task, and arguably the more difficult task, is to ensure the BSP is effectively implemented in the classroom. Treatment fidelity, or the extent to which interventions are delivered as planned and described (Gresham, Gansle, & Noell, 1993), is essential to produce targeted results. Ensuring high treatment fidelity is particularly important in the area of behavior support, where student problem behavior functions as a barrier to both social and academic success in school and beyond. Unfortunately, teachers are often not provided with sufficient preservice training that enables them to implement behavioral interventions with high fidelity in their classrooms (Begeny & Martens, 2006; Simonsen, MacSuga, Fallon, & Sugai, 2013), and they are not always provided with adequate organizational support (i.e., coaching, data systems, administrative assistance) to maintain implementation of acquired skills (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005).

Performance feedback is a strategy often used in staff training and consultation to improve the fidelity with which practitioners implement interventions. In a synthesis of the implementation literature, Fixsen et al. (2005) found performance feedback to be an essential training tool, concluding, "training by itself does not result in positive implementation outcomes (changes in practitioner behavior in the clinical setting) or intervention outcomes (benefits to consumers)" (p. 40–41). Similarly, in a meta-analysis of the professional development literature, Joyce and Showers (2002) emphasized that training must include opportunities for teachers to demonstrate skills and receive feedback from an expert or coach. Despite its effectiveness, performance feedback is often not realistic for schools, as most

[1]Utah State University, Logan, USA
[2]University of Oregon, Eugene, USA

Corresponding Author:
Sarah E. Pinkelman, Utah State University, 2865 Old Main Hill, Logan, UT 84322-1400, USA.
Email: sarah.pinkelman@usu.edu

Action Editor: Matt Tincani

schools do not have the means to implement such time-consuming and resource-extensive training practices.

Self-monitoring is another strategy that has been used to improve the fidelity with which interventions are implemented. Self-monitoring has been used to increase the number of praise statements delivered to students (Hager, 2012; Kalis, Vannest, & Parker, 2007; Keller, Brady, & Taylor, 2005; Simonsen et al., 2013; Sutherland & Wehby, 2001; Workman, Watson, & Helton, 1982), improve the fidelity of discrete trial instruction (Belfiore, Fritts, & Herman, 2008), increase the number of learning trials delivered to students (Lylo & Lee, 2013), and improve the fidelity with which staff implement token economies (Plavnick, Ferreri, & Maupin, 2010; Seligson-Petscher & Bailey, 2006). Given the challenges associated with delivering performance feedback in schools, self-monitoring might provide schools with a cost-efficient method to improve the fidelity with which staff implement function-based interventions.

In addition to implementing interventions with fidelity, another crucial factor to the effective implementation of function-based interventions is collecting data and reviewing those data for decision making. Although there is increased awareness that educators must use empirical evidence to inform decisions and monitor student performance (Mandinach & Gummer, 2013; Newton, Horner, Algozzine, Todd, & Algozzine, 2012), data-based decision making in schools continues to be a challenge (Dunn, Airola, Lo, & Garrison, 2013; Schildkamp, Ehren, & Lai, 2012; Telzrow, McNamara, & Hollinger, 2000). To make meaningful decisions from data, data must be collected on both independent and dependent variables (Gresham et al., 1993), meaning BSP implementation (independent variable) and student behavior (dependent variable). Both measures are essential and too often neglected (Detrich, 2014; Fixsen, Blase, Metz, & Van Dyke, 2013).

Purpose and Research Questions

The purpose of this study was to examine the effects of a treatment package in supporting school staff to implement function-based interventions in public school settings. As mentioned above, effective implementation requires high treatment fidelity, collecting data on fidelity and student behavior, and reviewing these data for decision making. To assist in the efficiency with which school personnel were able to carry out these activities, we used an online data management application, the Individual Student Information System: School-Wide Information System (ISIS-SWIS; www.pbisapps.org). ISIS-SWIS allowed participants to upload fidelity and student data and review these data in automatically generated graphs. ISIS-SWIS was not the intervention, but rather a system that provided the organization and structure for the intervention. It functioned as a *decision support data system* for continuous improvement, implementation support, and organizational functioning (Fixsen, Blase, Naoom, & Duda, 2013).

In this study, we developed a treatment package where staff (a) self-monitored treatment fidelity, (b) collected data on student problem behavior, (c) entered fidelity and student data into ISIS-SWIS, and (d) reviewed data in ISIS-SWIS for decision making. Specifically, we assessed whether there is a functional relation between the treatment package and BSP treatment fidelity, student problem behavior, and student academic engagement. The study also examined the extent to which self-monitoring of treatment fidelity (via ISIS-SWIS) aligned with observed treatment fidelity.

Method

Participants and Settings

Three dyads nested in two elementary schools participated in the study. Dyads consisted of a student and an educational assistant (EA) in self-contained special education classrooms, where students spent the majority of their day in the special education classroom but also participated in lunch, recess, specials, and some elective classes with their typically developing peers. Dyads were considered for inclusion in the study if (a) the student was currently receiving interventions to address problem behavior, (b) classroom staff were implementing the BSP with low treatment fidelity as judged by the district behavior specialist, and (c) student problem behavior remained at an unacceptable level. The district behavior specialist recommended potential dyads, and the primary researcher conducted preliminary observations to ensure the above inclusion criteria were met.

Table 1 outlines participating students, problem behavior, the hypothesized function of problem behavior, and components their BSP. Dyad 1 consisted of a fourth-grade student named Trent who attended school for the entire school day (8:45 a.m. to 2:45 p.m.) and an EA named Evie. Trent was male and identified with specific learning disability. Trent's problem behavior consisted of disruption, task refusal, property destruction, and off-topic talk/stalling (see operational definitions in the "Dependent Measures" section). An FBA hypothesized Trent's problem behavior was maintained by escape from demands. Trent's BSP had a total of 14 components, four of which were strategies that should be observed during each observation (the remaining components were those that could be scored as no opportunity). Evie had been working in Trent's classroom for 16 years. She previously received district-wide training in positive behavior support as well as in-classroom feedback from the classroom teacher. For Trent and Evie, observations took place during small group reading from 9:00 a.m. to 9:20 a.m. The small group consisted of Trent and a peer,

Table 1. BSP Components for Each Student.

Student	Problem behavior	Hypothesized function	BSP components
Trent	Disruption, task refusal, property destruction, off-topic talk/stalling	Escape from academic tasks	Provide 5 min break for every 30 min of work Prompt break/help at sign of precursor behavior Ensure "break" and "help" visual are on desk Precorrect break/help prior to teacher led activities Precorrect break/help/modification to task prior to independent work Provide verbal praise/color spot about every 5 min Grant 5 min of preferred activity for 10 color spots When he requests help, provide assistance When he requests break, allow 5 min break When he requests modified task, allow modification If problem behavior, redirect to request break/help/modification Repeat redirection every 2 min When he returns from break, precorrect break/help When he returns from break, have him complete the activity he was working on before break
Marin	Protest, aggression, property destruction, disruption	Escape from nonpreferred academic tasks	Prompt to check visual schedule at each transition Set visual timer to indicate how long work will last Build choice into work activities Intersperse easy tasks with difficult tasks Before transitions, indicate when transition will occur; deliver token immediately at destination Precorrect help/break before academic activities When he requests break/help, provide break/help and deliver verbal praise and a token Deliver tokens approximately every 3 min Provide descriptive verbal praise during token delivery Ask him to select backup reinforcer before work When he earns all tokens, provide access to reinforcer Allow him to put all tokens on and take all tokens off Increase rate of reinforcement during difficult tasks If he transitions from reinforcer at initial request, provide token and descriptive verbal praise Remind him how many tokens left to earn reinforcer If problem behavior, prompt break/help Differentially reinforce appropriate behavior Use no more than 2 sentences to prompt or redirect
Joey	Elopement, protest, off-topic talk	Escape from academic tasks	Provide 5 reinforcing statements to every corrective Provide descriptive verbal praise at least 3 times Have schedule on table and visible to student Schedule outlines academic activities and corresponding "student choice" activities Set timer for 6 to 8 min for student choice activities Complete schedule with student before period begins Place Pride Cards/color spots on table and visible Deliver Pride Card/color spot about every 5 min Once color spot card is complete, allow him to choose prize from treasure chest Allow him to cash in Pride Cards on Friday Precorrect break/help before transitioning to work If he requests break, allow break If he requests help, provide assistance Provide verbal praise for requesting break/help If problem behavior, redirect to finish task or request break/help Provide verbal praise and a color spot/Pride Card when he returns to work

Note. Full definitions of BSP components are available from the first author. BSP = behavior support plan.

and Evie led the reading lesson. During this group, students participated in remedial direct instruction reading lessons or completed independent work from their workbooks.

Dyad 2 consisted of a kindergarten student named Marin who attended school in the morning (8:30 a.m. to 11:30 a.m.) and an EA named Chris. Marin was male and identified with a communication disorder. Marin's problem behavior consisted of protest, aggression, property destruction, and disruption (see operational definitions in the "Dependent Measures" section). An FBA hypothesized Marin's problem behavior was maintained by escape from nonpreferred academic demands. Marin's BSP had 18 total components, eight of which should be observed during each observation. Chris had 7 years of experience working with students with special needs and attended a district-wide workshop in positive behavior support during her first year of employment. For Marin and Chris, observations took place from 9:30 a.m. to 9:50 a.m. when Marin worked one-on-one with Chris on Language for Learning® lessons.

Dyad 3 consisted of a first-grade student named Joey and an EA named Natalie. Joey was male and identified with other health impairment. He was on a modified schedule, where he attended school in the morning (8:00 a.m. to 11:45 a.m.). Joey's problem behavior consisted of elopement, protest, and off-topic talk (operational definitions in the "Dependent Measures" section). An FBA hypothesized Joey's problem behavior was maintained by escape from demands. Joey's BSP had 16 total components, seven of which should be observed during each observation. Natalie had 22 years of experience working with students with special needs. She received previous district-wide training in positive behavior support. For Joey and Natalie, observations took place during one-on-one academic instruction from 8:15 a.m. to 8:35 a.m. The content area addressed during this time varied depending on Joey's schedule for that day.

Dependent Measures

Direct observation data. Direct observations occurred 3 to 5 times per week and lasted 20 min in duration. Trained observers collected direct observation data on treatment fidelity and student problem behavior and academic engagement during the predetermined observation periods described above (small group for Trent and Evie, Language for Learning® lessons for Marin and Chris, and one-on-one academic instruction for Joey and Natalie).

Direct observation of treatment fidelity. To determine the extent to which EAs implemented student BSPs with fidelity, trained observers collected direct observation data on the number of BSP components implemented correctly. During observation periods, observers used a data sheet outlining components specific to each student's BSP and indicated whether each component was implemented or not implemented, or whether there was no opportunity (e.g., reinforcement of replacement behavior did not occur if replacement behavior was not emitted, or use of consequence-based strategies could not be assessed if problem behavior did not occur). From these data, the percentage of BSP components implemented was derived (i.e., number of components implemented divided by total possible components multiplied by 100%). For components that could be implemented more than once in an observation (e.g., reinforcement of replacement behavior), the EA had to implement the component for each opportunity (e.g., each time the student engaged in the replacement behavior) to receive the score of "implemented" for that component on the data sheet.

Direct observation of student behavior. Trained observers collected direct observation data on the occurrence of student problem behavior and academic engagement using 10-s partial interval recording. Trent's problem behavior consisted of disruption, task refusal, property destruction, and off-topic talk/stalling. Disruption was defined as tapping items on desk (e.g., pencil, book), touching peers with hands or other objects, or talking or making noises with mouth at a time when the expectation was to be quiet. Task refusal was defined as not initiating a staff request within 5 s. Property destruction was defined as tearing instructional materials or crumpling paper with hands. Off-topic talk/stalling was defined as asking questions or making comments that do not relate to the current task, or asking unnecessary/simple questions that relate to the task but do not result in information that is necessary to complete the task.

Marin's problem behavior consisted of protest, aggression, property destruction, and disruption. Protest was defined as saying "no," "I'm done," or other words to indicate he would not comply with a staff member's request, or not initiating a request within 5 s. Aggression was defined as grabbing any part of a staff member's body with his hand(s) and squeezing their body by clenching his hand(s). Property destruction was defined as throwing items, tearing instructional materials, or pushing materials off of his desk. Disruption was defined as making noises with mouth, yelling (volume of voice above that of a conversational level), or singing at a time when the expectation was to be quiet or respond to an instructional request.

Joey's problem behavior consisted of elopement, protest, and off-topic talk. Elopement consisted of walking or running more than 3 feet away from the designated area without staff permission. Protest was defined as yelling (volume of voice above that of a conversational level), saying "no" or other words to indicate he would not comply with staff request, or not initiating a staff request within 5 s. Off-topic talk consisted of asking questions or making comments that did not relate to the current task.

Trained observers also collected data on academic engagement for all students. Academic engagement during a 10 s interval was defined as the student's shoulders and eyes directed toward the assigned activity or instructional task for a minimum of 5 consecutive seconds. This included shoulders and eyes directed toward staff when staff were giving instructions or talking to the student. If staff instructed the student to orally respond, raise their hand, write, and so on, the student engaged in the requested behavior within 5 s.

ISIS-SWIS Data

ISIS-SWIS was the online data management application used in the study. ISIS-SWIS is part of PBISApps, a series of web-based educational tools designed to support the implementation of school-wide positive behavior interventions and supports (SWPBIS). ISIS-SWIS enables users to (a) upload and store documents, (b) collect and save treatment fidelity data, (c) collect and save student behavior data, and (d) graph these data.

Self-monitoring treatment fidelity. EAs logged in to ISIS-SWIS and recorded a self-assessment of the extent to which they implemented their student's BSP for that day. A 0 to 5 rating scale was used (0 = *0%–10% of components implemented*, 1 = *11%–30% of components implemented*, 2 = *31%–50% of components implemented*, 3 = *51%–70% of components implemented*, 4 = *71%–90% of components implemented*, 5 = *91%–100% of components implemented*). EAs were provided with a fidelity checklist during ISIS-SWIS training that outlined components of their student's BSP and the above rating scale. EAs rated their performance and entered these data into ISIS-SWIS daily.

Student outcome data. EAs entered student behavior data into ISIS-SWIS. The researcher worked with the teacher to define an appropriate measure for student behavior in ISIS-SWIS. For each dyad, EAs used an already existing method for collecting data on student behavior (pencil and paper) and transferred those data into ISIS-SWIS. Already existing measures were used for EAs rather than the measures used by data collectors because they were deemed by the teachers to have better contextual fit and because they were aligned with student Individualized Education Program (IEP) goals.

For Trent, Evie entered the percentage of points he earned per day. A classroom-wide token economy system was in place where students had the opportunity to earn a specified number of points for each activity throughout the day for following the school rules (be safe, be respectful, follow directions, do your personal best). For each rule during each classroom activity, staff used a rating scale to award students their points on individual point sheets (0 points = *needs work*, 1 point = *okay*, 2 points = *excellent*). Because 2 points were possible for each of the four rules, there were a total of 8 points possible per activity. There were eight different activities outlined on the point sheet, resulting in 64 possible points per day. Each day following student dismissal, Evie logged in to ISIS-SWIS and entered the total number of points available that day (which was 64, unless the school day was shortened or the student was absent for part of the day) and the total number of points Trent earned that day. From these data, ISIS-SWIS derived a percentage of points earned.

For Marin, Chris entered the frequency of problem behavior per day. Prior to the study, classroom staff were already using a data sheet to track frequency of problem behavior. Each day following student dismissal, Chris logged in to ISIS-SWIS and entered the total number of occurrences of problem behavior (including all topographies) observed that day.

For Joey, Natalie entered the number of EA-directed tasks he completed per day. Joey's individualized schedule was a two-column table that listed six to eight "teacher choice" activities in the left column and six to eight corresponding "student choice" activities in the right column. "Teacher choice" activities included tasks such as completing a workbook page, participating in a mainstream class or activity (e.g., music class or library), completing a reading lesson, and so on. For each "teacher choice" activity Joey completed, he was given access to the corresponding "student choice" activity listed on his schedule. Each day following student dismissal, Natalie referred to Joey's daily schedule, counted the number of "teacher choice" tasks he completed, and entered that number into ISIS-SWIS.

ISIS-SWIS data entry use. To ensure EAs were entering data into ISIS-SWIS on a daily basis, the researcher kept a record of ISIS-SWIS data entry. Each day, the researcher logged in to ISIS-SWIS and viewed fidelity and student behavior data for each dyad. The researcher then documented on a data sheet if fidelity data and student behavior data were entered (yes or no). EAs and teachers were unaware that the researcher was collecting these data. With the exception of 1 day, all EAs entered data for both measures into ISIS-SWIS daily. On this day, the researcher emailed the EA to remind her that data should be entered every day. The EA entered data the following morning and continued entering data daily throughout the duration of the study.

Social Validity

After data collection was complete, teachers, EAs, and the district behavior specialist completed a questionnaire to determine the acceptability of the study procedures and outcomes. This 12-item questionnaire asked questions regarding the importance of collecting and reviewing data on treatment fidelity and student behavior, data-based decision

making, and the usability and usefulness of the intervention. Participants rated each item using a Likert-type scale (1 = *strongly agree* to 5 = *strongly disagree*). A copy of the questionnaire is available from the first author.

Design and Procedures

We used a multiple baseline design across student–EA dyads to determine whether a functional relation existed between the treatment package and observed treatment fidelity of student BSPs and student behavior. The design also allowed for descriptive assessment of the extent to which EA self-monitoring of fidelity (via ISIS-SWIS) aligned with observed treatment fidelity.

Review of student records. Prior to data collection, the researcher reviewed student records to ensure the student's plan included necessary elements that would be sufficient in producing behavior change. Each FBA was evaluated using the *FBA Procedural Adequacy Checklist* (Loman & Horner, 2013). If the FBA met criteria (5/5 checklist items present), BSPs were evaluated for technical adequacy using the *Critical Elements of the BSP* checklist (Strickland-Cohen & Horner, 2015) and for contextual fit using the *Assessment of Contextual Fit in Schools* rating scale (Horner, Salentine, & Albin, 2003). All student plans used in the study met criteria for technical adequacy and contextual fit, were developed with teacher and district behavior specialist involvement, and were informed by an FBA consisting of indirect and direct assessment procedures.

Baseline. During baseline, EAs continued implementing student BSPs under typical classroom conditions. The researcher asked teachers and EAs to continue throughout their day as they typically would. Typical classroom conditions consisted of the EA working with the student during regularly scheduled times/activities and being responsible for implementing the BSP and collecting data on student behavior during those times/activities. The researcher provided no feedback regarding plan implementation, student progress, data collection, and so on. All teachers indicated that EAs received instruction on implementing student BSPs and collecting data on student behavior prior to baseline (e.g., review of BSP during classroom meetings, modeling, in vivo prompts/error correction, feedback at the end of the school day). This "business as usual" baseline was used to emulate typical classroom conditions to the greatest extent possible, because in practice these would be the conditions under which the treatment package would be introduced in schools. The only change during baseline from what typically occurred in the classroom prior to data collection was that observers were present in the classroom to collect data on treatment fidelity, problem behavior, and academic engagement.

Treatment. The treatment package consisted of EAs (a) self-monitoring treatment fidelity, (b) collecting data on student behavior, (c) entering fidelity and student behavior data into ISIS-SWIS, and (d) reviewing those data weekly using graphs generated by ISIS-SWIS.

The researcher met with EAs and provided them with a fidelity checklist outlining components of their student's BSP and the corresponding fidelity rating scale used in ISIS-SWIS. The researcher then trained EAs to enter self-monitored fidelity and student behavior data in ISIS-SWIS. The training lasted approximately 45 min and consisted of verbal instruction, modeling, practice, and feedback. Training was considered complete when the EA was able to log in and accurately enter fictitious data for both measures without researcher support once during that training session. The day following training, the researcher visited the classroom to check-in, answer questions regarding ISIS-SWIS data entry, and remind the EA to enter fidelity and student behavior data daily.

Three to 4 days following this check-in and weekly thereafter, the researcher met with the teacher and EA to provide feedback regarding ISIS-SWIS use and review data in ISIS-SWIS. Feedback was not provided regarding implementation of BSP strategies. During these meetings, the researcher followed a procedural fidelity checklist to provide feedback for ISIS-SWIS use, including (a) praise for using features of ISIS-SWIS regularly and accurately, (b) identifying features of ISIS-SWIS they were not using, (c) modeling, practice, and feedback regarding how to use neglected features, and (d) discussing and agreeing upon features they will begin to use. After providing feedback on ISIS-SWIS use, the researcher prompted the teacher and EA to review data collected by the EA (self-monitoring of fidelity and student behavior). The researcher asked the teacher or EA to generate graphs in ISIS-SWIS, and then prompted them to discuss (a) whether the plan was being implemented with fidelity, (b) whether the plan was effective in minimizing problem behavior, and (c) whether any changes need to be made. After each prompt, the researcher paused at which point the teacher and EA referred to the graphs and began discussing the data. During discussion, the researcher provided affirming comments, answered any questions, and provided additional verbal prompts as necessary to facilitate the discussion. Each discussion resulted in a consensus as to whether the plan was being implemented and whether student problem behavior was decreasing. No changes were made to BSPs during the study.

Interobserver Agreement (IOA)

A second trained observer was present for 33% of the weekly meetings to collect data on procedural integrity. The researcher and second observer used a procedural fidelity checklist to indicate whether the researcher addressed each

meeting item outlined above. Integrity was calculated by the total number of items in agreement divided by the number of agreements plus disagreements multiplied by 100%. Procedural integrity was 100% for each of the two meetings.

A second trained observer collected IOA data on direct observation measures for a minimum of 20% of the sessions per phase per dyad. Cohen's kappa was used to calculate IOA. To obtain average kappa per dyad, kappa values for all dependent variables (for that dyad) were summed and divided by the total number of IOA sessions for that dyad. To obtain an average kappa per dependent variable, kappa values for each dependent variable (across all dyads) were summed and divided by the total number of IOA sessions. For Dyad 1 (Trent and Evie), average Cohen's kappa was 0.77 (k = 0.77), ranging from 0.67 to 0.95. For Dyad 2 (Marin and Chris), the average was 0.90 (k = 0.90), ranging from 0.78 to 0.98. For Dyad 3 (Joey and Natalie), the average was 0.86 (k = 0.86), ranging from 0.85 to 0.98. Average Cohen's kappa for treatment fidelity was 0.81 (k = 0.81), ranging from 0.67 to 0.98. Average kappa for student problem behavior was 0.91 (k = 0.91), ranging from 0.85 to 0.95. Average kappa for student academic engagement was 0.86 (k = 0.86), ranging from 0.77 to 0.90.

Results

Direct Observation Data

Figure 1 displays the percentage of BSP components implemented (primary y-axis) and the percentage of 10-s intervals with problem behavior and academic engagement (secondary y-axis) for all three dyads across baseline and treatment conditions. The dashed lines on the x-axis denote spring break. On Days 5 and 6, baseline data were not collected for Trent and Evie because small reading group did not occur on those days during the scheduled observation period from 9:00 a.m. to 9:20 a.m. Data were not collected on Day 20 for Marin and Chris because there was an assembly that occurred during the scheduled observation period from 9:30 a.m. to 9:50 a.m. Data were not collected for Joey and Natalie on Day 18 because Joey was absent from school.

Treatment fidelity. During baseline, the mean levels of treatment fidelity were 22%, 45%, and 29% of BSP components implemented for Evie, Chris, and Natalie, respectively. A decreasing trend was observed for all participants. Following the introduction of the treatment package, each of the dyads demonstrated an immediate and sustained increase in level, with mean fidelity of 93%, 91%, and 85% for Evie, Chris, and Natalie, respectively. There were no overlapping data points between baseline and treatment, and variability following treatment was low with a range of 100% to 71% of BSP components implemented across the three dyads.

Problem behavior. During baseline, the mean levels of problem behavior were 34%, 22%, and 12% of 10-s intervals for Trent, Marin, and Joey, respectively. An increasing trend was observed for Trent and Joey. Following the introduction of the treatment package, there was an immediate decrease in level for Trent, and a decreasing trend for Chris and Natalie. Levels for all three students remained low and data were stable, with the exception of Observation 15 for Trent. Per teacher report, Trent had an altercation with a peer before school on this day, and he engaged in protest behavior most of the morning following the altercation, despite Evie implementing the BSP with high fidelity. Mean levels of problem behavior during treatment were 4%, 7%, and 2% of 10-s intervals for Trent, Marin, and Joey, respectively. There was one overlapping data point between baseline and treatment for Trent, and several overlapping data points for Chris and Joey.

Academic engagement. During baseline, the mean levels of academic engagement were 45%, 48%, and 18% of 10-s intervals for Trent, Marin, and Joey, respectively. A sharp decreasing trend was observed for Trent, and data for Marin and Joey were variable. Following the introduction of the treatment package, there was an immediate increase in level for all three dyads. Mean levels of academic engagement were 78%, 96%, and 85% of 10-s intervals for Trent, Marin, and Joey, respectively. For Trent, academic engagement data were variable during treatment. For Marin, academic engagement remained high and stable. For Joey, an increasing trend was observed and continued throughout the duration of the study. There were several overlapping data points between baseline and treatment conditions for Trent, and no overlapping data points for Marin and Joey.

Self-Monitoring of Treatment Fidelity

Correlational analyses were conducted to determine the extent to which observed fidelity data were correlated with self-monitoring data. For Evie, observed treatment fidelity and EA self-monitoring of fidelity were strongly correlated, r = .68, p < .01. For Chris, there was a weak correlation of .28 (p = ns). For Natalie, observed treatment fidelity and EA self-monitoring of fidelity were strongly correlated, r = .89, p < .05.

Social Validity

EAs indicated the greatest agreement with items indicating that collecting data on student behavior is important (M = 1.00), that using ISIS-SWIS made it more likely they would collect data on student behavior (M = 1.66), that using ISIS-SWIS made it more likely they would reflect on how accurately they implemented the student's plan (M = 1.66), and that ISIS-SWIS was easy to use (M = 1.66). The average rating across EA responses to all items was 1.94.

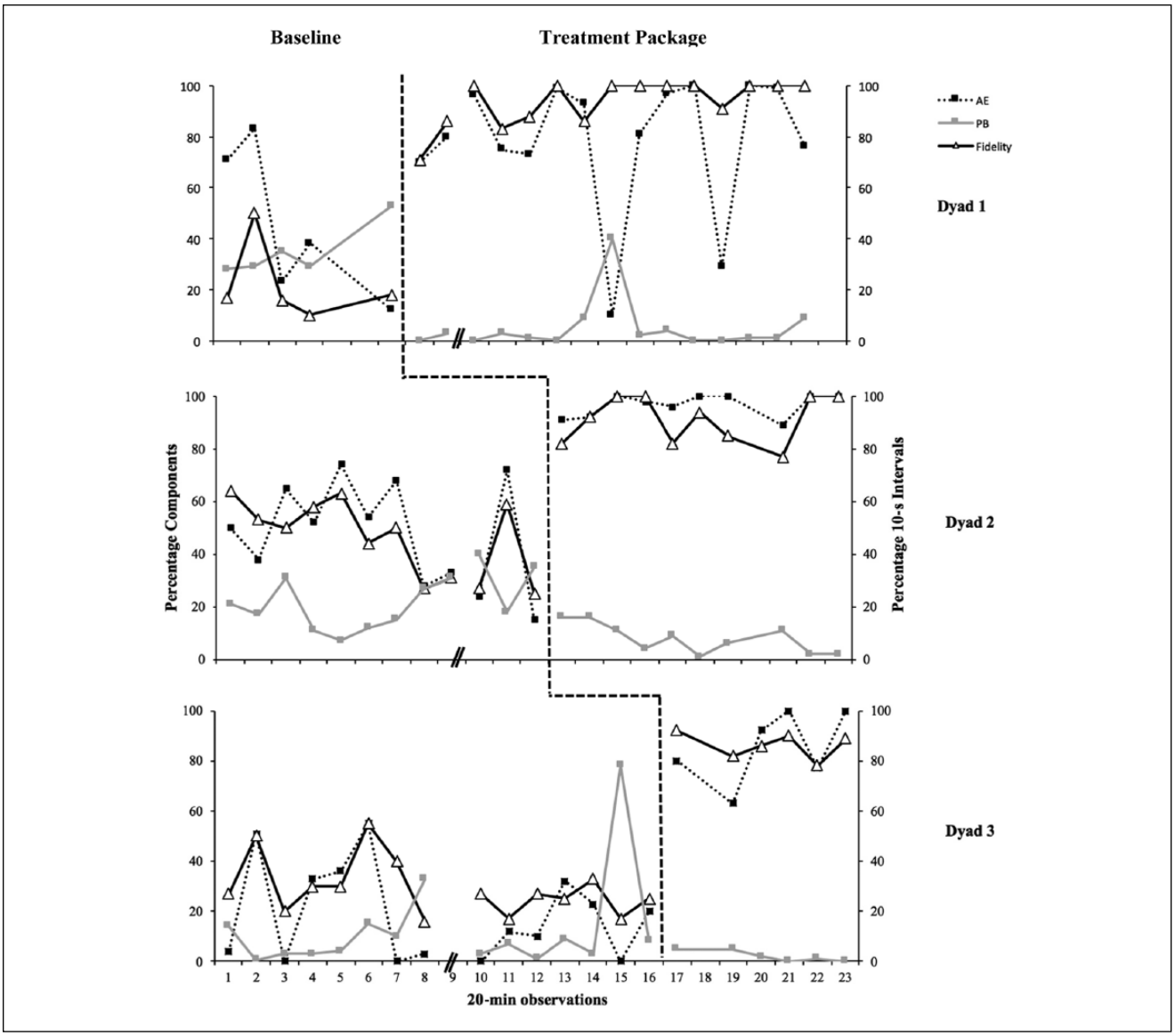

Figure 1. Percentage of BSP components implemented (treatment fidelity) and percentage 10 s intervals with PB and AE observed during 20-min observations.
Note. BSP = behavior support plan; PB = problem behavior; AE = academic engagement.

Discussion

The effective implementation of function-based interventions to teach socially appropriate behavior and decrease problem behavior is of utmost importance in schools. Unfortunately, staff are often not provided with adequate support that enables them to implement behavioral interventions with high fidelity, resulting in poor student outcomes. We used a multiple baseline design across three student–staff dyads to examine the effects of a treatment package consisting of (a) self-monitoring fidelity, (b) collecting data on problem behavior, (c) entering these data into ISIS-SWIS, and (d) reviewing these data weekly. During baseline, treatment fidelity for all dyads was low. Following the introduction of the treatment package, there was an immediate increase in level that remained high throughout the duration of the study and student behavior improved. Self-monitoring data for Evie and Natalie were strongly correlated with observed treatment fidelity. The social validity questionnaire indicated that overall, EAs found the treatment package to be beneficial and useful in assisting with the implementation of student BSPs.

Considerable efforts have been put forth over the past several decades to identify evidence-based practices (Fixsen, Fixsen, Blase, Metz, & Van Dyke, 2013), and this research has provided the field with invaluable information

regarding the efficacy of a variety of interventions. However, the effective use of these interventions in schools continues to be a challenge. This is of particular concern for students who engage in problem behavior given such behavior often results in restrictive educational placements and an array of social and academic difficulties. The disconnect between research and what actually occurs in classrooms has been referred to as the *research-to-practice gap*, and this phenomenon has spurred recent interest in developing a technology of implementation. Implementation includes clearly defining specific activities that are designed to assist in the effective use of an intervention (Fixsen et al., 2005). In the present study, implementation activities in the treatment package included (a) self-monitoring treatment fidelity, (b) collecting data on student behavior, (c) entering fidelity and student behavior data into ISIS-SWIS, and (d) reviewing these data weekly. These activities enabled staff to effectively implement student BSPs, therefore bridging the *research-to-practice gap* and improving student behavior. Results of the current study may be particularly important for populations in restrictive educational placements, such as the participants in this study, as a decrease in problem behavior could result in less restrictive placements (e.g., inclusion settings).

The collection of accurate, valid, and reliable data, and using those data for decision making, continues to be difficult for schools (Dunn et al., 2013; Newton et al., 2012; Schildkamp et al., 2012; Telzrow et al., 2000). This challenge may in part be due to the lack of efficient tools to assist schools in collecting, organizing, and summarizing data. A *decision support data system* has been identified as important in promoting effective implementation (Fixsen, Blase, Naoom, & Duda, 2013). An important caveat to collecting and using data is the specific measures that are used. Data must be collected on independent and dependent variables (Gresham et al., 1993) to make informed decisions that result in improved staff and student behavior. In the present study, ISIS-SWIS served as a *decision support data system*. EAs entered data on treatment fidelity and student behavior and reviewed these data with the team. ISIS-SWIS provided the structure and organization to support public school paraprofessionals in carrying out required implementation activities.

In this study, staff self-monitored treatment fidelity. It is important to tease apart the two potential uses of self-monitoring treatment fidelity: as a measure of fidelity or as a strategy to improve fidelity. Although self-monitoring to measure fidelity is easier and less intrusive than a supervisor collecting direct observation data, it is important to note that self-monitoring data may not be a valid measure of treatment fidelity, as individuals are not always accurate self-reporters of their behavior. However, self-monitoring can be used to improve fidelity. In schools, self-monitoring could offer a cost-effective alternative to resource-extensive performance feedback procedures. These two uses of self-monitoring fidelity (as a measure of fidelity or to improve fidelity) are distinct and not necessarily related, as previous research has indicated that the accuracy with which an individual self-monitors his or her behavior is not an important factor in producing a change in his or her behavior (see Broden, Hall, & Mitts, 1971; Fixsen, Phillips, & Wolf, 1972; Hayes & Cavior, 1977; Herbert & Baer, 1972; Lipinski & Nelson, 1974; Nelson & Hayes, 1981). These findings suggest that the extent to which an individual accurately self-monitors fidelity has no bearing on the effectiveness of self-monitoring in improving fidelity. In the current study, this was addressed descriptively rather than experimentally. Self-monitoring data were strongly correlated with observed fidelity data for Evie and Natalie, and improvement in EA behavior was observed for all three dyads. Although self-monitoring may used as a strategy to improve treatment fidelity, researchers and practitioners should exercise extreme caution when using self-monitoring data as a sole measure of treatment fidelity.

Future Research

Results of the present study present several important areas for additional research. First is the implementation of the treatment package in other settings (e.g., inclusion and general education classrooms), with larger units of analysis (i.e., school buildings, school districts, states), and with school teams that do not include the researcher. In addition, although a *decision support data system* is a factor that enables an organization's success (Fixsen, Blase, Naoom, & Duda, 2013), additional research is needed to guide the field on how to implement such data systems in a variety of school contexts. Wayman (2005) wrote how schools can be *data rich*, yet *information poor*. Regardless if schools are collecting data, they need to effectively use the data for decision making.

Results of the current study corroborate previous research indicating the potential benefits of self-monitoring to improve instructional effectiveness in an easy and cost-efficient manner. Further investigation is needed to define the features of self-monitoring that make it an effective strategy. For example, when an individual uses a self-monitoring checklist, the checklist could serve as a discriminative stimulus or prompt when the checklist is reviewed prior to implementation, or as a reinforcer or corrective feedback when reviewed following implementation. The present study did not include measures to assess when or how frequently EAs reviewed the checklist. Future research should explicitly examine the effects of reviewing the checklist before and after implementation, how often checklists are reviewed, and specific components of checklists that increase the likelihood of their effectiveness (e.g., number of items on the checklist, scoring of items, etc.).

Future research should also examine the effects of varying methods to measure treatment fidelity and use fidelity data for decision making. Questions regarding the best metric to measure fidelity in schools (i.e., rating scale, frequency count, etc.), who should measure fidelity, and the frequency with which fidelity should be measured and reviewed (i.e., daily, weekly, monthly) have yet to be answered. Considerations should be made with regard to accuracy, validity, and reliability of the data; response effort required by the data collector; and the degree to which the data can be easily summarized for decision making. The research is clear that data on both treatment fidelity and student behavior are crucial, yet research on how school teams should collect, summarize, and analyze these data is needed.

Limitations

There are several limitations of this study worth noting. First, given the intervention was a treatment package, there is no way to determine which individual component(s) of the package produced the change in EA behavior. To make this determination, a component analysis should be conducted to determine the individual effects of each component of the treatment package. Second, the topographies of problem behavior in the present study were mild to moderate in severity. EAs would likely require (and for ethical and safety reasons should receive) additional training and support in the form of expert-delivered performance feedback and coaching to implement strategies for more serious topographies of problem behavior such as self-injury and aggression. Third, all EAs in the present study had several years of experience working with students with special needs and previously received district-wide training in behavior support. Although baseline data indicated low levels of treatment fidelity, it is possible that this previous training affected their ability to implement BSPs with minimal coaching. In the current study, there was no content knowledge assessment to determine the extent to which EAs understood the basic principles of behavior or behavioral interventions. As such, it is possible that previous training provided EAs with necessary/foundational content knowledge, and the addition of the treatment package was sufficient in improving treatment fidelity. Fourth, to self-monitor treatment fidelity in the current study, the researcher gave EAs a fidelity checklist outlining components of the student's BSP and asked them to indicate the percentage of components they implemented. It is possible that similar results would not be observed had the EAs not received such a detailed fidelity checklist. Finally, student BSPs in the current study met criteria for technical adequacy and contextual fit and were informed by an FBA. If the BSPs did not meet such criteria, the treatment package would not produce an improvement in student behavior. The treatment package assisted in the implementation of student BSPs, but an inadequate BSP applied to the treatment package would produce poor results.

Conclusion

The current study provides evidence that a treatment package consisting of (a) self-monitoring treatment fidelity, (b) collecting data on student behavior, (c) entering fidelity and student data into an online data management application, and (d) reviewing these data on a weekly basis can improve the fidelity with which staff implement function-based interventions in typical school contexts. ISIS-SWIS served as a *decision support data system* (Fixsen, Blase, Naoom, & Duda, 2013) that provided structure for the implementation of activities in the treatment package. EAs and teachers indicated that ISIS-SWIS was a beneficial and easy-to-use tool that increased the likelihood they would collect data and implement behavioral interventions with fidelity. Results of this study have implications for policy, research, and practice regarding the efficiency with which schools can coordinate and monitor interventions for students.

Declaration of Conflicting Interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Development of this paper was supported by the Office of Special Education Programs U.S. Department of Education (H326S130004). Opinions expressed herein are those of the authors and do not necessarily reflect the position of the U.S. Department of Education, and such endorsements should not be inferred.

References

Begeny, J. C., & Martens, B. K. (2006). Assessing pre-service teachers' training in empirically-validated behavioral instruction practices. *School Psychology Quarterly*, *21*, 262–285.

Belfiore, P. J., Fritts, K. M., & Herman, B. C. (2008). The role of procedural integrity: Using self-monitoring to enhance discrete trial instruction. *Focus on Autism and Other Developmental Disabilities*, *25*, 95–102.

Broden, M., Hall, R. V., & Mitts, B. (1971). The effect of self-recording on the classroom behavior of two eighth-grade students. *Journal of Applied Behavior Analysis*, *4*, 191–199.

Detrich, R. (2014). Treatment integrity: Fundamental to education reform. *Journal of Cognitive Education and Psychology*, *13*, 258–271. doi:10.1891/1945-8959.13.2.258

Dunlap, G., & Carr, E. (2007). Positive behavior support and developmental disabilities. In S. Odom, R. Horner, M. Snell, & J. Blacher (Eds.), *Handbook of developmental disabilities* (pp. 469–482). New York, NY: Guilford Press.

Dunlap, G., Conroy, M., Kern, L., DuPaul, G., VanBrakle, J., Strain, P., . . . Ostrosky, M. (2003). *Research synthesis on effective intervention procedures: Executive summary*.

Tampa: Center for Evidence-Based Practice: Young Children With Challenging Behavior, University of South Florida.

Dunlap, G., & Fox, L. (2012). Function-based interventions for children with challenging behavior. *Journal of Early Intervention*, *33*, 333–343.

Dunn, K. E., Airola, D. T., Lo, W., & Garrison, M. (2013). Becoming data driven: The influence of teachers' sense of efficacy on concerns related to data-driven decision making. *The Journal of Experimental Education*, *81*, 222–241.

Fixsen, D. L., Blase, K. A., Metz, A., & Van Dyke, M. (2013). Statewide implementation of evidence-based programs. *Exceptional Children*, *79*, 213–230.

Fixsen, D. L, Blase, K., Naoom, S., & Duda, M. (2013). *Implementation drivers: Assessing best practices*. Chapel Hill: University of North Carolina at Chapel Hill.

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, FL: Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network, University of South Florida.

Fixsen, D. L., Phillips, E. L., & Wolf, M. M. (1972). Achievement place: The reliability of self-reporting and peer-reporting and their effects on behavior. *Journal of Applied Behavior Analysis*, *5*, 19–30.

Gresham, F. M., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis*, *26*, 257–263.

Hager, K. D. (2012). Self-monitoring as a strategy to increase student teachers' use of effective teaching practices. *Rural Special Education Quarterly*, *31*(4), 9–17.

Hayes, S. C., & Cavior, N. (1977). Multiple tracking and the reactivity of self-monitoring: I. Negative behaviors. *Behavior Therapy*, *8*, 819–831.

Herbert, E. W., & Baer, D. M. (1972). Training parents as behavior modifiers. *Journal of Applied Behavior Analysis*, *5*, 139–149.

Horner, R. H., Salentine, S., & Albin, R. W. (2003). *Self-Assessment of contextual fit in schools*. Eugene: University of Oregon.

Ingram, K., Lewis-Palmer, T., & Sugai, G. (2005). Function-based intervention planning: Comparing the effectiveness of FBA function-based and non-function-based intervention plans. *Journal of Positive Behavior Interventions*, *7*, 224–236.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, *27*, 197–209.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Kalis, T. M., Vannest, K. J., & Parker, R. (2007). Praise counts: Using self-monitoring to increase effective teaching practices. *Preventing School Failure*, *51*, 20–27.

Keller, C. L., Brady, M. P., & Taylor, R. L. (2005). Using self evaluation to improve student teacher interns' use of specific praise. *Education and Training in Developmental Disabilities*, *40*, 368–376.

Lipinski, D. P., & Nelson, R. O. (1974). The reactivity and unreliability of self-recording. *Journal of Consulting and Clinical Psychology*, *42*, 118–123.

Loman, S., & Horner, R. H. (2013). Examining the efficacy of a basic functional behavioral assessment training package for school personnel. *Journal of Positive Behavior Interventions*, *16*, 18–30.

Lylo, B. J., & Lee, D. L. (2013). Effects of delayed audio-based self-monitoring on teacher completion of learning trials. *Journal of Behavioral Education*, *22*, 120–138.

Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, *42*, 30–37.

Marquis, J. G., Horner, R. H., Carr, E. G., Turnbull, A. P., Thompson, M., Behrens, G. A., . . . Doolabh, A. (2000). A meta-analysis of positive behavior support. In R. M. Gerston & E. P. Schiller (Eds.), *Contemporary special education research: Syntheses of the knowledge base on critical instructional issues* (pp. 137–178). Mahwah, NJ: Lawrence Erlbaum.

Nelson, R. O., & Hayes, S. C. (1981). Theoretical explanations for reactivity in self-monitoring. *Behavior Modification*, *5*, 3–14.

Newcomer, L. L., & Lewis, T. J. (2004). Functional behavioral assessment: An investigation of assessment reliability and effectiveness of function-based interventions. *Journal of Emotional and Behavioral Disorders*, *12*, 168–181.

Newton, J. S., Horner, R. H., Algozzine, B., Todd, A. W., & Algozzine, K. (2012). A randomized wait-list controlled analysis of the implementation integrity of team-initiated problem solving processes. *Journal of School Psychology*, *50*, 421–441.

Plavnick, J. B., Ferreri, S. J., & Maupin, A. N. (2010). The effects of self-monitoring on the procedural integrity of a behavioral intervention for young children with developmental disabilities. *Journal of Applied Behavior Analysis*, *43*, 315–320.

Schildkamp, K., Ehren, M., & Lai, M. K. (2012). Editorial article for the special issue on data-based decision making around the world: From policy to practice to results. *School Effectiveness and School Improvement*, *23*, 123–131.

Seligson-Petscher, E., & Bailey, J. S. (2006). Effects of training, prompting, and self-monitoring on staff behavior in a classroom for students with disabilities. *Journal of Applied Behavior Analysis*, *39*, 215–226.

Simonsen, B., MacSuga, A. S., Fallon, L. M., & Sugai, G. (2013). The effects of self-monitoring on teachers' use of specific praise. *Journal of Positive Behavior Interventions*, *15*, 5–15.

Strickland-Cohen, M. K., & Horner, R. H. (2015). Typical school personnel developing and implementing basic behavior support plans. *Journal of Positive Behavior Interventions*, *17*, 83–94.

Sutherland, K. S., & Wehby, J. H. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with emotional and behavioral disorders. *The Journal of Special Education*, *35*, 161–171.

Telzrow, C. F., McNamara, K., & Hollinger, C. L. (2000). Fidelity of problem-solving implementation and relationship to student performance. *School Psychology Review*, *29*, 443–461.

Wayman, J. C. (2005). Involving teachers in data-driven decision making: Using computer data systems to support teacher inquiry and reflection. *Journal of Education for Students Placed at Risk*, *10*, 298–308.

Workman, E. A., Watson, P. J., & Helton, G. B. (1982). Teachers' self-monitoring of praise vs praise instructions: Effects on teachers' and students' behavior. *Psychological Reports*, *50*, 559–565.