Journal of Behavioral Education
https://doi.org/10.1007/s10864-018-9310-2

ORIGINAL PAPER

Effects of Teacher-Implemented Coaching to Increase the Accuracy of Data Collected by Paraeducators

Rose A. Mason[1] · Alana. G. Schnitz[2] · Stephanie Gerow[3] · Zhe G. An[4] · Howard P. Wills[2]

Abstract

The purpose of the present study was to assess the impact of coaching with performance feedback from teachers on accuracy of paraeducators' momentary time sampling (MTS) data of students' on-task behavior. Two lead teachers and three paraeducators participated in the study. The relation between coaching and accuracy of the data collection was evaluated using a multiple-baseline across paraeducators design. Baseline data from this study suggest that some paraeducators need explicit instruction in how to collect data with fidelity. Once coaching with performance feedback from teachers was implemented, there was an immediate increase in accuracy of data collection by paraeducators, as measured by inter-rater agreement, indicating a functional relationship between the independent and dependent variables. These results highlight that teacher-led coaching is feasible and effective for increasing paraeducators' MTS data collection accuracy and provides preliminary evidence that the accuracy maintained 1–2 months following cessation of formal coaching sessions. Implications for practice and future research are discussed.

Keywords Data collection · Paraeducators · Special education · Professional development

✉ Rose A. Mason
rmason3@purdue.edu

[1] Department of Educational Studies, College of Education, Purdue University, BRNG 5108B, 100 University St, West Lafayette, IN 47907, USA

[2] Juniper Gardens Children's Project, University of Kansas, Kansas City, KS, USA

[3] Department of Educational Psychology, School of Education, Baylor University, Waco, TX, USA

[4] School of Education, University of Wisconsin, Madison, WI, USA

Published online: 17 October 2018

Introduction

Accurate information about student performance is integral to the provision of effective instruction for children with disabilities. The law requires school professionals to develop Individualized Education Programs that include measurable goals and ongoing data collection of students' progress toward those goals (Individuals with Disability Education Improvement Act [IDEIA] 2004) 34 C.F.R. § 300.347(a)(7)). To adequately monitor students' progress as required by IDEIA, professionals must collect accurate data. Data can be used to validate initial assessment information (Sandall et al. 2004), monitor children's progress (Gunter et al. 2003a; Sandall et al. 2004; Stecker et al. 2008), communicate and collaborate with families and other professionals [Council for Exceptional Children (CEC) 2004; IDEIA 2004], and evaluate effectiveness of instruction and programming (Babkie and Provost 2004; Boardman et al. 2005). Teachers can use students' academic and behavioral data to identify effective and ineffective instructional practices and make educational decisions based on the data collected (Lewis-Palmer et al. 1999). Thus, accurate and reliable data collection by educators, including paraeducators, is essential for guiding decision-making and ensuring positive educational outcomes for students with disabilities, including those with autism (Gunter et al. 2003a).

Autism and On-Task Behavior

Monitoring on-task behavior, for instance, can provide the teacher with necessary information regarding the time a student spends attending to tasks, complying, and engaging in appropriate learning behaviors. As a "response class," on-task behavior is broadly defined as attending to the current activity (e.g., listening to teacher, completing independent work activity, working in groups) as evidenced by orienting toward the given task without extraneous movement (e.g., manipulating irrelevant objects, moving feet, rocking) or noise (Clare et al. 2000; Stahr et al. 2006) that distracts from the ability to actively engage. On-task behavior is a component of engagement, perhaps the most observable, and has been identified as a necessary behavior for learning and academic progress (Kartal and Oxkan 2015; Rabiner and Coie 2000).

Individuals with autism, however, often struggle with on-task behavior due to deficits in executive functioning (Hume et al. 2009), which impacts the ability to modulate behaviors, focus on relevant stimuli, and sustain attention (Finn et al. 2015; Southall and Gast 2011). This inattention, or off-task behavior, impacts the ability for the students to obtain full benefits from instruction. Additionally, the off-task behavior and inattention often interferes with educational professionals' ability to provide instruction given the additional time required to manage and redirect inattention and disruptive behaviors (Kartal and Ozkan 2015).

Effective instructional strategies and behavioral interventions can improve the on-task behavior of students with autism. For instance, high rates of teacher praise (Simonsen et al. 2013) and high levels of opportunity to respond (Sutherland and Wehby 2001) are instructional practices known for effectively increasing on-task

behavior for students including those with autism (Rivera et al. 2015). Additionally, if data indicate the on-task behavior of students with autism is unresponsive to effective instructional practices, individualized behavior interventions such as differential reinforcement (Jessel et al. 2017), physical activity (Luke et al. 2014), video modeling (Schatz et al. 2016), and self-monitoring (Finn et al. 2015) can be implemented. Given the importance of on-task behavior coupled with its responsiveness to intervention, tracking students' on-task behavior is important both for guiding changes in instructional behavior and development of individualized behavior intervention plans.

Training Paraeducators

The Council for Exceptional Children (CEC) has designated data collection as a basic skill needed for all paraeducators (2004). Paraeducators are frequently charged with collecting data on students' performance for both academic and behavior goals (Carter et al. 2009; Giangreco et al. 2002; Griffin-Shirley and Matlock 2004; Keller et al. 2007). In fact, paraeducators report that data collection is one of their most frequently assigned tasks (Carter et al. 2009). Additionally, assigning data collection to paraeducators allows the teacher to focus more on the primary responsibility of delivering instruction while still having the data needed to appropriately monitor the effectiveness of instructional practices and behavioral interventions. Yet, paraeducators receive little to no training in data collection procedures and report they are underprepared for this task (Carter et al. 2009), thus increasing the likelihood that the data they collect may be inaccurate. Training paraeducators in data collection procedures is imperative for obtaining accurate and reliable data that can be used to guide educational decision-making.

Momentary time sampling (MTS) is one such data collection procedure that would be particularly useful for paraeducators to learn. MTS can be advantageous over other data collection procedures such as continuous recording, as it only requires the recorder to attend to the presence or absence of the target behavior at the end of the interval rather than throughout the recording period (Alberto and Troutman 2003). Further, MTS is more feasible in the classroom as it allows for simultaneously collecting data while engaging in instruction (Alberto and Troutman 2003; Gunter et al. 2003a).

There are a number of models for paraeducator training used in research and practice. The most common is large-group professional development format, consisting primarily of written and verbal instructions, though research indicates this approach rarely produces sustainable improvement in performance (Schepis et al. 2003; Tate et al. 2005). This type of professional development is typically ineffective due to the lack of assessment (Collinson 2000; Lewis et al. 2004; Rockwell 2008; Stichter et al. 2006), individualized resources and materials (Brabec et al. 2004; Smith et al. 2007), and individualized feedback (Logan and Stein 2001; Shernoff and Kratochwill 2007), as well as a lack of emphasis on implementation and sustainability in specific contexts. However, professional development that includes direct instruction, modeling, opportunity for practice, and ongoing performance feedback can result in long-term improvement

in performance for both teachers (Snyder et al. 2015) and paraeducators (Rispoli et al. 2011). Although research indicates these components are effective, it is often challenging to provide all these components of professional development for paraeducators due to time and personnel resource constraints.

One option for provision of high-quality professional development for paraeducators is coaching with performance feedback. As an alternative to group training alone, coaching has strong evidence of improving teacher's implementation of high-quality practices (Snyder et al. 2015). However, only a handful of studies have implemented a coaching model to train paraeducators, typically utilizing a researcher-as-coach model rather than a typical school agent (Brock and Carter 2013). A few recent studies have trained teachers to implement coaching and performance feedback with paraeducators, which represents a more feasible and sustainable method of providing effective training. Brock and Carter (2015) implemented a teacher-delivered instruction package with a one-time teacher-implemented coaching and performance feedback session to increase paraeducators' accurate implementation of a peer support intervention for students with severe disabilities. Similarly, in another study, teachers were trained to coach their paraeducators to implement discrete trial training with fidelity for students with autism and moderate-to-severe developmental disabilities (Mason et al. 2017). Both studies provide promise for a teacher-as-coach model to increase the fidelity of paraeducators' implementation of evidence-based practices with children with disabilities.

There are gaps in the literature related to performance feedback and coaching of paraeducators. There is a need to evaluate the teacher-as-coach model with other skills, including data collection, to evaluate the effectiveness of this training model across the critical paraeducator competencies. Additionally, previous research on the teacher-as-coach model has not evaluated whether paraeducators continued to implement the instructional practices accurately once the coaching was discontinued. The purpose of the present study was to assess the impact of teacher-implemented coaching with performance feedback on the accuracy of paraeducators' data collection, as well as to evaluate whether or not the newly acquired skill would maintain following cessation of the coaching and performance feedback. The following research questions were addressed:

(a) Does teacher-delivered coaching with performance feedback increase the accuracy of momentary time sampling data of students' on-task behavior collected by paraeducators?
(b) Will improvements in paraeducators' momentary time sampling data collection maintain following coaching?

Method

Participants, Setting, and Materials

This study was conducted as a part of a research–practitioner partnership between a research university and a large urban school district in the Midwest, aimed at

identifying effective and efficient training methods for paraeducators. The study was conducted in two of the district's Title I elementary schools serving students in grades K-5, with an average enrollment of 438 students. Across the two schools, the average percentage of economically disadvantaged and minority student enrollment was 74% and 84%, respectively. Additionally, the average percent of students with disabilities across the two schools was 12%.

Participant Recruitment and Inclusion Criteria

After receiving institutional review board and district approval, teachers and paraeducators were recruited for the study. Teachers were recruited through the special education district coordinator. A research team member met with teachers who were interested in participating. Teacher participants were the first two teachers who consented to participate. To be included in the study, teachers had to be assigned to work with students with ASD and supervise at least one paraeducator. The paraeducators that worked with each participating teacher were given the opportunity to participate in the study. To be included in the study, paraeducators had to be assigned to work with students with ASD and be supervised by a teacher who had agreed to participate in the study. All three paraeducators who consented had previously been trained in several data collection methods including momentary time sampling (MTS), partial interval, and whole interval recording during a district-wide training approximately 2 months before the onset of the study. This training included instruction on each of the methods, demonstrations of how to do each method, and opportunity to practice each method. For practice, the instructor displayed videos of students in classrooms and then instructed the participating paraeducators to collect data on a defined target behavior. Data collection was done independently but simultaneously by each participating paraeducator and conducted for each of the methods being discussed. Following completion of the practice results were reviewed and questions answered in a whole group discussion format. This training was a large-group in-servicer training that did not include practice with follow-up and feedback.

Dyad 1: Jasmine (Paraeducator) and Bianca (Teacher)

Jasmine, who had a bachelor's degree, was a 52-year-old African-American female. She had been a paraeducator for 7 years in a self-contained classroom for third–fifth-grade students with autism. Jasmine's teacher and coach for this study, Bianca, was a 26-year-old Caucasian female who worked as a special education teacher for 2 years. Bianca had a bachelor's degree and was in the process of obtaining her master's degree. There were a total of 8 students in Bianca and Jasmine's classroom. Bianca reported she used data collection in her classroom to measure progress on Individualized Education Plan goals for her students, and this task was not assigned to Jasmine as her data were not reliable. Bianca dropped out of the study after the first coaching session. The district lead teacher for specialized programs, *Eleanor*, who had been previously trained in the coaching protocol, served as the coach for Jasmine for her last two coaching sessions. Eleanor had a master's degree in special education and had previous experience supervising paraeducators.

Dyad 2: Olive (Paraeducator) and Nancy (Teacher)

Olive was a 60-year-old African-American female and had been working as a paraeducator for 15 years. She indicated that she had limited experience with data collection noting only one recent occasion when she was assigned the task of collecting frequency data for a student's spitting behavior. Olive was supervised by Nancy, a 29-year-old Caucasian female, who had worked as a special education teacher for 3 years. Nancy and Olive worked together in a self-contained classroom with six, first–second-grade students with autism. At the time of study, Nancy reported she had experience collecting data on students' behavior and academic progress. She also noted that she did not assign data collection to her paraeducators as they did not collect accurate data, yet indicated it would be helpful if she could add this to their assigned tasks to track challenging behavior and progress toward meeting individualized educational plan goals.

Dyad 3: Amy (Paraeducator) and Nancy (Teacher)

Amy, a 24-year-old African-American female, had 6 months experience as a paraeducator. Amy had a bachelor's degree and was pursuing a master's degree in counseling. Amy reported no experience collecting data in the classroom. Both Olive and Amy were assigned to the same classroom, with Nancy as the teacher for the classroom.

Setting and Materials

The current study was conducted in two self-contained classrooms for students with autism in two of the district's elementary schools. As noted 6–8 students were enrolled in each classroom. Sessions were conducted during one-to-one instruction during which the paraeducator was providing instruction to the student for whom she was collecting data. The target student had to be engaged in an activity for which on-task behavior was relevant, that is completing a task as assigned by the paraeducator and not engaged in a preferred activity (e.g., play) or downtime. The activities of the students varied as the students were engaged in activities relevant to his or her individualized education plan (e.g., completing tasks related to one-to-one correspondence, letter identification, counting, letter writing practice). Although this did limit session control and increase the complexity of data collection, the sessions were conducted in this manner as the lead teachers indicated this would be the expectation in typical practice. The materials used in the study were instructional materials that were already present and typically used in the classroom.

Paraeducators were given a data sheet by the research team for use during the study. The data sheet included a table with a blank box in which the paraeducator could record data for each 10-s intervals within a 5-min session. Researchers and paraeducators downloaded an interval timer application on their smartphones to use during data collection sessions. The interval timer provided prompts for participants to record the student's target behavior every 10 s. Participants collected data using momentary time sampling (MTS). Momentary time sampling was chosen,

in collaboration with the participating teachers, because it does not require the observer to watch the target student the whole time, but rather to make a judgment at the presence or absence of the target behavior at the end of the 10-s interval. As previously noted, MTS is easier for use in the classroom to measure behavior than the other time sampling systems (Alberto and Troutman 2003; Gunter et al. 2003b). In addition, MTS is a more accurate measurement of behavior than partial or whole intervention when used to estimate duration (Ledford et al. 2015). MTS was chosen to estimate the percentage of intervals the children engaged in the target behavior in the classroom.

Experimental Design

A multiple-probe design across paraeducators was used to evaluate the functional relation between implementation of coaching with performance feedback delivered by the teacher and improvements in the accuracy of paraeducator's MTS data, as measured by percentage of agreement with a secondary observer, of students' on task behavior. The intervention was systematically introduced at different points in time across the three participants, following a minimum of five data points and establishment of a clear pattern of behavior, meeting the design standards for rigorous single-case research (Kratochwill et al. 2013). Visual analysis was used to assess the functional relation between the independent and dependent variable.

Measurement

Dependent Variable

The dependent variable was measured two times per week during the baseline and intervention phases of the study. Data were collected on the accuracy of paraeducators' data collection throughout the study, using MTS with 10-s intervals. Accuracy of paraeducators' data collection was calculated as the percentage of intervals with agreement with another observer on occurrence of child on-task behavior. An agreement was scored if both the paraeducator and the research team member rated the child as on-task during the interval or if neither the paraeducator nor research rated the child as on-task. An individualized, operational definition, developed collaboratively by the supervising teacher and the first author was used to measure on-task behavior for each child. On-task behavior was chosen given its importance to learning and instructional decisions (Kartal and Oxkan 2015; Rabiner and Coie 2000). On-task behavior was broadly defined, for the purposes of this study, as the student engaged in the task assigned which included looking at instructional materials and/or looking at the instructor (teacher or paraeducator) and exhibiting behavior required to complete the current task. For example, if the student was working on a task of letter writing, on-task behavior required the student to be looking at the writing paper and holding the pencil while actively manipulating the pencil in a manner required for letter completion. On the other hand, if the paraeducator was providing instruction, the student was counted as on-task if the student was quiet and oriented,

facing, toward the paraeducator. Conversely, if the student was engaged in behaviors incompatible with completing the assigned task (e.g., leaving the work area) or listening to the paraeducator (e.g., vocalizing, looking around the room with body oriented away from the paraeducator).

Coaching Procedural Fidelity

Teachers' recorded all coaching sessions and uploaded the resulting videos to a password-protected cloud storage site. Teachers' coaching fidelity was measured utilizing a researcher-developed measure of coaching fidelity. The measure was based on the nine steps for implementation of the coaching procedure as outlined in Fig. 1. The nine items were rated yes/no/not applicable based on the presence or absence of each of the steps. The second author, who has previous experience coaching teachers to implement teaching practices, assessed the coaching fidelity for 33% each of the coaching sessions for Olive and Amy and 50% of the coaching sessions for Jasmine. The average coaching fidelity for Olive with Nancy and Amy was 100% and 94%, respectively. The average coaching fidelity for Bianca with Jasmine was 88%. The overall fidelity of coaching across dyads was 94% ($R = 88–100\%$).

Inter-observer Agreement

Data collectors, graduate research assistants, were trained in MTS by the first two authors and practiced collecting data until they reached 90% reliability for three consecutive sessions. The data from one of the data collectors served to assess the accuracy of paraeducator data collection (the dependent variable). A second, independent observer collected inter-observer agreement (IOA) data on a minimum of 20% of data points in each phase for each participant. The point-by-point method of agreement was used to calculate occurrence, nonoccurrence and overall IOA ((agreement/agreements + disagreements) * 100). Percentage of IOA for each IOA session within and across participants is graphically displayed in Fig. 2 (indicated by open triangles). Table 1 indicates the average percentage of occurrence, nonoccurrence, and overall agreement for each phase across all paraeducators. The combined average percentage of agreement for occurrence and nonoccurrence was 89% and

1. Open with a positive greeting
2. Review current action plan goal
3. Share most recent data applicable to action plan goal (i.e 85% IOA for 3 sessions)
4. Provide supportive feedback regarding progress towards goal
5. Provide suggestions for strategies to meet goal
6. Ask the paraeducator if she has any questions
7. Respond to any questions
8. Only provide instruction and coaching relevant to current goal
9. Schedule the day/time for next coaching session

Fig. 1 Steps for implementation of coaching procedure

90%, respectively. The overall average reliability across participants and phases was 93% ($R = 80–100\%$).

Inter-rater agreement was also measured on fidelity of coaching. A second, independent observer trained in the coaching procedure and fidelity measure viewed and evaluated the fidelity of coaching for 100% of the videos for which coaching fidelity was assessed as described above. The inter-rater agreement for fidelity of coaching was 99% ($R = 89–100\%$).

Social Validity

Social validity of the coaching intervention was measured using two researcher-developed questionnaires. The *Teacher Social Validity* measure (see Table 2 for included items) was administered to determine the teachers' satisfaction with the training they received on coaching, satisfaction with the coaching procedures and dosage, and the effectiveness of coaching procedures to change the paraeducators

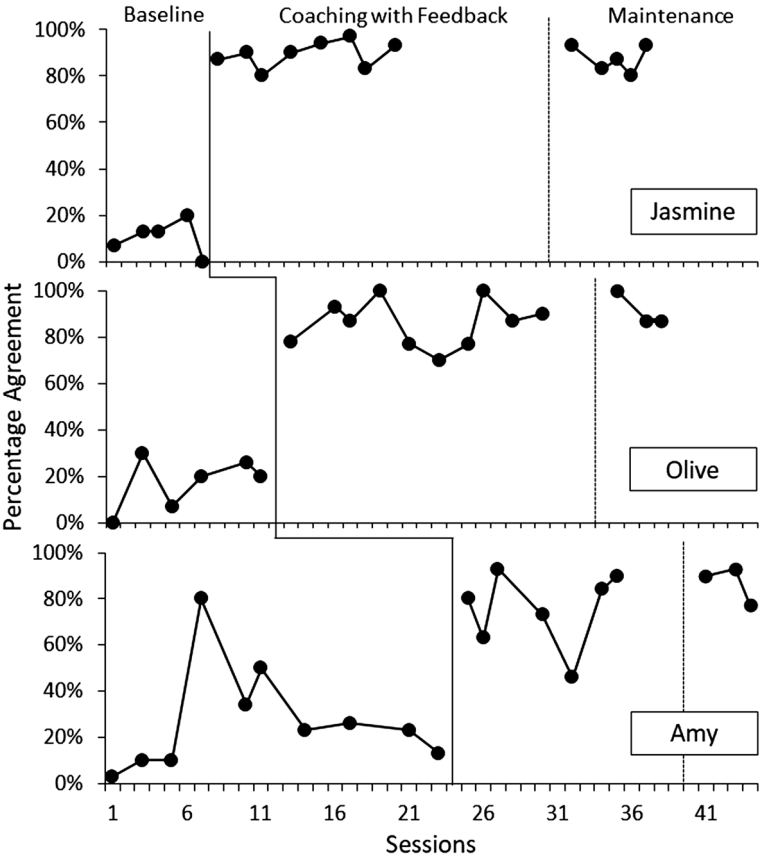

Fig. 2 Percentage agreement between the paraeducator and the primary data collector (researcher)

Table 1 Average percent of researcher agreement by phase and paraeducator

Phase	Overall (%)	Occurrence (%)	Nonoccurrence (%)
Jasmine			
Baseline	96	100	94
Intervention	90	88	85
Maintenance	92	81	93
Olive			
Baseline	88	87	86
Intervention	95	95	81
Maintenance	87	81	93
Amy			
Baseline	93	77	98
Intervention	97	96	88
Maintenance	95	100	94

behavior. The *Paraeducator Social Validity* measure (see Table 3 for included items) contained questions on the appropriateness of dosage and format of coaching, usefulness of feedback from the teacher, importance of collecting data in the classroom, and usefulness of MTS as a data collection method. Each questionnaire consisted of 11 Likert-type questions, rated on a scale of 1 (strongly disagree) to 5 (strongly agree).

Procedures

Baseline

Baseline data were collected during convenient times for the classroom as determined by the teachers and daily schedules. During each data collection session at least one member of the research team went to the classroom and gave the paraeducator the data collection sheet. The paraeducator identified the target child they would collect data on for the session. As the intervention was targeting paraeducators' use of data collection and not student behavior, it was determined that who the paraeducator observed was less important than how the paraeducator conducted the observation and the accuracy of her data. Thus, the paraeducator was able to identify the target child for whom she would collect data each session. Thus, over the course of the study, different students were chosen to watch, which also provided an opportunity for the paraeducator to experience collecting data on different children. The variation in students, including differences in behavior, did not affect the outcome as there were no systematic differences in the students chosen from baseline to intervention. Anecdotally, all paraeducators chose the student with whom they were already working when the research arrived to begin the study rather than randomly selecting from the classroom. Once the target student was identified, the paraeducator and researcher then simultaneously started their timers. At the end of each 10-s interval, the paraeducator and researcher marked the presence or absence of on-task

Table 2 Teacher Social Validity by Participant

Item	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1. Coaching with performance feedback was an effective means to train paraeducators	0	0	0	1	1
2. Providing feedback during coaching sessions on my paraeducators actual performance was helpful	0	0	0	1	1
3. I liked having the opportunity to meet with my para(s) after we completed the data collection	0	0	0	1	1
4. Coaching sessions were frequent enough to support para's accurate data collection in the classroom	0	1	0	1	0
5. It was possible to find time to coach my paraeducators	0	0	1	1	0
6. I will use coaching in the future to provide training and feedback to my para(s) regarding other practices in the classroom	0	0	0	1	1
7. Coaching is an efficient means for me to provide supervision and feedback to my paraeducators	0	0	0	1	1
8. I believe data collection is one of my para's job responsibilities	0	0	0	1	1
9. It is feasible for my paraeducators to collect momentary time sampling data in the classroom	0	0	1	0	1
10. As a result of this project my paraeducators have learned how to collect momentary time sampling data and will be able to accurately collect this data in the future when asked	0	0	0	1	1
11. I will ask my paraeducators to collect data in the future using momentary time sampling	0	0	0	1	1

Table 3 Paraeducator social validity by participant

Item	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1. Coaching from my supervising teacher was effective	0	0	0	0	3
2. The suggestions I received from my supervising teacher during coaching sessions were useful	0	0	0	0	3
3. Feedback on my actual performance during coaching sessions was helpful	0	0	0	0	0
4. I liked having the opportunity to meet with the supervising teacher after we completed the data collection	0	0	0	1	2
5. Supervisor teacher coaching sessions were frequent enough to support my implementation of data collection in the classroom	0	0	1	0	2
6. I would like coaching from my supervising teacher in the future about other practices I use in the classroom	0	0	0	0	3
7. Compared to other training methods I have experienced as a para, coaching was more helpful	0	0	0	2	1
8. I believe data collection is one of my job responsibilities	0	0	1	1	1
9. It is feasible for me to collect momentary time sampling data in the classroom	0	0	1	0	2
10. I have learned how to collect momentary time sampling data and will be able to accurately collect this data in the future when asked	0	0	0	0	3
11. Receiving verbal feedback about my data collection was helpful	0	0	0	0	3

behavior at that moment. Recording continued for 5 min. Once the session was over, the researcher collected the paraeducator's data collection sheet. The accuracy of the paraeducator's data collection was then calculated using number of agreements divided by agreements plus disagreements times 100, and subsequently graphed. Throughout baseline, the participants did not receive any feedback regarding how to collect MTS data or the accuracy of their data collection. Following a minimum of 5 data points and a stable pattern of data, the intervention was introduced.

Teacher-as-Coach Training

A researcher with expertise in coaching trained the teachers in the coaching procedure including appropriate use of performance feedback. The trainer followed a structured training protocol which included background information on the strength of coaching as it relates to adult learning and improving performance. The training, which lasted approximately 1 h, also included specific training on procedures for implementing the coaching protocol. Examples of supportive and constructive feedback (video and verbal) were provided including both live demonstrations and video models. Teachers then practiced giving feedback via role play with the trainer. All participant questions were answered and participants reported feeling comfortable with the coaching procedures.

Coaching Intervention

Following teacher training, each teacher provided coaching with performance feedback to the paraeducators. Intervention data were collected in an identical manner to baseline data collection. Coaching sessions occurred approximately two times a week during times that were most convenient for the teacher. During each coaching session the teacher completed a coaching log that included recording the length of session, number of participants, training needed to reach goal, next meeting date, and any barriers that were present in the session. As researchers collected the reliability data, the teachers were given each paraeducator's reliability percentage prior to each coaching session. The researcher also provided notes of what child behavior was occurring during the discrepancies and provided that information to the teacher. The teacher used these reliability percentages to give the paraeducator feedback on their data collection in the classroom. The first coaching session was based on the data collected in the last baseline session for each participant. Teachers began each session with a positive greeting and then stated that the goal of the study was to reach a minimum of 80% reliability of data collection for at least three sessions. Teachers then gave their paraeducators the current reliability percentage and provided supportive feedback on their progress. The teacher gave feedback on the percentage and discussed challenges as described by the paraeducator. The teacher and paraeducator talked about issues and difficulties with collecting data and problem-solved any issues. For instances, looking at the target student at the exact moment the 10-s interval sounded was suggested to avoid uncertainty regarding the on-task behavior of the target student at the exact moment. The teacher and paraeducator then set up action steps for the next couple of days, and ended with a positive

closing. Coaching sessions occurred in dyads (teacher–paraeducator) until the paraeducators met criterion for the study. Each coaching session lasted approximately 10 min with a range of 6–15 min. Coaching continued until all paraeducators had at least 80% fidelity of for three sessions. Once criterion was reached, participants were not given any further feedback on their data collection fidelity. Both teachers and paraeducators were given a social validity form to complete and return to the research team in a sealed envelope.

Maintenance

Maintenance data were collected 2 months after each paraeducator met criterion. During the 1- to 2-month period, the teachers did not coach the paraeducators in data collection. Procedures during the maintenance phase were the same as in Baseline with a minimum of three data points collected per participant.

Data Analysis

Visual analysis of the graphical display of data was used to evaluate the functional relation between implementation of performance feedback delivered by the supervising teacher and improvements in paraeducators' accuracy of MTS data of students' on task behavior, based on the guidelines described in Kratochwill et al. (2013). The consistency of data within and across phases, as well as changes in level, trend, and variability were evaluated.

Results

Coaching Intervention

Figure 2 shows a graphical display of percentage of agreement of MTS data for Jasmine, Olive, and Amy during baseline and coaching phases. Figure 3 shows a graphical display of the percentage of on-task behavior as rated by each individual rater (paraeducator, primary researcher, and secondary researcher). It is important to note that agreement between the paraeducator and the primary research was based on a conservative point-by-point agreement, as displayed in Fig. 2, rather than overall agreement in the percentage of off-task behavior as seen in Fig. 3. Thus, review of each data series within and across paraeducators in Fig. 3 seems to indicate a higher level of agreement in on-task behavior, as evidence by the overlapping data paths, than what is indicated in Fig. 2. This is because there could be agreement in the number of intervals with on-task behavior; however, each of those could be different intervals. For example, two raters could both rate a given student as being on-task for 50% of intervals but have 0% agreement on which of those intervals the student was on-task. Thus, the more consistent point-by-point reliability measure was utilized to ensure agreement on the occurrence or nonoccurrence of behavior within specific intervals.

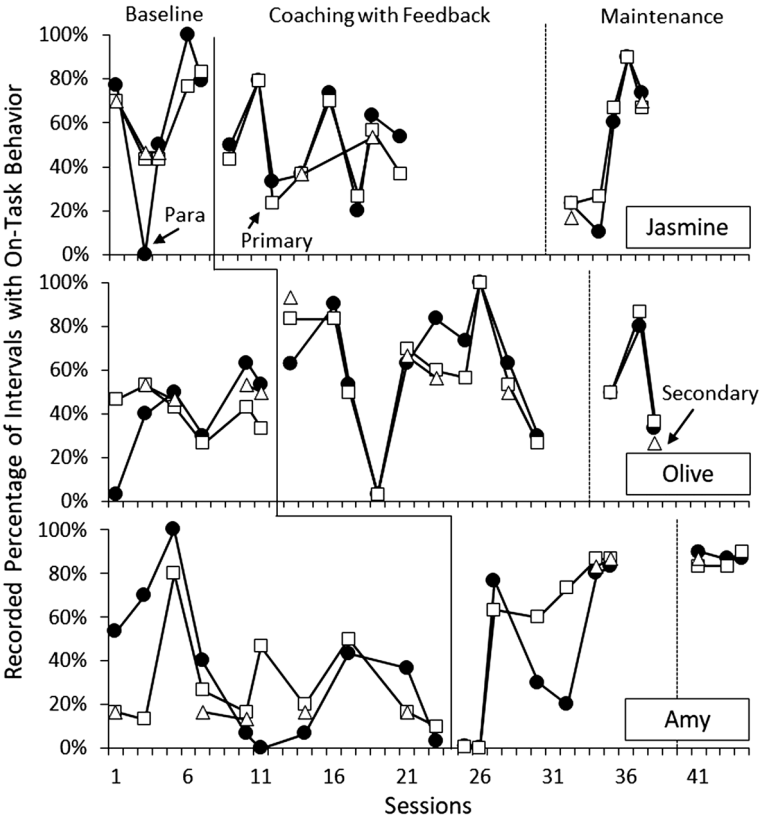

Fig. 3 Recorded percentage of 10-s intervals with on-task behavior for the paraeducator ("para"), primary researcher ("primary"), and secondary researcher ("secondary")

Baseline data in Fig. 2 indicate percentage of inter-observer agreement of MTS of on-task behavior remained well below 80% across all participants for most baseline sessions. Introduction of coaching yielded an immediate increase in inter-rater reliability as is noted by the intercept gap and increase in level between the last baseline data point and first coaching data point across participants. Olive and Jasmine required 3 coaching sessions to meet the criteria of three sessions at or above 80% reliability, whereas Amy required 4 coaching sessions.

Jasmine's baseline data were stable and remained below 20%. Upon introduction of coaching there was an immediate increase in level and trend. Her percentage of agreement with the researcher remained above 80% throughout the phase, and there was no overlap in data between the baseline and intervention conditions. Jasmine's data collection reliability remained at a high level in the maintenance condition (1 and 2 months later), similar to the level of the intervention condition.

Olive's percent of agreement was below 30% through the baseline phase. Introduction of the coaching intervention yielded an immediate increase in level with the introduction of coaching, and again there was no overlap in data between the

baseline and intervention conditions. Maintenance data were collected 2 months after intervention ended and Olive's data remained high with a level comparable to the intervention condition.

Amy's baseline data remained below 50% with the exception of session 4 (80%). This rather high data point in baseline is likely attributed to the fact that the target student was engaging in overt, challenging behavior (e.g., screaming, spitting) for the entire interval for all but four intervals of the session and was clearly not on-task. This only occurred for 1 session throughout all phases. Amy's intervention data show an increasing trend and immediate increase in level upon implementation of the intervention. Furthermore, there was minimal overlap (two data points) across conditions. Accuracy of data collection remained high during the 1-month maintenance probes.

Occurrence and Nonoccurrence Agreement

In an attempt to further understand the nature of agreements and disagreements that occurred, percentage of occurrence and nonoccurrence agreement was calculated at the end of the study in a manner similar to that described for IOA calculations. Average percent of occurrence, nonoccurrence and overall agreement in each phase for each paraeducator is indicated in Table 4. It is important to note that these occurrence and nonoccurrence data were calculated for descriptive purposes only and thus, not shared with the teacher or paraeducator and not utilized during coaching session or decision-making regarding phase changes.

As can be seen in Table 4, Jasmine's sources of agreement appear to be pretty evenly dispersed across occurrence and nonoccurrence of on-task behavior for each phase of the study, noting agreement was below 15% for each during baseline with significant improvement during the intervention phase. However, agreements for Olive and Amy were notably higher for occurrence of on-task behavior rather than nonoccurrence during baseline and intervention, as well as maintenance for Amy. Increases in average percent of agreement during intervention when compared to baseline are noted for both occurrence and nonoccurrence agreement as well as overall agreement for all paraeducators.

Social Validity

Both teachers, Nancy and Eleanor, either agreed or strongly agreed that the coaching and feedback were helpful in changing their paraeducators' behavior, that coaching was beneficial, and they would use coaching in the future (see Table 2). In regard to frequency of coaching, Eleanor did not think coaching was frequent enough, whereas Nancy did. Additionally, Nancy neither agreed nor disagreed that it was possible to find time to coach her paraeducators or that coaching was feasible, whereas Eleanor indicated agreement to both. Both teachers agreed that data collection was one of their paraeducators' responsibilities, that their paraeducators had learned to collect MTS data, and that they would ask them to collect data in the future utilizing MTS.

Table 4 Average percent of agreement for each para across phases

Phase	Overall (%)	Occurrence (%)	Nonoccurrence (%)
Jasmine			
Baseline	9.5	13	6
Intervention	88	88	88
Maintenance	76	78	73
Olive			
Baseline	16	30	2
Intervention	74	86	61
Maintenance	88.5	89	88
Amy			
Baseline	36	46	27
Intervention	60	69	51
Maintenance	71	95	47

A summary of the paraeducators' responses to the *Social Validity* measure is given in Table 3. Overall the paraeducators agreed to strongly agreed that coaching and receiving feedback from their supervising teacher was beneficial, suggestions from the teacher were helpful, and they liked meeting with their supervising teacher regarding their data collection accuracy. Although Olive and Jasmine strongly agreed that the coaching sessions were frequent enough and feasible, Amy neither agreed nor disagreed. All agreed or strongly agreed that coaching was more helpful than other training methods they had previously experienced and that they would like coaching on other practices. All indicated they had learned how to collect MTS data and would be able to do it in the future. Jasmine neither agreed nor disagreed that collecting data was one of her job responsibilities, whereas Olive and Amy agreed to strongly agreed that it was one of their responsibilities.

Discussion

Summary

Given the importance of progress monitoring for students with disabilities, ensuring all education staff are well trained to accurately collect data using a variety of methods, such as MTS, is essential. However, paraeducators are woefully underprepared to engage in this task despite its prominence in their job responsibilities (Carter et al. 2009). Identification of efficient training methods that include instruction, modeling, opportunity for practice and ongoing performance feedback (Snyder et al. 2015) is necessary. The current study sought to evaluate the impact of a teacher-as-coach model (Mason et al. 2017) to train paraeducators to collect accurate momentary time sampling data on students' on-task behavior. Overall results demonstrate a functional relationship between implementation of the teacher-as-coach model and improvements in the paraeducators' accurate data collection. Furthermore, the mean

of the intervention phase was above 80% accuracy for two of three paraeducators and the mean of the maintenance phase was above 80% for all three paraeducators. Based on field standards regarding reliability (e.g., Kratochwill et al. 2013), the paraeducators achieved adequate reliability with the research team by the maintenance probes, suggesting that the data can be used to make instructional decisions.

It is important to note that all of the participating paraeducators had previous training in data collection following the typical group, in-service training model without follow-up and feedback. Despite this, the accuracy of MTS data of students' on-task behaviors across all participants was unreliable during baseline. This performance is consistent with previous research indicating that large-group, speaker–listener trainings rarely translates to practice (Schepis et al. 2003; Tate et al. 2005). On a more positive note, implementation of the teacher-as-coach model yielded an immediate increase in the accuracy of data collection following the first coaching session. Further, achieving the criterion of three sessions at or above 80% reliability or higher took an average of 3, 15-min coaching sessions. Thus, the paraeducators were able to implement accurate MTS data collection which maintained once the coaching ended. Given the efficiency and effectiveness of the teacher-as-coach model, it is seemingly a more viable approach to training paraeducators than typical in-service professional development.

Another important factor to consider is the limited amount of time required to train the teachers to implement the coaching model. The teacher-as-coach training lasted only 1 h for each teacher and teachers were subsequently able to implement the model with high rates of procedural fidelity when provided with reliability data from the research team. As teachers are charged with the responsibility of supervising paraeducators, typically without training on how to effectively carry out this charge (Kratz et al. 2015), having a brief yet effective training model with which to equip supervising teachers such as the teacher-as-coach model implemented in this study, would likely assist with this responsibility.

Although this study constitutes progress in terms of our evidence regarding methods to prepare supervising teachers to train and supervise paraeducators, the procedure utilized in this study is best-described as collaboration between researchers and teachers. Researchers carried out a considerable portion of the work in this study including collecting the reliability data with the paraeducator, calculating the reliability, documenting relevant student behavior for the teacher to reference during coaching sessions, and providing the necessary data to the teacher. Had the teacher been charged with completing these tasks, in addition to the coaching and feedback, it is estimated this would necessitate an additional 30 min minimum of the teacher's time per coaching session. Given this, it may have proven impossible for the teacher to carry out.

This is particularly true considering the social validity data. It is highly likely that the disagreement between the two teachers regarding the feasibility of the coaching model was related to the fact that Nancy was coaching two paraeducators while also having responsibility for the classroom, whereas Eleanor was only coaching one paraeducator and did not have responsibility for the overall classroom. It was more challenging for Nancy to find time to meet with her paraeducators, suggesting the perceived efficiency of the teacher-as-coach model is likely negatively correlated

with the number of paraeducators the teacher is charged with supervising as well as the different job responsibilities of each of the participating teachers. That is, with increasing numbers of paraeducators, the ability to find time to coach each would likely decrease. Additionally, Eleanor's main job responsibility was supporting staff in various classrooms thus her schedule was seemingly more flexible. Eleanor may have been more willing to indicate coaching needed to be more frequent as training one paraeducator was feasible and increasing frequency seemed less daunting, perhaps, than it did for Nancy. Considering both teachers and paraeducators indicated the coaching and feedback was effective for improving paraeducators' MTS data collection skills and that they would like to utilize it for improving other instructional strategies, identifying mechanisms to increase the efficiency is warranted.

Limitations

In addition to the limitation previously discussed regarding the researchers' role compared to the teachers', this study has some additional limitations. First is the attrition of Jasmine's supervising teacher, Bianca from the study, which is a significant limitation. Although this ultimately did not affect study procedures or the results, her withdrawal from the study may be an indication of her lack of acceptability of the coaching procedure. Qualitatively, Bianca did indicate she was overwhelmed with other, non-work-related expectations and did not want the additional responsibility. Given the study participation was optional, it could be that it was the easiest task to relinquish and not purely a reflection of the procedure itself. Regardless, this further limits the confirmation of the feasibility of the teacher-as-coach model. Therefore, we have evidence from this study and others (Mason et al. 2017) that provision of coaching and feedback to paraprofessionals by the supervising teachers is an effective training model, yet we continue to need evidence regarding its feasibility.

Additionally, it is important to point out that MTS data collection is only one type of data collection and, further, on-task behavior is only one type of behavior for which a paraprofessional might be asked to collect data. Thus, given the rather narrow range encompassed, there is no evidence to support generalization of the findings to other types of data collection systems or monitoring of student performance.

A final limitation is that this study collected data on the outcome, agreement with researcher collected data, rather than the process, fidelity, of data collection. First, this focus did not take into account the inherent error in the researcher collected data and attributes all disagreements to paraeducator error. It is possible that focusing on steps for collecting data in addition to the accuracy of data collection would have decreased the amount of time that it took the paraeducators to achieve the 80% criterion. Further, such data would allow for more focused coaching. That is, the teacher could have focused on specific procedures the paraeducators were implementing inaccurately, such as scoring the interval before or after the exact end of the interval, rather than utilizing the coaching session to explore reasons why the data may not have agreed with the researcher collected data. Additionally, based on the post hoc data analysis regarding occurrence and nonoccurrence agreement, it is possible that

some of the challenges in data collection may have been related to clarification of the definition of on-task behavior rather than inaccurate data collection procedures. Thus, a combination of procedural fidelity and percentage of agreement data would likely have assisted in identifying the breakdowns more rapidly. Additionally, given the operational definition of on-task behavior typically varies from student to student, particularly when working with students with autism who may engage in stereotypy (i.e., hand-flapping, rocking, vocalizations), defining "on-task" based on the individual student's manifestation of on-task behavior rather than a broad, universal definition, may have further increased the accuracy of data collection.

Implications for Research and Practice

This study provides support for the use of coaching with performance feedback to increase paraeducators' data collection skills. Knowledge of effective training strategies is critical for special education teachers, who are often tasked with supervising paraeducators. Practitioners should evaluate the extent to which teachers use effective training strategies and provide PD for teachers who do not currently use effective training strategies with their paraeducators. Additionally, allotting time for the teacher to meet with paraeducators specifically to engage in supervision tasks such as coaching in effective practices and targeted paraeducator goals would increase the feasibility of ongoing training for paraeducators.

The current study has focused exclusively on training paraeducators in discrete skills. Future research should assess the extent to which teachers are able to provide training to paraeducators on a more comprehensive set of skills such as the paraeducator standards established by CEC (CEC 2004). This would likely necessitate establishment of a comprehensive paraeducator assessment to identify specific "target skills" for which the paraeducator needs training. Additionally, implementation protocols that allow the teacher to measure fidelity of implementation for each of the skills would ease the burden for supervising teachers and allow for systematic, data-driven paraeducator training and supervision.

Future research should consider replicating this study in different school settings. For this study, paraeducators collected data during one-on-one instructional time with students in a special education classroom. It is unclear if the same results would be obtained during group instruction or in a general education classroom, with more children present. In addition, future research should target different data collection skills, such as measuring student progress on an academic goal.

Conclusion

Given the frequent use of paraeducators to support students in special education classrooms and the requirement to monitor and demonstrate student progress, it is imperative that paraeducators collect accurate data on student behavior. The present study indicates that with support teachers can effectively train paraeducators to accurately collect data within a special education classroom. Teachers and administrators

should encourage paraeducators to collect data on IEP objectives and the accuracy of data collection should be monitored. For those paraeducators who need additional training in data collection, teacher-implemented coaching and performance feedback may be one option to quickly increase the reliability of paraeducator's data collection.

Acknowledgements This work was supported by the Institute of Education Sciences, National Center for Education Research (Grant Number R305H140048).

References

Alberto, P. A., & Troutman, A. C. (2003). *Applied behavior analysis for teachers* (6th ed.). Columbus, OH: Charles E. Merrill.

Babkie, A. M., & Provost, M. C. (2004). Teachers as researchers. *Intervention in School and Clinic, 39,* 260–268.

Boardman, A. G., Arguelles, M. E., Vaughn, S., Hughes, M. T., & Klingner, J. (2005). Special education teachers views of research-based practices. *The Journal of Special Education, 39,* 168–180.

Brabec, K., Fisher, K., & Pitler, H. (2004). Building better instruction: How technology supports nine research-proven instructional strategies. *Learning & Leading with Technology, 31*(5), 6–11.

Brock, M. E., & Carter, E. W. (2013). A systematic review of paraprofessional-delivered educational practices to improve outcomes for students with intellectual and developmental disabilities. *Research and Practice for Persons with Severe Disabilities, 38*(4), 211–221.

Brock, M. E., & Carter, E. W. (2015). Effects of a professional development package to prepare special education paraprofessionals to implement evidence-based practice. *The Journal of Special Education, 49*(1), 39–51.

Carter, E., O'Rourke, L., Sisco, L. G., & Pelsue, D. (2009). Knowledge, responsibilities, and training needs of paraprofessionals in elementary and secondary schools. *Remedial and Special Education, 30,* 344–359.

Clare, S. K., Jenson, W. R., Kehle, T. J., & Bray, M. A. (2000). Self-modeling as a treatment for increasing on-task behavior. *Psychology in the Schools, 37*(6), 517–522.

Collinson, V. (2000). Staff development by any other name: Changing words or changing practices? *Educational Forum, 64,* 124–132.

Council for Exceptional Children. (2004). *What every special educator must know: Ethics, standards and guidelines for special educators* (5th ed.). Arlington, VA: Author.

Finn, L., Ramasamy, R., Dukes, C., & Scott, J. (2015). Using WatchMinder to increase the on-task behavior of students with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(5), 1408–1418.

Giangreco, M. F., Broer, S. M., & Edelman, S. W. (2002). "That was then, this is now!" Paraprofessional supports for students with disabilities in general education classrooms. *Exceptionality, 10*(1), 47–64.

Griffin-Shirley, N., & Matlock, D. (2004). Paraprofessionals speak out: A survey. *RE: View, 36,* 127–137.

Gunter, P. L., Callicott, K., Denny, K., & Gerber, B. L. (2003a). Finding a place for data collection in classrooms for students with emotional/behavioral disorders. *Preventing School Failure, 48*(1), 4–8.

Gunter, P. L., Venn, M. L., Patrick, J., Miller, K. A., & Kelly, L. (2003b). Efficacy of using momentary time samples to determine on-task behavior of students with emotional/behavioral disorders. *Education and Treatment of Children, 26*, 400–412.

Hume, K., Loftin, R., & Lantz, J. (2009). Increasing independence in autism spectrum disorders: A review of three focused interventions. *Journal of Autism and Developmental Disorders, 39*(9), 1329–1338.

Individuals with Disabilities Education Improvement Act. (2004). 20 U.S.C. § 1400.

Jessel, J., Ingvarsson, E. T., Whipple, R., & Kirk, H. (2017). Increasing on-task behavior of an adolescent with autism using momentary differential reinforcement. *Behavioral Interventions, 32*(3), 248–254.

Kartal, M. S., & Ozkan, S. Y. (2015). Effects of class-wide self-monitoring on on-task behaviors of preschoolers with developmental disabilities. *Education and Training in Autism and Developmental Disabilities, 50*(4), 418.

Keller, C. L., Bucholz, J., & Brady, M. P. (2007). Yes, I can! Empowering paraprofessionals to teach learning strategies. *Teaching Exceptional Children, 39*(3), 18–23.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., et al. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*(1), 26–38.

Kratz, H. E., Locke, J., Piotrowski, Z., Ouellette, R. R., Xie, M., Stahmer, A. C., et al. (2015). All together now: Measuring staff cohesion in special education classrooms. *Journal of Psychoeducational Assessment, 33*(4), 329–338.

Ledford, J. R., Ayres, K. M., Lane, J. D., & Lam, M. F. (2015). Identifying issues and concerns with the use of interval-based systems in single case research using a pilot simulation study. *The Journal of Special Education, 49,* 104–117.

Lewis, T. J., Hudson, S., Richter, M., & Johnson, N. W. (2004). Scientifically supported practices in emotional and behavioral disorders: A proposed approach and brief review of current practices. *Behavioral Disorders, 29*(3), 247–259.

Lewis-Palmer, T., Sugai, G., & Larson, S. (1999). Using data to guide decisions about program implementation and effectiveness. *Effective School Practices, 17*(4), 47–53.

Logan, K. R., & Stein, S. S. (2001). The research lead teacher model: Helping general education teachers deal with classroom behavior problems. *Exceptional Children, 33*(3), 10–15.

Luke, S., Vail, C. O., & Ayres, K. M. (2014). Using antecedent physical activity to increase on-task behavior in young children. *Exceptional Children, 80*(4), 489–503.

Mason, R. A., Schnitz, A. G., Wills, H. P., Rosenbloom, R., Kamps, D. M., & Bast, D. (2017). Impact of a teacher-as-coach model: Improving paraprofessionals fidelity of implementation of discrete trial training for students with moderate-to-severe developmental disabilities. *Journal of Autism and Developmental Disorders, 47*(6), 1696–1707.

Rabiner, D., & Coie, J. D. (2000). Early attention problems and children's reading achievement: A longitudinal investigation. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(7), 859–867.

Rispoli, M., Neely, L., Lang, R., & Ganz, J. (2011). Training paraprofessionals to implement interventions for people autism spectrum disorders: A systematic review. *Developmental Neurorehabilitation, 14*(6), 378–388.

Rivera, C. J., Mason, L. L., Jabeen, I., & Johnson, J. (2015). Increasing teacher praise and on task behavior for students with autism using mobile technology. *Journal of Special Education Technology, 30*(2), 101–111.

Rockwell, S. (2008). Working smarter, not harder: Reaching the tough to teach. *Kappa Delta Pi Record, 44,* 154–159.

Sandall, S. R., Schwartz, I. S., & Lacroix, B. (2004). Interventionists' perspectives about data collection in integrated early childhood classrooms. *Journal of Early Intervention, 26,* 161–174.

Schatz, R. B., Peterson, R. K., & Bellini, S. (2016). The use of video self-modeling to increase on-task behavior in children with high-functioning autism. *Journal of Applied School Psychology, 32*(3), 234–253.

Schepis, M. M., Reid, D. H., Ownbey, J., & Clary, J. (2003). Training preschool staff to promote cooperative participation among young children with severe disabilities and their classmates. *Research and Practice for Persons with Severe Disabilities, 28*(1), 37–42.

Shernoff, E. S., & Kratochwill, T. R. (2007). Transporting an evidence-based classroom management program for preschoolers with disruptive behavior problems to a school: An analysis of implementation, outcomes, and contextual variables. *School Psychology Quarterly, 22,* 449–472.

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*(1), 5–15.

Smith, S. W., Daunic, A. P., & Taylor, G. G. (2007). Treatment fidelity in applied educational research: Expanding the adoption and application of measures to ensure evidence-based practice. *Education and Treatment of Children, 30,* 121–134.

Snyder, P. A., Hemmeter, M. L., & Fox, L. (2015). Supporting implementation of evidence-based practices through practice-based coaching. *Topics in Early Childhood Special Education, 35,* 133–143. https://doi.org/10.1177/0271121415594925.

Southall, C. M., & Gast, D. L. (2011). Self-management procedures: A comparison across the autism spectrum. *Education and Training in Autism and Developmental Disabilities, 46*, 155–171.

Stahr, B., Cushing, D., Lane, K., & Fox, J. (2006). Efficacy of a function-based intervention in decreasing off-task behavior exhibited by a student with ADHD. *Journal of Positive Behavior Interventions, 8*(4), 201–211.

Stecker, P. M., Lembke, E. S., & Foegen, A. (2008). Using progress-monitoring data to improve instructional decision making. *Preventing School Failure, 52*(2), 48–58.

Stichter, J. P., Lewis, T. J., Richter, M., Johnson, N. W., & Bradley, L. (2006). Assessing antecedent variables: The effects of instructional variables on student outcomes through in-service and peer coaching professional development models. *Education and Treatment of Children, 29,* 665–692.

Sutherland, K. S., & Wehby, J. H. (2001). Exploring the relationship between increased opportunities to respond to academic requests and the academic and behavioral outcomes of students with EBD: A review. *Remedial and Special Education, 22*(2), 113–121.

Tate, T. L., Thompson, R. H., & McKerchar, P. M. (2005). Training teachers in an infant classroom to use embedded teaching strategies. *Education & Treatment of Children, 28,* 206–221.