Exceptional Children
2021, Vol. 87(2) 217–235

DOI: 10.1177/0014402920947641
journals.sagepub.com/home/ecx
SAGE

Efficacy of Tiered Training on Paraeducator Implementation of Systematic Instructional Practices for Students With Severe Disabilities

Matthew E. Brock[1], Mary A. Barczak[1], Eric J. Anderson[1], and Nichole M. Bordner-Williams[1]

Abstract

Existing approaches for training paraeducators rely heavily on intensive one-to-one coaching and may not be feasible in practice. In this study, we test a tiered training model in which all paraeducators first received group training, and then coaching was provided only for the subset who did not meet performance criteria after group training. Using a concurrent multiple-probe design staggered across classrooms, we demonstrated a functional relation between the tiered model and implementation fidelity of two systematic prompting strategies across 13 paraeducators in five schools. Nine paraeducators achieved the performance criterion for both practices with group training alone, and the remaining four met the criterion after teacher-delivered coaching. In addition, paraeducators generalized implementation to new situations, and students with severe disabilities who received instruction made progress on individualized goals. Based on these findings, a tiered training model is a feasible and promising means to train paraeducators.

All students deserve to be taught with instructional practices that have been proven to be effective. When teachers use evidence-based practices that are supported by rigorous scientific research, they increase the likelihood that their students will learn (Odom et al., 2019). In addition to being the most ethical way to educate students, evidence-based practices are also mandated by federal law. Specifically, the Every Student Succeeds Act (2015–2016) requires schools to use evidence-based practices that have been shown to improve student outcomes.

Although all students should be taught using evidence-based practices, this may be particularly important for students with severe disabilities—students with intellectual disability, autism, or multiple disabilities who qualify for their state's alternate assessment for students with severe cognitive disabilities. Compared to their peers, students with severe disabilities are at the highest risk for poor postschool outcomes, including unemployment, lack of independent living, and limited opportunities to demonstrate self-determination (Rusch et al., 2009). Even more than their peers, these students are in desperate need of highly effective instruction. To address these needs, scholars have identified evidence-based practices that promote improved outcomes for students with severe disabilities (Spooner et al., 2019; Wong et al., 2015).

[1]The Ohio State University

Corresponding Author:
Matthew E. Brock, Crane Center for Early Childhood Research and Practice, The Ohio State University, A334 PAES Building, 305 Annie & John Glenn Ave, Columbus, OH 43210, USA.
Email: brock.184@osu.edu

Some of the most basic evidence-based practices for this population are associated with systematic instruction—task analytic instruction and systematic prompting (Neitzel & Wolery, 2009c). Task-analytic instruction is based on breaking down a multistep skill into its component steps, and teaching and measuring progress on individual steps. Once broken down into components steps, these skills can be taught using systematic prompting. Systematic prompting involves providing assistance that enables students to be successful as well as a plan for fading the assistance over time. Two specific systematic instructional strategies—simultaneous prompting and least-to-most prompting—are especially well suited for initially teaching skills and then fading support. Simultaneous prompting, a form of errorless teaching, is a strategy that ensures initial student success and eliminates errors (Collins, 2012; Neitzel & Wolery, 2009b). Each time the student is given an opportunity to respond, the instructor immediately delivers a prompt that is sufficient to ensure a correct response (i.e., a controlling prompt). Independent responding is then assessed periodically using probe trials in which the prompt is removed. Least-to-most prompting, also known as system of least prompts, is a strategy that promotes student independence by fading support contingent on student responding (Collins, 2012; Neitzel & Wolery, 2009a). First the student is given an opportunity to respond independently, and progressively intensive prompts are provided, if necessary. As the student provides more independent responses, less intensive prompts are provided. These systematic instructional strategies are highly versatile and can be used to target a range of outcomes, including academic, social, communication, vocational, and adaptive skills (Shepley et al., 2019; Wong et al., 2015).

Despite the utility of these practices and their strong evidence base, task-analytic instruction and systematic prompting—like many other evidence-based practices—do not always make their way into classrooms. Scholars have highlighted a wide gap between evidence-based practices and the instruction that actually occurs in many schools (Odom et al., 2019). In surveys of special education teachers, many teachers did not understand the term *evidence-based practice* (Stahmer et al., 2005), reported implementing evidence-based practices infrequently (Brock, Dynia, et al., 2019), or reported implementing unproven practices more often than evidence-based practices (Hess et al., 2008).

Bridging the research-to-practice gap is a challenge across all educational fields but may be especially challenging when educating students with severe disabilities who often receive instruction from paraeducators. Special education paraeducators outnumber licensed special education teachers (U.S. Department of Education, 2019), with teachers who serve students with severe disabilities often supervising multiple paraeducators (Suter & Giangreco, 2009). Nearly all paraeducators spend at least some time delivering instruction to students, despite having received little or no training in evidence-based practices (Carter et al., 2009).

Bridging the research-to-practice gap is a challenge across all educational fields but may be especially challenging when educating students with severe disabilities who often receive instruction from paraeducators.

Federal law allows for paraeducators to deliver instruction when given appropriate training and supervision, but the law does not delineate *how* to appropriately provide this training and supervision (Individuals With Disabilities Education Improvement Act, 2006). This puts teachers in a difficult situation, as they are left with little guidance about how to best train and supervise paraeducators on their team (Wermer et al., 2018). Fortunately, scholars have been working to address this problem by testing different models of paraeducator training. Over 50 studies have been published that focus on how to train paraeducators to implement interventions with students with developmental disabilities (cf. Brock & Carter, 2013; Brock & Anderson, 2020). Across these studies, a number of con-

vergent findings have emerged. First, when paraeducators are trained to implement specific evidence-based practices, they are far more effective than when they are asked to provide general support. This is exemplified by the contrast between business-as-usual and treatment conditions in these studies. Second, effective paraeducator training typically includes a combination of an implementation checklist that delineates implementation steps, didactic instruction on each step, modeling of the steps, and repeated performance feedback on the paraeducator's implementation with students (Brock & Anderson, 2020; Brock et al., 2017)

These studies take important steps in moving the field forward, but there are a number of challenges that have not yet been addressed. First, the existing literature relies heavily on one-to-one coaching for all paraeducators (Brock & Anderson, 2020; Brock et al., 2017). Using coaching as a stand-alone tool to train every paraeducator in every practice associated with their job is resource-intensive and time-consuming (Russo, 2004). This lack of feasibility may be one of the reasons why very few schools have embraced research-based approaches for training paraeducators (Carter et al., 2009). Similar to one-on-one tutoring for students, we think that coaching is a powerful tool that should be used judiciously and not as a stand-alone solution. Second, nearly all paraeducator training studies focus narrowly on training a paraeducator to implement an intervention in one context with one student and do not measure the degree to which paraeducators are able to generalize their implementation to other students or teaching situations (Brock & Anderson, 2020; Brock et al., 2017). Understanding the degree to which skills generalize is critical. Without measuring generalization, it is unclear if paraeducators would require additional training before they can implement a practice with fidelity with a new student or in a new situation. Third, most studies involve training that is delivered by researchers (Brock & Anderson, 2020; Brock et al., 2017). Therefore, it is unclear if the teachers—the individuals responsible for paraeducator training and supervision—could feasibly implement the training procedures.

Similar to one-on-one tutoring for students, we think that coaching is a powerful tool that should be used judiciously and not as a stand-alone solution.

In this study, we address all three of these challenges. First, we address the challenge of feasibility by testing a tiered model in which group training is delivered to all paraeducators and one-to-one coaching is delivered only if and when paraeducators do not meet a criterion for implementation fidelity. Similar to a response-to-intervention (RtI) model, the most intensive approach (i.e., one-to-one coaching) would be used only when less intensive approaches are insufficient. Second, we address the challenge of generalization by measuring paraeducator implementation in two different contexts: (a) teaching the instructional target for which they were observed and received feedback and (b) teaching a different instructional target for which they did not receive feedback. Third, we involved teachers in the coaching process. This enables us to gauge the degree to which it is feasible and effective for teachers to deliver coaching and leverage the knowledge that teachers have about individual students. We designed our study to address the following research questions:

1. What are the effects of a tiered training on paraeducator implementation of systematic prompting strategies (i.e., simultaneous prompting and least-to-most prompting) with students with severe disabilities?
2. What are the effects of the individual tiers (i.e., Tier 1 = group training; Tier 2 = one-to-one coaching)?
3. When directed to use systematic prompting strategies in a new teaching situation for which they have received no coaching or feedback, to what degree do paraeducators generalize their implementation fidelity?

4. What progress do students make on individualized goals after receiving instruction from paraeducators who have participated in tiered training?

Method

Participants and Settings

After securing approval from a university institutional review board and school districts, we recruited 13 paraeducators from five different schools who delivered instruction to students with severe disabilities (i.e., students eligible for their state's alternate assessment) from three school districts that served students in rural and suburban communities in a midwestern state. First, we asked administrators in these districts to identify teams of special education teachers and paraeducators who served students with severe disabilities and might be willing to participate in a research study. The administrators directed us to teachers at five different schools who all consented to participate and arranged for us to meet with their teams of paraeducators. Next, we sought consent from the paraeducator, permission from the family of a target student with a severe disability to whom the paraeducator regularly delivered instruction, and assent from the target student. Characteristics of teachers, paraeducators, and students are summarized in Table 1. Students included in this study spent time in both general and special education settings. Our observations of paraeducator-implemented instruction took place in self-contained special education classrooms, although in several cases, instruction was designed to build on skills that would be utilized in general education classrooms. In these cases, teaching targets focused on content from the general education curriculum that was identified by general education teachers.

Experimental Design and Procedures

We implemented an experimental single-case-design study. Specifically, we implemented a concurrent multiple-probe-across-participants design (Gast et al., 2018) in which introduction of the intervention was staggered across the five schools, but data were collected at the level of individual paraeducators and students. We randomized the order in which the intervention would be introduced across schools using a computerized random-number generator. This randomization of tiers enhances the strength of the experimental design (Kratochwill & Levin, 2014).

Before the study began (i.e., prebaseline), we held one training session with paraeducators that focused on differentiating between discrete and chained skills and on conducting a task analysis of chained skills. Next, all paraeducators began the study in a "written directions" condition, which served as the baseline condition. Once all paraeducators demonstrated a stable pattern of responding and we had collected at least five data points for each paraeducator at School 1, we introduced group training at School 1. We introduced group training at each subsequent school after paraeducators in the previous tier demonstrated a clear change in level toward criterion-level implementation fidelity for simultaneous prompting and least-to-most prompting for the target goal (i.e., at least two data points demonstrating a trend toward criterion). After the group training condition, paraeducators entered a maintenance condition in which no training was delivered. If a paraeducator demonstrated criterion-level performance (i.e., >85% implementation fidelity) across both prompting procedures and both teaching situations, they remained in the maintenance condition for the rest of the experiment. If paraeducator performance dropped below criterion level for two consecutive sessions for either procedure for either target, we introduced teacher-delivered coaching. Next, we detail procedures for each condition.

Student Goal Selection and Task Analysis of Chained Skills (Prebaseline). We reviewed each students' individualized education program (IEP) with the supervising special education teacher and worked together to select two goals that could be appropriately targeted with systematic prompting. In several cases, we

Table 1. Characteristics of Teachers, Paraeducators, and Students.

	Teachers			Paraeducators				Students				
School number	Gender	Years experience	Highest education	Number	Gender	Years experience	Highest education	Age	Grade	Educational label	IQ score	Adaptive behavior score
1	Female	15	Master's	1	Female	4	Bachelor's	11	5	ID	52[a]	NR
				2	Female	6	Master's	11	5	ID	NR	NR
2	Female	5	Bachelor's	1	Female	1	Bachelor's	7	1	ASD	NR	NR
				2	Female	4	Bachelor's	8	3	OHI	NR	NR
				3	Female	1	Bachelor's	7	1	OHI	NV[a]	NR
3	Female	7	Master's	1	Female	5	Master's	5	K	ASD	NR	69[b]
				2	Female	3	Bachelor's	10	5	ASD	78[c]	71[d]
				3	Female	1	Bachelor's	6	K	ASD	NR	NR
4	Female	2	Bachelor's	1	Male	1	Bachelor's	16	11	ASD	42[c]	41[b]
				2	Female	8	Master's	17	11	ID	NR	NR
				3	Female	9	Bachelor's	15	9	OHI	71[a]	NR
5	Female	16	Master's	1	Male	2	Bachelor's	15	9	MD	NV[e]	NR
				2	Female	4	Bachelor's	11	18	ASD	NV[c]	48[d]

Note. Schools and paraeducators are numbered consistently in this fashion for ease of cross-referencing throughout manuscript. Only primary educational labels are reported. ASD = autism spectrum disorder; ID = intellectual disability; OHI = other heath impairment; NR = not reported; NV = no valid score.

[a]Weschler Intelligence Scale for Children–Fifth Edition.

[b]Vineland Adaptive Behavior Scales–Third Edition.

[c]Comprehensive Test of Nonverbal Intelligence–Second Edition.

[d]Adaptive Behavior Assessment System–Third Edition.

[e]Naglieri Nonverbal Ability Test.

Table 2. Targeted and Generalization Instructional Goals for Paraeducators.

School number	Paraprofessional number	Target instructional goal	Generalization instructional goal
1	1	Blending four-letter words	Letter identification
	2	Spelling	Sight words
2	1	Counting from array	Blending three-letter words
	2	Sight words	Blending four-letter words
	3	Sight words	Blending three-letter words
3	1	Spelling[a]	Sight words[a]
	2	Sight words	Spelling
	3	Sight words	Spelling
4	1	Identifying greater number	Identifying coins and coin values
	2	Science vocabulary[b]	Counting money (bills)
	3	Money-words problems	Nutrition vocabulary[b]
5	1	Identifying emotions	Folding T-shirt
	2	Identifying greater number	Coin identification

[a]The student mastered the original teaching targets partway through the study, and new targets within the same category were selected.
[b]Special education teachers worked with general education teachers to identify new vocabulary that would be used in upcoming units in general education classrooms where students were included.

consulted with general educators to select instructional targets from the general education curriculum that were relevant to upcoming instruction in the general education classroom. Next, we used a computerized random-number generator to randomly determine which goal would be the target goal (i.e., paraeducators would receive performance feedback on their implementation of prompting to teach this goal during group training) and the generalization goal (i.e., paraeducators would not receive any feedback during group training related to teaching this goal). Target and generalization goals are reported for each student in Table 2.

In one 50-min training session with paraeducators, the first author trained paraeducators to differentiate between discrete and chained skills and to conduct a task analysis of chained skills. Training strategies included (a) sharing written descriptions and examples of discrete skills, chained skills, and task analyses; (b) providing didactic instruction on these concepts; and (c) supporting paraeducators to record task analyses of chained skills they were teaching, including feedback as appropriate. At the end of this prebaseline condition, task analyses were completed for all chained skills (both target and generalization). Having complete task analyses is a prerequisite to implementing systematic prompting (Collins, 2012; Neitzel & Wolery, 2009a, 2009b).

Written Directions (Baseline). We distributed written directions for how to implement both simultaneous prompting and least-to-most prompting. This included implementation checklists of all steps associated with each procedure. We distributed these checklists so that we would have a baseline in which paraeducators would have a conceptual understanding of the procedures we were asking them to perform but would not have yet received the experimental tiered training. In addition, we could isolate the effects of the tiered training above and beyond the effects of simply providing written directions. Based on findings in previous studies, written directions alone have limited or no effect on paraeducator implementation (Brock, Barczak, et al., 2019).

Group Training. The first author provided three 50-min didactic training sessions over 3 weeks. In two schools (Schools 3 and 4), teachers indicated that it would not be possible for all paraeducators to meet during the same uninterrupted 50-min period, so we split the single 50-min weekly session into two 25-min training sessions across two consecutive days. The sequence of topics built on one another while incorporating feedback on any previously taught strategies.

In the first week, training focused on simultaneous prompting. After distributing a workbook that included didactic materials, implementation checklists, and worksheets, the first author (a) provided a rationale for simultaneous prompting, (b) explained the steps of simultaneous prompting and provided examples, (c) modeled all implementation steps, (d) directed paraeducators to practice with each other through role-play, and (e) provided feedback to paraeducators during the role-play by praising steps that were followed with fidelity and offering constructive suggestions for fixing mistakes. At the end of the session, the first author left a tablet computer (i.e., Amazon Fire) and mini tripod in the classroom and directed paraeducators to video record their implementation of simultaneous prompting with the target student so that they could share it with the group at the next training session. Paraeducators were instructed to video record their implementation that focused on the (randomly selected) target instructional goal. Instruction that targeted the generalization goal was not video recorded, and the trainer did not provide any modeling or feedback that focused on the generalization goal during group training.

In the second week, training focused on feedback on the video recording of simultaneous prompting and an introduction to least-to-most prompting. First, the first author briefly reviewed the implementation steps associated with simultaneous prompting and explained that these would be the basis for feedback. Each paraeducator shared their video with the group. After each video, the first author invited feedback from the group and provided additional constructive feedback if it was not offered by peers. Next, the first author introduced least-to-most prompting in the same fashion as described before for simultaneous prompting. The first author instructed the paraeducators to collect data on student performance during least-to-most prompting using the data sheets provided by the special education teacher and modeled how to complete data collection. If the special educator did not have data sheets, we developed data sheets for them. At the end of the session, the first author directed paraeducators to video record their implementation of both simultaneous prompting and least-to-most prompting.

The third and final week of group training focused on feedback on both procedures. Video feedback was provided in the same fashion as described for simultaneous prompting.

Maintenance. No additional training was provided in the maintenance condition. Procedures were identical to the baseline condition in which paraeducators had access to written directions. The maintenance condition began immediately after the group training condition and continued until all procedures were implemented in the last tier. The longest time between the end of training and a maintenance probe was 46 school days (i.e., School 2 Paraeducators 1 and 2).

Coaching. If a paraeducator failed to demonstrate criterion-level (>85%) fidelity for two consecutive sessions for either procedure (i.e., simultaneous prompting or least-to-most prompting) for either student goal (i.e., target or generalization) during the maintenance condition, the special education teacher delivered two 15-min one-to-one coaching sessions. Coaching focused on implementation of both strategies for the student goal for which the paraeducator did not demonstrate criterion-level fidelity. Coaching sessions involved watching video recordings of both procedures for both student goals. Teachers provided praise for steps followed with fidelity and constructive suggestions for fixing mistakes. The first author was present for all coaching sessions. Prior to the coaching condition, the first author provided each teacher with a 45-min training session that focused on (a) implementation steps of simultaneous prompting and least-to-most prompting and (b) how to provide positive and

constructive feedback on all implementation steps. Unlike the group training, coaching focused on feedback related to both targeted and generalization situations.

Dependent Measures and Recording

Classroom Observations. We observed paraeducators in their classrooms as they were delivering instruction to target students. Observations occurred at the time of day when the paraeducators would naturally be working with target students and were conducted at a separate time that was distinct from any training sessions. All observations involved the paraeducator delivering instruction focused on the preselected goals from the student's IEP and then the student being probed on the goal. Observations were conducted in the same fashion across all experimental conditions. We observed paraeducators delivering both simultaneous prompting and least-to-most prompting for both student goals. Directions and measurement were identical for both targeted and generalization situations. These directions were designed to teach the paraeducator to pair the procedures with the situations in which they were most appropriate. First we directed the paraeducator to implement simultaneous prompting as if they were targeting the skill for the first time. Next we directed the paraeducator to implement least-to-most prompting as if they had already targeted the skill for several days and they anticipated fewer student errors (Collins, 2012). Last, we directed the paraeducator to deliver a student probe to gauge student progress. We directed the paraeducator to withhold any prompts during this probe so that we could gauge independent student performance. We measured implementation fidelity of systematic prompting in two ways.

Adherence to steps. First, we measured adherence to implementation steps for simultaneous prompting and least-to-most prompting using implementation checklists. Steps for simultaneous prompting included delivering a prompt within 1 s of providing a cue or task direction, delivering the same topography of controlling prompt in each trial, delivering specific praise after correct responses, and repeating a trial after incorrect responses. Steps for least-to-most prompting included providing 3 to 5 s for the student to respond independently, delivering specific praise after correct responses, and delivering prompts of increasing intensity after incorrect responses. These steps were based on implementation checklists developed by the National Professional Development Center on Autism Spectrum Disorders (Neitzel & Wolery 2009a, 2009b). We calculated the percentage of steps implemented correctly.

Quality of implementation. We measured implementation quality as a second, descriptive dimension of implementation fidelity. We scored six items on a 4-point Likert-type scale with 3 being the highest quality and 0 being the poorest quality. Items focused on pacing of instruction, consistency of delivering the same cue or task direction, immediacy of prompting after an error, verbal praise, immediacy of reinforcement after a correct response, and overall quality of implementation. We computed an average score across items as a measure of overall quality. The possible range of this average was 0 to 3.0. The complete quality instrument is published in Brock, Barcak et al., 2019.

Student Progress. At the end of each classroom observation, we administered probes to students in which they had the opportunity to provide 10 discrete responses or 10 chained responses associated with the target and generalization goals described in Table 2. We calculated the percentage of correct responses.

Social Validity Survey. We used a 10-item paper-pencil questionnaire to measure paraeducator perceptions of their own competence, the training, and their likelihood to use systematic instruction and participate in similar training opportunities in the future. Responses were provided on a 5-point Likert-type scale. Both the questions and the scaling are reported in Table 3.

Observer Training and Interobserver Agreement

The second author, a graduate student in special education, was the primary data collector for this study. She was already trained by the first author on the same observation protocol in a prior study. The second author trained

Table 3. Social Validity Questionnaire Ratings by Participants.

Category and question	*M*	*SD*
Paraeducator feedback		
Perception of skill after training		
How skilled are you in implementing simultaneous prompting?	4.07	0.62
How skilled are you in implementing least-to-most prompting?	4.21	0.58
How skilled are you in data collection?	3.93	0.62
Perception of training		
How effective was the training at enabling you to implement new strategies with your student?	4.14	0.66
How much do you think that the new strategies you learned contributed to your student's progress on his/her goal?	3.93	0.73
Likelihood of future implementation and training		
How likely would you be to continue to use the strategies that you learned in the future with the same student?	4.50	0.65
How likely would you be to use the strategies that you learned in the future with a different student?	4.50	0.65
How likely would you be to participate in a similar training opportunity in the future?	4.07	0.92
How likely would you be to recommend a similar training opportunity to a colleague?	3.93	1.07
Teacher feedback		
Providing feedback to paraeducators		
How difficult was it to provide video-based feedback to your paraeducators?	1.33	0.52
How confident are you that you could provide video-based feedback to paraeducators on your own?	3.50	1.52
How willing would you be to provide video-based feedback to your paraeducators in the future?	4.00	0.58

Note. Response options included 1 = *not at all*, 2 = *slightly*, 3 = *somewhat*, 4 = *moderately*, and 5 = *extremely*.

five additional observers: two graduate students and three undergraduate students. Before collecting data, these students were required to (a) review the coding manual, (b) score 100% on a written test of coding definitions, (c) achieve at least 95% agreement with the second author when coding a training video, and (d) achieve at least 95% agreement with the second author in a live observation.

Two observers collected data during 26% of all observations across participants and conditions. We computed point-by-point agreement for each behavior. Average overall agreement across paraeducator implementation behavior was 96% (range = 67%–100%) and student behavior was 99% (range = 80%–100%).

Procedural Fidelity

During the written-directions condition, all paraeducators reported reading the written directions. During the group training sessions, the first author used a written checklist to ensure that he (a) provided a rationale for each strategy, (b) provided didactic instruction on each step of the implementation checklist, (c) modeled each implementation step, (d) demonstrated exemplars and common errors, and (e) provided feedback to paraeducators during the role-play by praising steps that were followed with fidelity and offering constructive suggestions for fixing mistakes, (f) invited peer feedback during video-based feedback, (g) provided feedback on videos by praising steps that were followed with fidelity and offering constructive suggestions for fixing mistakes, and (h) modeled implementation of any incorrectly implemented steps. These steps were followed with 100% fidelity for both prompting strategies. The second author independently scored the same procedural checklist for 20% of all training sessions and had 100% agreement with the first author.

During coaching sessions, the special education teachers used a written checklist to ensure that they delivered either positive or corrective feedback for every implementation step associated with each procedure. These steps were followed with 100% fidelity based on scoring by the first author, who attended all coaching sessions.

Results

Next, we report on results of (a) paraeducator implementation of systematic instructional strategies, (b) student performance on individualized goals that were targeted by paraeducators, and (c) paraeducator and teacher perceptions of the training.

Paraeducator Implementation of Systematic Instructional Strategies

We measured two dimensions of implementation fidelity: (a) adherence to steps and (b) implementation quality. We report results from each, including our visual analysis of the data in terms of level, trend, variability, and immediacy of effects. Adherence to steps for both simultaneous prompting and least-to-most prompting across targeted and generalized situations is graphed in Figure 1. Implementation quality is described in Table 4. On the basis of our visual analysis, we concluded that there is a functional relation between the tiered training and adherence to steps for both simultaneous prompting and least-to-most prompting in both targeted and generalization situations. All 13 paraeducators acquired criterion levels of fidelity after all training was complete; nine of them 13 acquired and maintained criterion-level performance with group training alone, whereas four required coaching to maintain performance at criterion in targeted or generalization situations. Specifically, School 1 Paraeducator 1 required coaching for both simultaneous prompting and least-to-most prompting for the generalization situation, School 1 Paraeducator 2 required coaching for only least-to-most prompting in the generalization situation, School 2 Paraeducator 3 required coaching for both simultaneous prompting and least-to-most prompting in the targeted situation, and School 4 Paraeducator 2 required coaching for simultaneous prompting in the generalization situation.

Adherence to Steps for Simultaneous Prompting. We measured adherence to steps for simultaneous prompting in both targeted situations (i.e., trainer provided feedback on performance in this situation during group training) and generalization situations (i.e., trainer did not provide feedback on performance in this situation during group training).

Targeted training situations. The percentage of steps of simultaneous prompting implemented with fidelity in targeted training situations is represented in filled circles in Figure 1. During the baseline condition, the levels of the data were low, with high variability across all paraeducators. Twelve of the 13 paraeducators had at least one session with 0% steps implemented with fidelity; the lowest performance for Paraeducator 2 in School 4 was 25%. Although baseline performance was variable, in no case was performance consistently trending toward improved implementation fidelity. Immediately after introduction of the group training condition, data increased to criterion level (i.e., >85%) across all 13 paraeducators. For 12 of the 13 paraeducators, this level maintained in the maintenance condition. For the remaining paraeducator (i.e., School 2 Paraeducator 3), consecutive sessions below criterion triggered coaching. After coaching, this paraeducator demonstrated criterion-level performance.

Generalization situations. The percentage of steps of simultaneous prompting with fidelity in generalized situations is represented with open circles in Figure 1. During the baseline condition, the levels of data were low, with high variability across all paraeducators. All 13 paraeducators had at least one session with 0% steps implemented with fidelity. Although baseline performance was variable, in no case was performance consistently trending toward improved implementation fidelity. Immediately after introduction of the group training

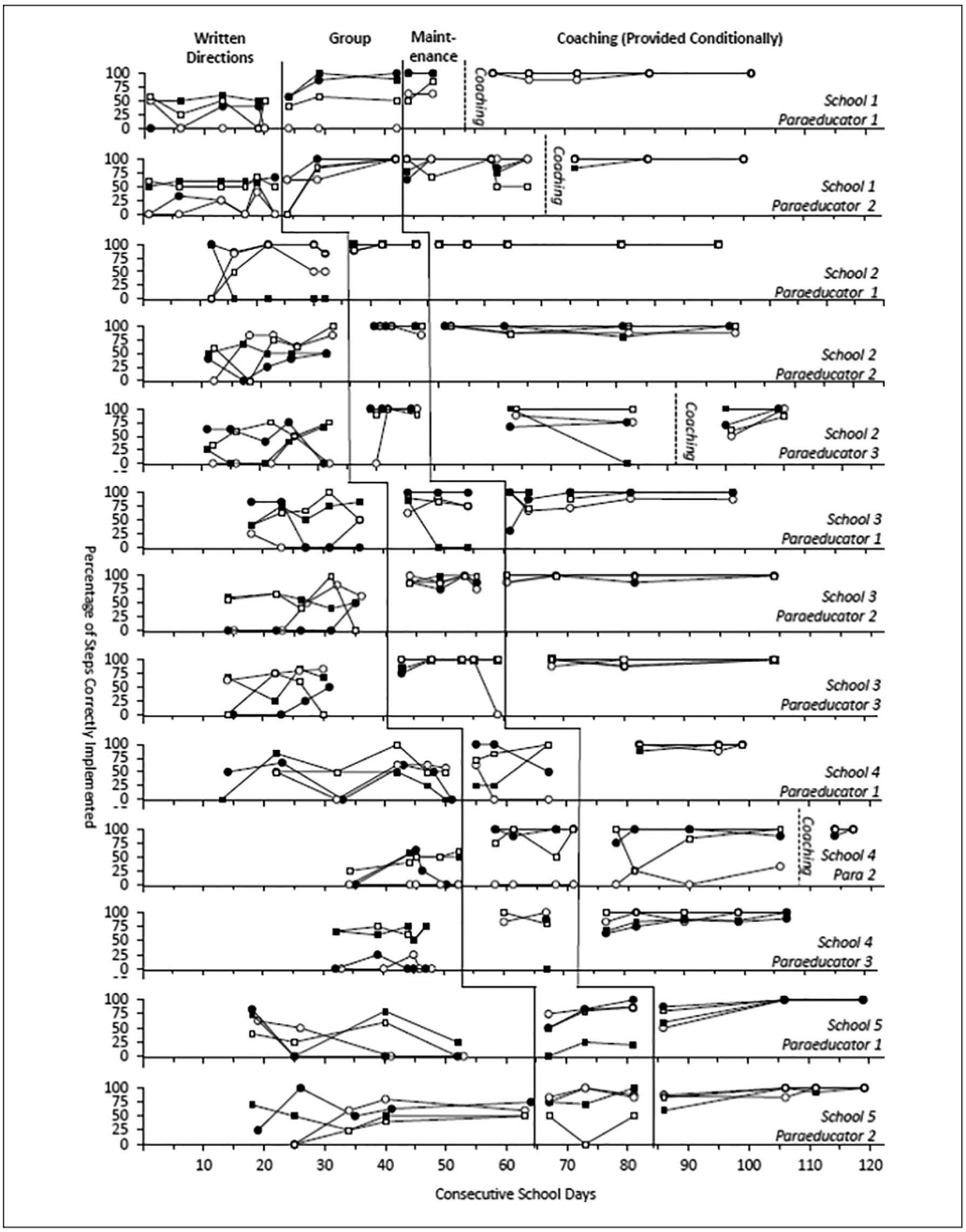

Figure 1. Effects of tiered training on paraeducator implementation of systematic prompting strategies. Implementation fidelity of simultaneous prompting (circles) and least-to-most prompting (squares) given written directions, group training, and one-to-one coaching. Introduction of one-to-one coaching occurred only if a paraeducator did not implement a practice with criterion-level fidelity after group training (see dotted lines on graph for the four participants who received coaching). Filled shapes represent situations in which paraeducators received direct feedback during group training. Open shapes represent generalization situations in which paraeducators did not receive direct feedback prior to coaching. One paraeducator (School 2 Paraeducator 3) received coaching on the targeted situation, and three paraeducators received coaching on the generalization situation (School 1 Paraeducator 1, School 1 Paraeducator 2, and School 4 Paraeducator 2).

Table 4. Average Overall Quality of Implementation Scores for Paraeducators Across All Experimental Conditions.

School number	Paraeducator number	Written directions	Group training	Maintenance	Coaching
1	1	2.0 (1.5–2.5)	2.4 (2.0–2.8)	2.3 (2.0–2.7)	2.4 (2.2–2.7)
	2	2.2 (1.5–2.7)	2.4 (1.7–2.4)	2.4 (2.2–2.8)	3.0 (3.0–3.0)
2	1	2.6 (2.0–2.8)	3.0 (3.0–3.0)	3.0 (3.0–3.0)	—
	2	2.2 (1.5–2.5)	3.0 (3.0–3.0)	3.0 (3.0–3.0)	—
	3	1.9 (1.2–2.7)	2.3 (1.3–3.0)	2.6 (1.7–2.8)	2.8 (2.2–3.0)
3	1	2.3 (2.0–2.6)	2.7 (2.0–3.0)	2.6 (1.7–3.0)	2.8 (2.5–3.0)
	2	2.3 (1.8–2.5)	2.5 (1.9–2.8)	2.7 (2.0–2.8)	—
	3	2.2 (1.8–2.8)	2.1 (2.0–2.8)	3.0 (3.0–3.0)	—
4	1	1.9 (1.3–2.2)	3.0 (3.0–3.0)	3.0 (3.0–3.0)	—
	2	1.8 (1.5–2.2)	2.6 (1.7–2.8)	2.6 (2.2–3.0)	2.7 (1.6–3.0)
	3	1.9 (1.8–2.0)	2.2 (1.9–2.4)	2.4 (1.5–2.9)	—
5	1	2.0 (1.7–2.5)	2.8 (2.1–3.0)	3.0 (3.0–3.0)	—
	2	2.2 (1.7–2.5)	2.7 (2.2–3.0)	2.6 (2.4–3.0)	—

Note. The ranges of overall scores across all sessions within the phase are reported in parentheses. Overall scores represent an average of the six items described in Table 1. Scores range from 0 (*poor quality*) to 3 (*high quality*).

condition, data increased to criterion level (i.e., >85%) across all 11 of the 13 paraeducators. Two paraeducators (School 1 Paraeducators 1 and 2) had consecutive session below criterion, which triggered coaching. After coaching, both paraeducators demonstrated criterion level-performance.

Adherence to Steps for Least-to-Most Prompting. We measured adherence to steps for least-to-most prompting both in targeted situations and generalization situations.

Targeted training situations. The percentage of steps of least-to-most prompting implemented with fidelity in targeted training situations is represented with filled squares in Figure 1. During the baseline condition, the levels of data were low or medium, with high variability across all paraeducators. Five of the 13 paraeducators had at least one session with 0% steps implemented with fidelity; the other eight paraeducators had at least one session with 50% or less steps implemented with fidelity. Variability was very high, with one paraeducator having sessions at both 0% and 100% (i.e., School 2 Paraeducator 2). Although baseline performance was variable, in no case was performance consistently trending toward improved implementation fidelity. After introduction of the group training condition, data trended to criterion level (i.e., >85%) within 3 data points across all 13 paraeducators. We anticipated a delay in effect because the least-to-most prompting was not covered in detail until the second week of training. For 12 of the 13 paraeducators, this level maintained in the maintenance condition. For the remaining paraeducator (i.e., School 2 Paraeducator 3), consecutive sessions below criterion triggered coaching. After coaching, this paraeducator demonstrated criterion-level performance.

Generalization situations. The percentage of steps of least-to-most prompting with fidelity in generalized situations is represented with open circles in Figure 1. During the baseline condition, the levels of data were low or medium, with high variability across all paraeducators. Six of the 13 paraeducators had at least one session with 0% steps implemented with fidelity, and the other 11 had at least one session with 50% or less of steps implemented with fidelity. In 10 of 13 cases, performance was highly variable and did not trend toward improved implementation fidelity. In three cases (School 2 Paraeducator 1, School 2 Paraeducator 2, and School 5 Paraeducator 2), implementation did trend toward criterion during baseline; in these three cases, it is not

possible to establish an experimental effect. After the introduction of group training, 11 paraeducators implemented with criterion-level performance (i.e., >85%). Two paraeducators (School 1 Paraeducators 1 and School 5 Paraeducator 2) had consecutive session below criterion, which triggered coaching. After coaching, both paraeducators demonstrated criterion level-performance.

Implementation Quality. Quality was measured as a secondary, descriptive indicator of implementation fidelity. The measure spans from 0 to 3. Mean quality improved from 2.0 to 2.7 across all paraeducators between the written directions and maintenance conditions. Mean quality and range for each paraeducator are reported by experimental condition in Table 4. Each paraeducator improved their implementation quality between the written directions and maintenance conditions.

Student Progress on Individualized Goals That Were Targeted by Paraeducators

Student performance was measured as a secondary, descriptive variable. Student performance in situations that were directly targeted in paraeducator group training (filled triangles) and in generalization situations (open triangles) are graphed in Figure 2. Because we manipulated the independent variable based on paraeducator implementation fidelity data, there are cases in which we changed conditions when student data were displaying therapeutic trends. In addition, one student (School 3 Student 1) demonstrated mastery of both of his original goals, so we directed the paraeducator to move on to more difficult teaching targets in the maintenance condition. In these situations, it is not possible to conclude that experimental effects were demonstrated.

Therefore, we describe these data but make no claims of causal inference. Nine students met the performance criterion (i.e., >80% accuracy) in the targeted training situation and nine in the generalization situation. The only students who did not meet the performance criterion for either goal were at School 5, where paraeducators implemented the practices with fidelity for the shortest period of time before the end of the study.

When we conducted visual analysis of student data, we identified 14 cases in which therapeutic baseline trends did not enable us to make a judgement of effect, eight cases in which we detected an effect, and four cases with stable baseline data in which there was no evidence of an effect. The 14 cases with therapeutic baseline trends included School 2 Paraeducator 1 (both goals), School 2 Paraeducator 2 (primary), School 2 Paraeducator 3 (generalization), School 3 Paraeducator 1 (both), School 3 Paraeducator 3 (both), School 4 Paraeducator 1 (primary), School 4 Paraeducator 2 (both), School 4 Paraeducator 3 (both), and School 5 Paraeducator 1 (generalization). The eight cases in which we detected an effect included School 1 Paraeducator 1 (both), School 1 Paraeducator 2 (generalization), School 2 Paraeducator 2 (primary), School 2 Paraeducator 3 (primary), School 3 Paraeducator 2 (primary), School 4 Paraeducator 1 (generalization), and School 5 Paraeducator 2 (generalization). The four cases in which there were stable baseline trends and no evidence of an effect included School 1 Paraeducator 2 (primary), School 3 Paraeducator 2 (generalization), School 5 Paraeducator 1 (primary), and School 5 Paraeducator 2 (primary).

Teacher and Paraeducator Perceptions of the Training

Mean scores and standard deviations for all items on the paraeducator and teacher questionnaires are reported in Table 3. Paraeducators reported that after the training, they felt moderately skilled at implementing simultaneous prompting, least-to-most prompting, and data collection. Paraeducators perceived the training as moderately effective in enabling them to implement new strategies with their student and felt that they strategies moderately contributed to student progress. They reported that they would be extremely likely to continue using the strategies with the same student and with different students. They reported that

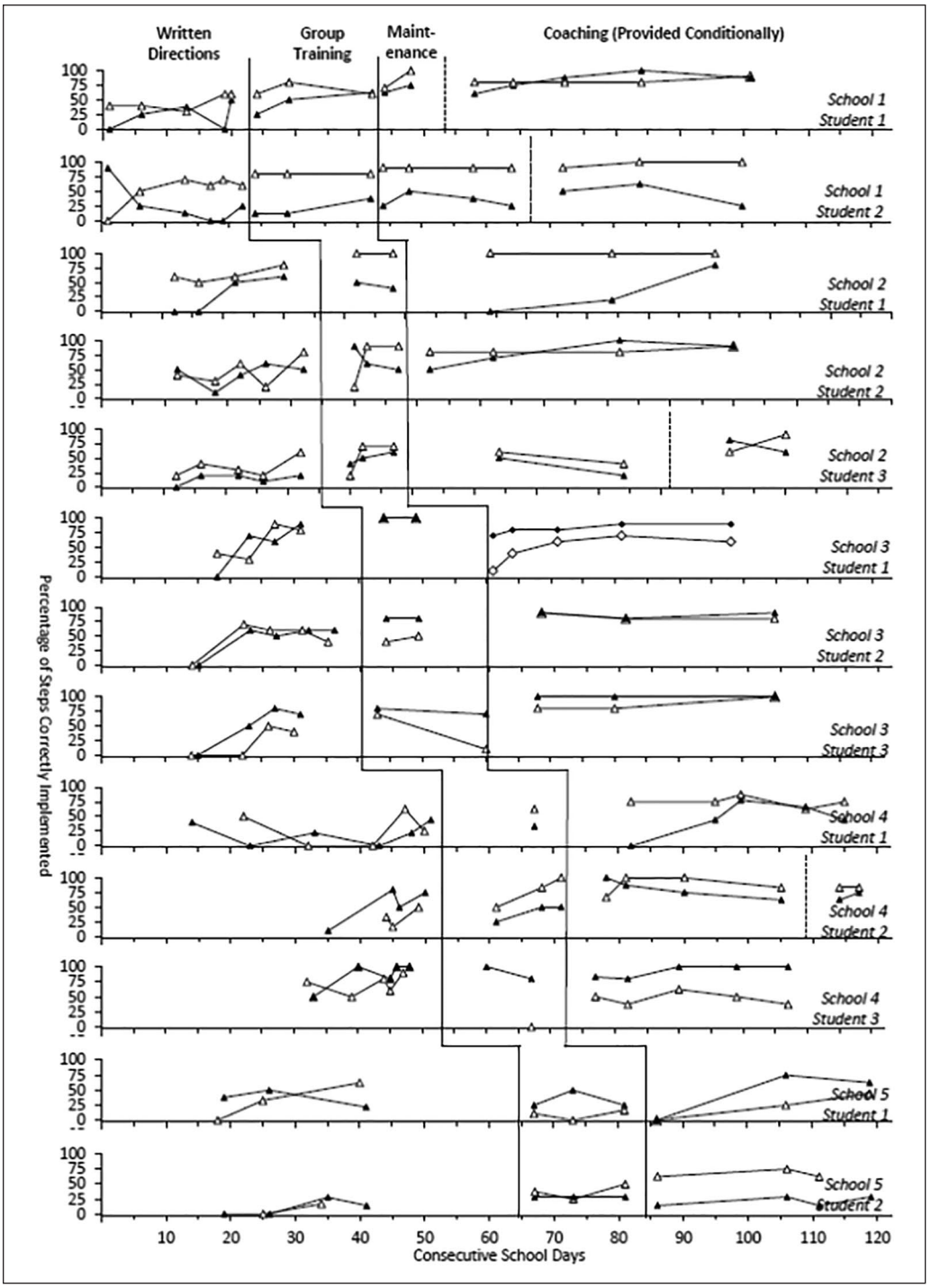

Figure 2. Student progress on individualized goals targeted by paraeducator-implemented instruction. Student performance on goals targeted by paraeducators. Filled shapes represent situations in which paraeducators received direct feedback during group training. Open shapes represent generalization situations in which paraeducators did not receive direct feedback. Diamond shapes (see maintenance condition for School 3 Student 1) signify that new instructional targets were provided because the student demonstrated mastery of original targets.

they would be moderately likely to participate in a similar training opportunity in the future and to recommend a similar opportunity to a colleague.

Teachers reported that the video-based feedback was not at all difficult to provide, that they were somewhat-to-moderately confident that they could provide video-based feedback on their own, and that they would be moderately willing to provide video-based feedback to paraeducators in the future.

Discussion

Students with severe disabilities deserve to be taught with evidence-based practices, but they often receive instruction from paraeducators who are not trained in these practices. Previous research has led to identification of training and supervision strategies that enable paraeducators to implement evidence-based practices with fidelity but rely heavily on resource-intensive one-to-one coaching and may not enable paraeducators to generalize instructional strategies to new situations. In the current study, we tested a tiered training model that involves group training for all paraeducators and one-to-one coaching only in the subset of cases that a paraeducator is struggling to implement with fidelity after the group training. We found that tiered training was effective in enabling paraeducators to implement with fidelity, most paraeducators were able to generalize their implementation to new situations with only group training, and most students made progress on individualized goals after receiving instruction from paraeducators who had been trained. These findings add to the existing literature in a number of key ways.

First, tiered training is an effective approach for training paraeducators to implement systematic instructional practices with fidelity while reducing reliance on one-to-one coaching. This is consistent with emerging evidence for tiered models of professional development for teachers (Gage et al., 2017). Although many studies have demonstrated the efficacy of one-to-one coaching for paraeducators, this is the first published study to show the efficacy of a tiered model where coaching is used contingent on paraeducator performance. In only one case did a paraeducator require one-to-one coaching to meet the performance criterion for the teaching situations that were directly targeted in the group training. As would be expected based on previous findings, coaching was an effective tool for enabling that paraeducator to acquire implementation fidelity.

tiered training is an effective approach for training paraeducators to implement systematic instructional practices with fidelity while reducing reliance on one-to-one coaching.

Second, a tiered approach enabled some paraeducators to generalize their implementation across different teaching situations. In nine of 13 cases, paraeducators were able to implement systematic prompting strategies in new teaching situations without any additional training or feedback. Generalization is a challenge for all learners, and researchers have previously documented the challenge of enabling teachers or paraeducators to generalize a teaching strategy to a new situation (Brock, Barczak, et al., 2019). The approach for promoting generalization in this study was multiple exemplars (Stokes & Baer, 1977). Specifically, in the group training, paraeducators observed correct implementation in each context that their colleague implemented. They observed either (a) correct implementation in their colleague's video (i.e., if there were no mistakes) or (b) trainer modeling of correct implementation during corrective feedback. Although observing a colleague receive feedback does not necessarily enable paraeducators to initially acquire implementation fidelity (Brock, Barczak, et al. 2019), viewing multiple exemplars may be an approach that promotes generalization to new situations after implementation fidelity has already been initially acquired.

Third, prior to group training, paraeducator-delivered instruction in these schools was inconsistent and lacked any systematic approach. Our baseline condition involved giving paraeducators written directions so that

we could establish their implementation after having some basic familiarity with the prompting procedures we were targeting. Data across all paraeducators were highly variable, suggesting that they were changing their instructional approaches over time. This inconsistency stands in stark contrast to evidence-based, systematic approaches. Systematic instruction enables teachers to analyze whether a given instructional approach is effective and whether instructional changes are warranted. It is not possible to make valid data-based decisions when a teacher believes a paraeducator is consistently implementing a given instructional approach but in fact their implementation fidelity is highly variable. Instead, teachers must ensure that the intervention is implemented as intended before they make a judgment about the effect of the intervention on student progress.

Fourth, students with severe disabilities made progress when paraeducators implemented systematic prompting strategies. The student data in this study were descriptive and simply show that students made progress over time as they received instruction from paraeducators. However, when findings from this study are contextualized within the broader literature that demonstrates robust student-level effects for systematic prompting (cf. Spooner et al., 2019; Wong et al., 2015), one can conclude that students with severe disabilities benefit when paraeducators deliver these evidence-based practices with fidelity.

students with severe disabilities benefit when paraeducators deliver these evidence-based practices with fidelity.

Implications for Practice

Findings from this study point to implications for administrators and faculty members in teacher training programs. Administrators should ensure that teachers have the time needed to adequately train and supervise paraeducators. Our findings illustrate that this can be done in a limited amount of time. Specifically, after reading a four-page handout, paraeducators received 2.5 total hours of training time (i.e., three 50-min training sessions), with only a few paraeducators requiring an additional 30 min of coaching (i.e., two 15-min coaching sessions). However, at several of our training sites, coordinating a meeting time required either paraeducators staying before or after school beyond the time that they are paid to work, splitting training sessions between 2 days, and/or scheduling at times that staff from other classrooms could be reallocated to support students. We recommend that in practice, administrators help teachers schedule time that they can meet with paraeducators during the school day or ensure that paraeducators are paid for any training time that falls outside of their regular work hours. Furthermore, we recommend that administrators mentor their teachers so that they can more confidently take on the role of supervisor and instructional leader.

Faculty members in teacher training programs should ensure that future teachers are well prepared to train and supervise paraeducators. Teachers reported that prior to this study they had not provided frequent feedback and paraeducators had not received any focused training on instructional practices. This aligns with previously reported findings that teachers are not well prepared to train and supervise paraeducators (Wermer et al., 2018). Findings from this study illustrate an effective and feasible model that teachers could be trained to implement with their paraeducators.

Limitations and Future Directions for Research

Limitations from our study provide opportunities for future directions for research. First, we were able to measure generalization for paraeducator implementation during group training but not during one-to-one coaching. Specifically, it is unclear if, after receiving one-to-one coaching focused on the second teaching situation, paraeducators could have independently generalized to a third teaching situation. In future studies, researchers might include multiple contexts for generalization to address this issue. Second, we trained teams

of two or three paraeducators, but descriptive studies have reported that teachers sometimes supervise even larger numbers of paraeducators (Suter & Giangreco, 2009). Researchers might replicate our training model with larger teams. Third, we treated student performance as a secondary descriptive variable and made decisions about changing conditions based solely on paraeducator implementation fidelity. Therefore, there are instances in which it is not possible to demonstrate student effects (e.g., unstable student baseline data or a change in student targets). This means we can present student data only descriptively but cannot make strong claims about experimental effects of paraeducator implementation on student progress. Fourth, there were multiple components to the group training so we cannot isolate effects related to specific components (e.g., verbal instruction, video-based feedback). In future studies, researchers might conduct a component analysis to disentangle these effects. Fifth, in some cases, we were unable to complete scheduled observations because students or paraeducators were absent. For example, for School 5 Paraeducator 1, we were not able to collect data on Day 63. Sixth, like any study in which live observations are conducted, it is likely that the presence of observers served as a prompt for paraeducators to implement practices they might not have otherwise implemented. In future studies, researchers might attempt to observe maintenance of implementation without the presence of an observer from the research team (e.g., report from supervising teacher, covert video recording). Finally, baseline observations did not begin on the same day across tiers. In future studies, researchers should attempt to begin data collection at the same time across all tiers.

Conclusion

One-to-one coaching is not feasible as a standalone model for training all paraeducators who work with students with severe disabilities to implement evidence-based instruction. In this study, we demonstrated that in a tiered training model, most paraeducators acquired and generalized evidence-based practices with group training alone, and the remainder did so after teacher-delivered coaching. Although further replication is needed, this evidence provides initial promise that tiered training may present a feasible approach for scaling up research-based paraeducator training.

References

Brock, M. E., & Anderson, E. J. (2020). Training paraprofessionals who work with students with intellectual and developmental disabilities: What does the research say? *Psychology in the Schools*. Advance online publication. http://dx.doi.org/10.1002/pits.22386

Brock, M. E., Barczak, M. A., & Dueker, S. A. (2019). Effects of delayed video-based feedback and observing feedback on paraprofessional implementation of evidence-based practices for students with severe disabilities. *Focus on Autism and Other Developmental Disabilities*. Advance online publication. https://doi.org/10.1177/1088357620902492

Brock, M. E., Cannella-Malone, H. I., Seaman, R. L., Andzik, N. R., Schaefer, J. M., Page., E. J., Barczak, M., & Dueker, S. (2017). Findings across practitioner training studies in special education: A comprehensive review and meta-analysis. *Exceptional Children*, *84*, 7–26. https://doi.org/10.1177/0014402917698008

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*, 211–221. https://doi.org/10.1177/154079691303800401

Brock, M. E., Dynia, J. M., Dueker, S., & Barczak, M. (2019). Teacher-reported priorities and practices for students with autism: characterizing the research-to-practice gap. *Focus on Autism and Other Developmental Disabilities*. Advance online publication https://doi.org/10.1177/1088357619881217

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. https://doi.org/10.1177/0741932508324399

Collins, B. C. (2012). *Systematic instruction for students with moderate and severe disabilities*. Brookes.

Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015–2016).

Gage, N. A., MacSuga-Gage, A. S., & Crews, E. (2017). Increasing teachers' use of behavior-specific praise using a multitiered system for professional development. *Journal of Positive Behavior Interventions*, *19*, 239–251. https://doi.org/ 0.1177/1098300717693568

Gast, D. L., Lloyd, B. P., & Ledford, J. R. (2018). Multiple baseline and multiple probe designs. In J. R. Ledford, & D. L. Gast (Eds.). *Single case research methodology: Applications in special education and behavioral sciences* (pp. 239–282). Routledge.

Hess, K. L., Morrier, M. J., Heflin, L. J., & Ivey, M. L. (2008). Autism treatment survey: Services received by children with autism spectrum disorders in public school classrooms. *Journal of Autism and Developmental Disorders*, *38*, 961–971. https://doi.org/10.1007/s10803-007-0470-5

Individuals With Disabilities Education Improvement Act, U.S.C. §§ 1400 et seq. (2006 & Supp. V. 2011)

Kratochwill, T. R., & Levin, J. R. (2014). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods*, *15*, 124–144. https://doi.org/10.1037/a0017736

Neitzel, J., & Wolery, M. (2009a). *Implementation checklist for least-to-most prompts*. National Professional Development Center on Autism Spectrum Disorders, FPG Child Development Institute, University of North Carolina.

Neitzel, J., & Wolery, M. (2009b). *Implementation checklist for simultaneous prompting*. National Professional Development Center on Autism Spectrum Disorders, FPG Child Development Institute, University of North Carolina.

Neitzel, J., & Wolery, M. (2009c). *Overview of prompting*. National Professional Development Center on Autism Spectrum Disorders, FPG Child Development Institute, University of North Carolina.

Odom, S. L., Hall, L. J., & Steinbrenner, J. R. (2019). Implementation science research and special education. *Exceptional Children*. Advance online publication. https://doi.org/10.1177/0014402919889888

Rusch, F. R., Hughes, C., Agran, M., Martin, J. E., & Johnson, J. R. (2009). Toward self-directed learning, post-high school placement, and coordinated support constructing new transition bridges to adult life. *Career Development for Exceptional Individuals*, *32*, 53–59. https://doi.org/10.1177/0885728809332628

Russo, A. (2004). School-based coaching. *Harvard Education Letter*, *20*, 1–4.

Shepley, C., Lane, J. D., & Ault, M. J. (2019). A review and critical examination of the system of least prompts. *Remedial and Special Education*, *40*, 313–327. https://doi.org/10.1177/0741932517751213

Spooner, F., Root, J. R., Saunders, A. F., & Browder, D. M. (2019). An updated evidence-based practice review on teaching mathematics to students with moderate and severe developmental disabilities. *Remedial and Special Education*, *40*, 150–165. https://doi.org/10.1177/0741932517751055

Stahmer, A. C., Collings, N. M., & Palinkas, L. A. (2005). Early intervention practices for children with autism: Descriptions from community providers. *Focus on Autism and Other Developmental Disabilities*, *20*, 66–79. https://doi.org/10.1177/10883576050200020301

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *10*, 349–367.

Suter, J. C., & Giangreco, M. F. (2009). Numbers that count: Exploring special education and paraprofessional service delivery in inclusion-oriented schools. *The Journal of Special Education*, *43*, 81–93. https://doi.org/10.1177/0022466907313353

U.S. Department of Education. (2019). *IDEA Section 618 data products*. https://www2.ed.gov/programs/osepidea/618-data/index.html

Wermer, L., Brock, M. E., & Seaman, R. L. (2018). Efficacy of a teacher coaching a paraprofessional to promote communication for a student with autism and complex communication needs. *Focus on Autism and Other Developmental Disabilities*, *33*, 217–226. https://doi.org/10.1177/1088357617736052

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., Brock, M. E., Plavnick, J. B., Fluery, V. P., & Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum

disorder: A comprehensive review. *Journal of Autism and Developmental Disorders*, *45*, 1951–1966. https://doi.org/10.1007/s10803-014-2351-z

Authors' Note

The research reported here was supported, in whole or in part, by the Institute of Education Sciences (IES), U.S. Department of Education, through Grant R324B160009 to The Ohio State University. The opinions expressed are those of the authors and do not represent the views of the IES or the U.S. Department of Education.

ORCID iD

Matthew E. Brock https://orcid.org/0000-0001-7197-2120

Manuscript received December 2019; accepted July 2020.