Hammill Institute on Disabilities

Focus on Autism and Other Developmental Disabilities
2020, Vol. 35(3) 153–164

DOI: 10.1177/1088357620902492
focus.sagepub.com
SAGE

Effects of Delayed Video-Based Feedback and Observing Feedback on Paraprofessional Implementation of Evidence-Based Practices for Students With Severe Disabilities

Matthew E. Brock, PhD[1], Mary A. Barczak, MS[1], and Scott A. Dueker, PhD[2]

Abstract

Coaching with live observation and immediate performance feedback is an effective means to train paraprofessionals, but might not always be feasible. We used a multiple baseline across participants design with six paraprofessionals who taught elementary students with severe disabilities to test the efficacy of two innovations designed to improve the feasibility of delivering feedback. We found a functional relation between delayed, video-based performance feedback and paraprofessional implementation fidelity of two systematic prompting strategies. Observing a colleague receive feedback did result in some improvement, but did not enable all paraprofessionals to meet the training criterion. These findings suggest that delayed, video-based feedback is an effective and feasible training tool, but only observing a colleague receive feedback might be insufficient.

Keywords

paraprofessionals, severe disabilities, performance feedback, observational learning, personnel preparation

Paraprofessionals play a large role in the education of students with severe disabilities. Paraprofessionals outnumber licensed special education teachers in the United States by a ratio of 1.1:1 (U.S. Department of Education, 2017). Although federal data do not disaggregate paraprofessionals and teachers who serve only students with severe disabilities, state and local-level data suggest that the ratio is even higher for this subgroup (Suter & Giangreco, 2009). According to one descriptive study, 99% of paraprofessionals report that they spend at least some time providing one-to-one instruction to students (Carter et al., 2009). For paraprofessional-delivered instruction to promote student progress, it is critically important for paraprofessionals to be well-trained and supported to implement evidence-based practices that promote student learning (Brock & Carter, 2013).

Scholars have raised serious concerns about whether paraprofessionals are sufficiently trained and supervised in these instructional roles, and whether paraprofessional-delivered instruction is appropriately contextualized as supplemental to teacher-delivered instruction (e.g., Suter & Giangreco, 2009). They point to problematic cases when paraprofessionals deliver instruction with insufficient teacher oversight (Giangreco & Broer, 2005). Others have pointed out that paraprofessionals are rarely given the training needed to deliver instruction effectively (Brock & Carter, 2013; Carter et al., 2009). Scholars also suggest that in some cases, paraprofessional-delivered instruction might be inappropriately used as a substitute for teacher-delivered instruction, which deprives students of the opportunity to learn from highly qualified teachers (Giangreco et al., 2001). Indeed, these kinds of practices are ethically problematic (Giangreco et al., 2001), inconsistent with federal law (Individuals with Disabilities Education Improvement Act, 2004), and unlikely to promote optimal progress for students with severe disabilities (Brock & Carter, 2013).

To address these concerns, it is critically important for schools to only task paraprofessionals with instructional roles that fall within certain parameters. Specifically, paraprofessionals should only deliver instruction (a) when a teacher has

[1]The Ohio State University, Columbus, USA
[2]Ball State University, Muncie, USA

Corresponding Author:
Matthew E. Brock, Crane Center for Early Childhood Education and Policy, Department of Educational Studies, The Ohio State University, 334 PAES Building, 305 West 17th Avenue, Columbus, OH 43210, USA.
Email: brock.184@osu.edu

designed the instruction; (b) when the paraprofessional has been given adequate training to implement the instruction as designed by the teacher; (c) when teachers provide adequate oversight and feedback; (d) when paraprofessionals share student data with teachers, and teachers decide if and how to adjust instruction; and (e) when this instruction is supplemental to teacher-delivered instruction (Uitto et al., 2016).

Researchers have demonstrated how paraprofessionals can be tasked with delivering instruction and support within these parameters. For example, Brock et al. (2016) showed that special education teachers can create individualized plans that outline how paraprofessionals should facilitate peer support arrangements in middle school general education classrooms, and then direct and supervise paraprofessionals to implement the plans. They provided evidence that paraprofessional implementation of peer support arrangements improved social outcomes for students with severe disabilities. In another study, Wermer et al. (2017) demonstrated that a teacher can use coaching and feedback to enable a paraprofessional to elicit and promote communication for an elementary student with autism and complex communication challenges. These examples illustrate that it is indeed possible for paraprofessionals to provide instruction that is teacher-directed, supplemental to teacher-delivered instruction, and effective in improving outcomes for students with severe disabilities.

Despite these examples from the research literature, there is little evidence that similar models have been adopted in everyday practice (Brock & Carter, 2013). One possible explanation might be that these training and supervision models are intensive and time-consuming. Specifically, research-based paraprofessional training has focused heavily on repeated live observations and immediate performance feedback in a one-to-one format (Brock & Carter, 2013). This approach involves a trainer visiting the classroom to observe attempted implementation of a practice in real time, and then meeting with the paraprofessional to share constructive feedback. This feedback focuses on aspects of implementation that were well-executed, and how others might be improved in the future. Researchers have recommended that performance feedback be delivered during or immediately after a teaching session because this has tended to be more effective in the research literature (Scheeler et al., 2016). Furthermore, coaching is often extended over time. In the existing literature, trainers sometimes delivered five or more training sessions to paraprofessionals to ensure sustained implementation fidelity of a single practice (e.g., Leblanc et al., 2005).

There are a number of reasons why a teacher would find it challenging to deliver this kind of one-to-one coaching to paraprofessionals. These challenges stem from a teacher's limited time during the school day, which they report is a major barrier to training paraprofessionals (French, 2001). This limited time makes it difficult for teachers to (a) observe paraprofessionals during the school day, (b) deliver performance feedback immediately after an observation, (c) schedule multiple training sessions within a short window of time, and (d) replicate this one-to-one training across all paraprofessionals whom they supervise. Innovations are needed to improve the feasibility of performance feedback.

In this study, we investigated some practical solutions that might mitigate or eliminate these challenges. First, we eliminated the need for trainers to conduct live observations by video recording paraprofessional implementation, providing trainers the flexibility to observe implementation at a more convenient time, such as during a teacher planning period or before or after school. Second, we examined whether performance feedback could result in rapid acquisition of implementation fidelity even if delivered after a short delay (i.e., 2–3 school days). Third, we tested whether paraprofessionals would be able to acquire and maintain implementation fidelity given only two sessions of direct performance feedback. Fourth, we examined whether a paraprofessional might benefit from observing a colleague receive performance feedback. If paraprofessionals can improve their implementation by observing a colleague receive feedback (without directly receiving feedback on their own practice), trainers could more efficiently train paraprofessionals in small groups. We hypothesized that the observing paraprofessional would make progress by viewing additional exemplars that could be generalized to her own implementation (Stokes & Baer, 1977). For each step, the observing paraprofessional would see an exemplar from her colleague, or the trainer point out the error and model an exemplar for that step. Furthermore, the trainer and the colleague would practice steps on which errors had been made, providing additional opportunities to observe correct implementation. In addition, the observing paraprofessional would see the trainer distinguish whether each step was followed with fidelity, which could generalize to the paraprofessional more accurately judging her own implementation fidelity. To ensure that the feedback they observed would be directly relevant to their own implementation, we designed the study so that each pair of paraprofessionals would be targeting very similar types of skills; specifically, all students were labeling discrete targets; see "Student progress" in the "Method" section for further description.

Although performance feedback has been identified as a key active ingredient in paraprofessional training, it is rarely delivered in isolation of other training strategies (Brock et al., 2017). Researchers often pair performance feedback with written instructions and an implementation checklist (Brock & Carter, 2013). In an effort to isolate the effects of performance feedback, we first tested the effects of written instructions alone before introducing performance feedback. We did not anticipate that written instructions alone would promote implementation fidelity (Brock et al., 2017). We used these strategies to train

paraprofessionals to implement simultaneous prompting and least-to-most (LTM) prompting, because these two strategies have been well established as evidence-based practices (Wong et al., 2015).

Specifically, we addressed the following research questions: First, is there a functional relation between written instructions and paraprofessional implementation of two systematic prompting strategies? Second, is there a functional relation between delayed video-based feedback and paraprofessional implementation? Third, is there a functional relation between observing a colleague receiving feedback and paraprofessional implementation? Fourth, do the effects of delayed video-based feedback maintain after the training is over? Fifth, what progress did students who receive instruction from paraprofessionals make on individualized goals?

Method

Paraprofessionals and Students

After receiving Institutional Review Board and district approval, we recruited six paraprofessionals and five students with severe disabilities. To be included, a paraprofessional must have provided one-to-one instruction to an elementary student with a severe disability (i.e., received special education services under the category of intellectual disability, autism, or multiple disabilities; and was eligible for the state's alternate assessment). All paraprofessional reported that they were unfamiliar with and had not received training on simultaneous prompting, LTM prompting, or any other instructional strategy prior to the study. In addition to providing one-to-one instruction to the target students, all paraprofessionals also had other responsibilities at other times of the school day (e.g., providing one-to-one instruction to other students; supporting students in general education classrooms; supervising students during lunch, recess, and transitions). Paraprofessionals and students are described in the following.

Anna and Adam. Anna was a 50-year-old White female paraprofessional with 6 years of experience in special education and a bachelor's degree in an unrelated field. Anna had received district-provided training on crisis prevention. She taught Adam, an 11-year-old White male student with autism spectrum disorder (ASD). Adam had a standard score of <50 on the Developmental Profile–Third Edition (Alpern, 2007).

Belva and Brenda. Belva was a 46-year-old White female paraprofessional with 6 years of experience in special education and a high school diploma. Brenda had received district-provided training on crisis prevention and behavior management. She taught Brenda, a 10-year-old White female student with ASD. Brenda had a Global Adaptive Composite score of 66 on the Adaptive Behavior Assessment System–Second Edition (ABAS-II; Harrison & Oakland, 2003).

Clair and Carl. Clair was a 51-year-old White female paraprofessional with 1 year of experience in special education and a high school diploma. She had received district-provided training on managing problem behavior, crisis prevention, and first aid. She taught Carl, a 7-year-old Black male student with ASD. Carl had a standard score of 32 on the *Vineland Adaptive Behavior Scales–Third Edition* (Vineland-III; Sparrow et al., 2005). Carl frequently exhibited challenging behavior (i.e., hitting, pinching, and spitting).

Denise. Denise was a 57-year-old White female paraprofessional with 12 years of experience in special education and a high school diploma. She had received district-provided training on managing problem behavior, crisis prevention, and first aid. Denise also worked with Carl (described above).

Emily and Eric. Emily was a 53-year-old White female paraprofessional with 21 years of experience in special education and a high school diploma. She had received district-provided training on school safety. She taught Eric, a 13-year-old White male student with ASD and intellectual disability. Eric had an IQ score of 40 on the *Wechsler Intelligence Scale for Children–Fourth Edition* (WISC-IV; Wechsler, 2003) and a standard score of 63 on the Vineland-III.

Felicia and Frank. Felicia was a 59-year-old White female paraprofessional with 7 years of experience in special education and an associate's degree. She had received district-provided training on school safety. She taught Frank, a 9-year-old White male student with multiple disabilities. Frank had an IQ score of 56 on the *Stanford–Binet Intelligence Scales* (Roid, 2003) and a standard score of 80 on the ABAS-II.

Settings and Materials

Participants were recruited from two elementary schools in two school districts. The first served nearly 600 students in rural and suburban communities, approximately 65% of whom were White and about one sixth received free or reduced-priced meals. The second served about 400 students in a rural area. More than 95% of the students were White and about two thirds received free or reduced-price meals. Both schools were located in the Midwest. All instruction was delivered in self-contained special education classrooms. All training was provided to paraprofessional in these same classrooms, where we either sat together at a

cluster of desks (i.e., Anna and Belva) or at a table (Clair, Denise, Emily, and Felicia). Training materials included a 6-page handout and an inexpensive computer tablet (i.e., Amazon Fire) with a mini-tripod.

Dependent Measures and Recording

We collected data on both paraprofessional implementation fidelity and individualized student outcomes. All data were collected live using a pencil-and-paper system.

Paraprofessional implementation fidelity. We collected data on implementation of simultaneous prompting and LTM prompting. Simultaneous prompting involves delivering a controlling prompt (i.e., the least intrusive prompt that results in the student performing the skill correctly) immediately after delivering a cue or task direction. In contrast, LTM prompting involves giving an opportunity to respond independently, and then providing increasing levels of assistance until the student performs the skill correctly.

We focused on two dimensions of fidelity: adherence to the protocol and quality of implementation (O'Donnell, 2008). Both dimensions were measured at every observation. The primary dimension was adherence to the protocol. First, we told the paraprofessional to use simultaneous prompting. Then we used a checklist to measure whether the paraprofessional adhered to the implementation of each step associated with simultaneous prompting, including (a) delivering a cue or task direction; (b) immediately providing a controlling prompt; and (c) providing specific praise after a correct response, or repeating the controlling prompt after an incorrect response. Each step was scored as implemented or not implemented. After we scored three trials, we directed the paraprofessional to teach using LTM prompting as if the student had already worked on the skill for several days. Then we measured the degree to which the paraprofessional adhered to the implementation steps associated with LTM prompting, including (a) delivering a cue or task direction; (b) providing a 3 to 5 s interval for an independent student response; and (c) providing specific praise after a correct response, or providing prompts of increasing intensity after an incorrect response. For each prompting procedure, we calculated the percentage of steps implemented correctly out of the total number of steps across the three trials. In our previous work, we have found data from three trials provides a representative sample of fidelity (Brock, Seaman, & Downing, 2017). For both procedures, our goal was to train paraprofessionals to at least 80% fidelity across three consecutive sessions.

Quality of implementation was treated as a secondary, descriptive dimension of implementation fidelity. At the end of each observation, we scored six items on a 4-point Likert-type scale with three being the highest quality and zero 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 (see Table 1). This is an instrument that was constructed by our research team. We selected areas of focus based on our anecdotal notes from observations in previous studies, and we based the specific quality indicators on evidence-based recommendations for implementation (i.e., Neitzel & Wolery, 2009). We computed an average score across items as a measure of overall quality. Higher scores represented higher quality instruction.

Student progress. We tracked student progress as a secondary dependent variable. Student learning targets were selected from each student's Individualized Education Program (IEP) in collaboration with the student's teacher of record. Adam was labeling the names and values of four real coins (i.e., penny, nickel, dime, and quarter). Brenda was defining 10 vocabulary words printed on flashcards from the general education science curriculum. These words included carnivore, herbivore, omnivore, decomposer, producer, consumer, food chain, food web, predator, and prey. Carl was identifying the letters in his own name printed on laminated pieces of paper (with Clair) and identifying colors printed on laminated paper (with Denise). The colors were red, blue, and yellow. Eric and Frank were each reading 10 high-frequency sight words on flashcards. Eric's words included their, other, many, been, called, your, have, over, sound, and where. Frank's words included these, down, made, with, what, water, were, place, years, and little.

Experimental Design and Procedures

We used a multiple baseline across participants design (Gast & Ledford, 2014) to test the efficacy of (a) written directions relative to no training; (b) observing feedback relative to written directions; (c) direct video-based feedback relative to written directions; and (d) direct video-based feedback relative to observing feedback. We required a minimum of three data points in the no training condition, and five data points in all other conditions. Paraprofessionals were paired so that they would begin the video-based feedback session simultaneously and one could receive direct feedback while the other observed. We changed conditions for a given pair of paraprofessionals when the minimum number of points were collected for those paraprofessionals, trends were flat across all paraprofessionals, and we were able to schedule any training required in the subsequent condition. We did not require any level of minimum or criterion-level performance before changing conditions.

No training. We did not deliver any training to paraprofessionals in this condition. We simply measured the dependent variables as described in the "Dependent Measures and Recording" section.

Table 1. Items Scored for Measure of Quality of Implementation Fidelity.

Item focus	Score	Descriptor
Pacing	0	Virtually no pause (<1 s between all trials) or long pause (>5 s) between all trials
	1	Brief pause for less than half of trials
	2	Brief pause for at least half, but not all trials
	3	Brief pause (1–5 s) between all trials
Consistency of cue	0	Topography differed across multiple trials
	1	Topography differed for one trial
	2	Identical in topography, but differed in presentation (e.g., verbal directions with different wording)
	3	Topography and presentation consistent across all trials
Immediacy of prompt after error	0	No prompt is delivered
	1	Prompt is very delayed after error is apparent (>3 s)
	2	Prompt is somewhat delayed after error is apparent (1–3 s)
	3	Prompt is delivered immediate after error is apparent (≤1 s)
Verbal praise	0	Some praise has negative tone that sounds insincere and disingenuous
	1	All praise has a neutral tone that is neither negative nor positive
	2	Tone is inconsistent; sometimes positive and sometimes neutral
	3	All praise has a positive tone that is sincere and genuine
Immediacy of reinforcement	0	Reinforcement is never delivered
	1	Reinforcement is very delayed (>3 s) after a correct response
	2	Reinforcement is somewhat delayed (1–3 s) after a correct response
	3	Reinforcement is immediate (≤1 s) after a correct response
Overall impression	0	Poor: There is one significant problem with implementation quality (e.g., negative tone, poorly chosen prompt, ineffective reinforcer)
	1	Fair: Overall, there were no significant problems with implementation quality, but there are two or more minor issues that could be improved (e.g., sometimes reinforcement is slightly delayed, wording of task direction is sometimes slightly different)
	2	Good: Overall, there were no significant problems with implementation quality, but there is one minor issue that could be improved
	3	Excellent: Overall, there are no significant or minor issues related to quality of implementation that could be improved

Note. All items were scored on a 4-point scale with 0 reflecting the *poorest quality* and 3 reflecting the *highest quality*. Scores across items were averaged to compute an overall score.

Written directions. We provided paraprofessionals with a 6-page handout. The first page defined simultaneous prompting, and explained when and why it should be used. The second page defined a controlling prompt, explained how to select a controlling prompt based on the skill and student, and provided examples of controlling prompts. The third page defined LTM prompting, and explained when and why it should be used. The fourth page explained how to select levels in a prompting hierarchy based on the skill and student, and provided examples of prompting hierarchies. The fifth and sixth pages were detailed implementation checklists for simultaneous prompting and LTM prompting. We instructed paraprofessionals to read the handouts and do their best to implement the procedures.

Delayed video-based feedback. There were two different ways delayed video-based feedback was provided: Direct feedback (i.e., feedback on one's own performance 2–3 days after implementation) and indirect feedback (i.e., observing someone else receive feedback without receiving any feedback on one's own performance). Within each pair of paraprofessionals, one paraprofessional was selected randomly to receive direct feedback. The selected paraprofessional used a computer tablet with an attached mini-tripod to record her own instruction. Two to three school days after the video was recorded, the first author provided feedback to the selected paraprofessional on her performance (i.e., direct feedback) while the second paraprofessional observed the feedback session (i.e., indirect feedback). The second paraprofessional observed all aspects of the feedback session, and was only permitted to ask questions that related directly to the colleague's video. Feedback sessions involved the first author (a) reviewing all implementation steps prior to watching the video; (b) watching the video and providing specific praise and/or corrective feedback for each implementation step; and (c) engaging in role-play (with only the paraprofessional receiving direct feedback) to confirm that the paraprofessional could correctly implement each step.

Feedback was not scripted. On average, feedback sessions lasted 29 min (range = 26–36). Each paraprofessional received two sessions of direct feedback. After both paraprofessionals demonstrated stable patterns of responding (i.e., flat trends), the second paraprofessional transitioned to the direct feedback phase.

Maintenance. We terminated all feedback sessions, but allowed paraprofessionals to keep and refer to the written directions.

Procedural Fidelity

For the written directions condition, all paraprofessionals reported that they had read all written materials. For the video-based performance feedback condition, the trainer completed a 16-item checklist during each feedback session. An observer also completed this checklist during 83.3% of sessions. Steps were followed with 100% fidelity across all sessions, and agreement was 100% between the trainer and the observer.

Observer Training and Reliability

The second author was the primary data collector, and the third author and one other observer collected secondary data to measure reliability. All observers were doctoral students in special education and were trained by the first author, a faculty member in special education. Before collecting data, observers were required to (a) score 100% on a written test of coding definitions and rules, (b) achieve 95% agreement with the first author when coding a training video, and (c) achieve 95% agreement with the first author in a live setting.

Two observers collected data during 32% of all observations, balanced across participants and conditions. We computed point-by-point agreement on each behavior. Total agreements (i.e., identical codes for both observers) were divided by opportunities for agreement. Overall agreement across all paraprofessional behaviors was 98.7% (range = 80.0%–100.0%). Total agreement across all student behaviors was 99.5% (range = 83.3%–100.0%).

Social Validity

We used a 10-item paper-pencil questionnaire to measure paraprofessional 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. In addition to these forced choice items, we invited paraprofessionals to write what they liked most about the training, and what they thought could be improved.

Results

Paraprofessional Adherence to Steps

Adherence data for simultaneous prompting and LTM prompting are displayed in Figure 1 and described below for each participant in terms of trend, level, and variability. No basic effects were demonstrated for written directions alone. Observed feedback resulted in a change in data patterns across three participants, but did not consistently produce the desired effect of sustained criterion-level adherence (i.e., three consecutive data points $\geq$80%). In one case, a paraprofessional did reach criterion-level adherence with observed feedback. In all other cases, direct video-based feedback enabled participants to rapidly acquire criterion-level adherence for both procedures, demonstrating a functional relation. Effects maintained for all participants after all feedback was withdrawn.

Anna. During the no training condition, simultaneous prompting data began at 0%, increased to 50%, and then remained flat at 50%. LTM prompting data were flat and between 40% and 50%. During the written instructions condition, simultaneous prompting slowly trended upward to peak at 75% and then decreased to 63%. LTM prompting data were more variable compared with no training, and lower in level. No effects were demonstrated for written directions. During the direct feedback condition, simultaneous prompting data immediately increased in level and trended to 100% and remained at that level. LTM prompting data immediately increased to 100% and maintained at that level. Therefore, basic effects were demonstrated for direct feedback. Data for both procedures remained at 100% throughout the maintenance condition.

Belva. During the no training condition, data for both procedures had a flat trend and ranged 25% to 40%. During the written directions condition, data for both procedures were initially variable. Simultaneous prompting data stabilized at 63%, and LTM prompting data remained variable between 25% and 67%. During the observed feedback condition, simultaneous prompting data initially decreased, trended upward to 88%, and then trended downward to 50%. LTM prompting data immediately increased in level to 83%, and then fluctuated between 60% and 100%. Although there was a clear shift in data patterns, and Belva did demonstrate criterion-level implementation fidelity for both procedures, this level of fidelity was not sustained across three consecutive sessions. During the direct feedback condition, data for both procedures immediately increased to 100% and remained at that level; therefore, basic effects were demonstrated for direct feedback. Data remained at 100% throughout the maintenance condition.

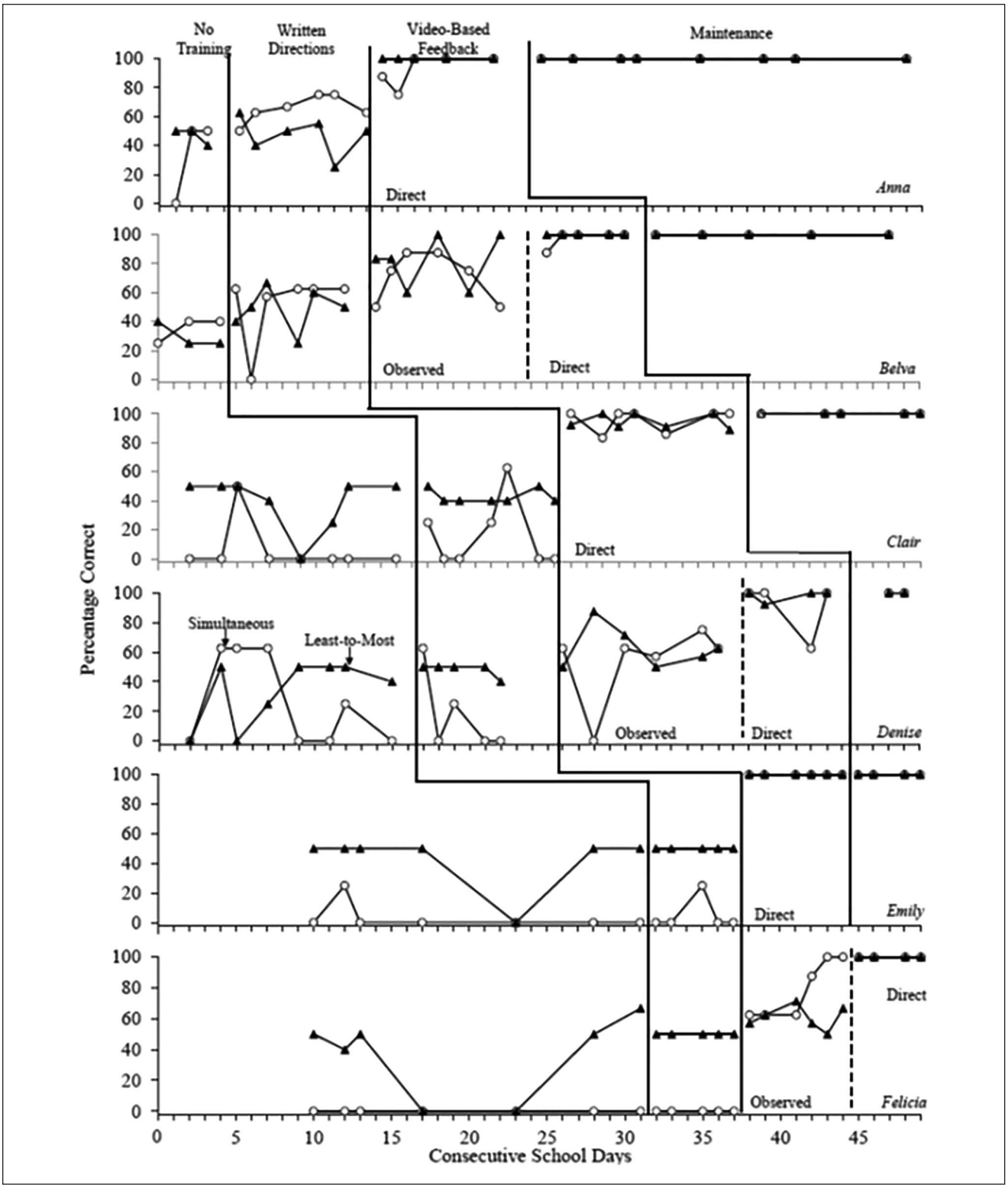

Figure 1. Paraprofessional adherence to steps for simultaneous prompting (open circles) and least-to-most prompting (closed triangles). Observed feedback involved watching a colleague receive feedback on her performance; direct feedback involved receiving feedback on one's own performance.

Clair. During the no training condition, simultaneous prompting data were flat at 0% with the exception of one data point at 50%. LTM data were initially flat at 50%, trended downward to 0%, and then back up to 50%. During the written directions condition, simultaneous prompting data were more variable, but four of the seven points were still 0%. LTM prompting data were flat, with all data between 40% and 50%. No basic effects were demonstrated for written directions. During the direct feedback condition, data for both conditions increased immediately, and were somewhat variable between 83% and 100%; therefore, basic effects were demonstrated for direct feedback. Data remained at 100% throughout the maintenance condition.

Denise. During the no training condition, simultaneous prompting data trended from 63% down to 0%. LTM prompting data trended toward 50% and remained flat between 40% and 50%. During the written directions condition, simultaneous prompting data initially increased and then trended down to 0%. LTM prompting data remained flat at 40% to 50%. No basic effects were demonstrated for written directions. During the observed feedback condition, simultaneous prompting data increased in level, with all but one point between 57% and 75%. LTM data initially trended upward to 88% before trending back down to 50% to 63%. Although Denise's performance improved relative to the previous condition, observed feedback did not result in sustained criterion-level implementation fidelity for either procedure. During the direct feedback condition, data for both procedures immediately increased to 100%, with some variability; therefore, basic effects were demonstrated for direct feedback. Data remained at 100% throughout the maintenance condition.

Emily. During the no training condition, simultaneous prompting data were flat at 0%, and LTM prompting data were flat at 50%, with some variability. These patterns were unchanged in the written directions condition; no basic effects were demonstrated. During the direct feedback condition, data immediately increased to 100% for both procedures and remained at that level; therefore, basic effects were demonstrated for direct feedback. Data remained at 100% throughout the maintenance condition.

Felicia. During the no training condition, simultaneous prompting data were flat at 0%, and LTM prompting data were highly variable between 0% and 67%. During the written directions condition, simultaneous prompting data remained unchanged, and LTM-prompting data decreased in variability but did not change in level; no basic effects were demonstrated. During the observed feedback condition, simultaneous prompting data immediately increased in level to 63% and then trended to 100%. LTM prompting data increased slightly in level and variability (range = 50%–71%). Both data patterns changed relative to the previous condition, but sustained criterion-level fidelity was demonstrated only for simultaneous prompting. During the direct feedback condition, data maintained at 100% for simultaneous prompting, and increased to 100% and maintained at that level for LTM prompting.

Paraprofessional Implementation Quality

Overall quality of implementation for each paraprofessional in each experimental condition is reported in Table 2. All paraprofessionals improved their quality of implementation after observed feedback or direct feedback, mirroring improvements in adherence data. Three paraprofessionals maintained high quality after training was withdrawn.

Student Progress

Student progress was tracked as a secondary dependent variable, and is reported in Figure 2. Below, data are described in terms of trend, level, and variability.

Adam's data were variable between 0% and 50% across phases. No basic effects were demonstrated for any experimental condition. Brenda's data were flat between 10% and 20% in the no training and written directions conditions. They increased in level slightly during the observed feedback condition. They trended upward to 90% during the direct feedback condition, and then to 100% during the maintenance condition. Basic effects were demonstrated for both observed and direct feedback. Carl's data were variable for both of his instructional goals, with no basic effects demonstrated for any experimental condition. Eric's data trended to 100% during no training, and maintained at 100% during the written directions condition. We selected a new learning target for Eric, and his data for the new target trended upward in the direct feedback condition and flattened in the maintenance condition. Because of the upward trend in the no training condition and the change in learning targets, there was not an opportunity to demonstrate basic effects. Frank's data trended upward to 70% during the no training condition, and were variable between 60% and 90% in all subsequent conditions. Due to the upward trend in the no training condition, there was not an opportunity to demonstrate basic effects.

Social Validity

In response to close-ended questions, paraprofessionals reported a high degree of satisfaction with the training and with the systematic teaching strategies (see Table 3). For open-ended written responses, paraprofessionals reported that the aspects of training that they liked best included the actual strategies being taught, receiving video-based feedback, and role-playing correct implementation of the strategies. Paraprofessionals thought the

Table 2. Average Overall Quality of Implementation Scores for Paraprofessionals Across All Experimental Conditions.

Paraprofessional	Baseline	Written directions	Observed feedback	Direct feedback	Maintenance
Anna	2.1 (1.0–3.0)	2.2 (1.2–2.3)	—	2.8 (2.0–3.0)	3.0 (3.0–3.0)
Belva	1.4 (0.0–1.7)	2.2 (1.0–3.0)	2.3 (1.3–3.0)	3.0 (2.8–3.0)	3.0 (3.0–3.0)
Clair	1.6 (0.6–3.0)	2.6 (1.6–3.0)	—	2.7 (1.6–3.0)	2.0 (0.2–3.0)
Denise	1.7 (0.7–2.8)	2.2 (1.0–3.0)	2.6 (1.7–3.0)	2.4 (1.5–3.0)	1.8 (0.0–3.0)
Emily	2.1 (1.1–2.9)	2.7 (1.6–3.0)	—	3.0 (3.0–3.0)	3.0 (3.0–3.0)
Felicia	1.3 (1.0–2.4)	1.9 (1.0–2.8)	2.7 (2.0–3.0)	2.8 (2.5–3.0)	—

Note. The range of overall scores across all sessions within the phase is 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*).

training might be improved by providing feedback sooner and more frequently.

Discussion

There is strong evidence that repeated delivery of immediate performance feedback after live observations can improve practitioner implementation fidelity of evidence-based practices (Scheeler et al., 2016), but this form of professional development would be difficult for teachers to replicate. In this study, we tested innovations that would improve the efficiency and feasibility of delivering performance feedback. We identified a functional relation between delayed, video-based feedback, and practitioner implementation fidelity of two systematic prompting strategies. These findings extend the research literature in a number of key ways.

First, delayed, video-based feedback is an effective means to enable paraprofessionals to quickly acquire and maintain implementation fidelity of systematic prompting strategies. We demonstrated improvements in both adherence to steps and quality of implementation. Although immediacy of performance feedback has been previously identified as a factor that might moderate the efficacy of performance feedback (Scheeler et al., 2016), our findings suggest that delayed, high-quality feedback can produce desired outcomes. Delayed video-based feedback provides the trainer with flexibility about the timing and location of feedback sessions, increasing feasibility. This flexibility could be helpful for both supervising teachers and administrators who wish to provide feedback to paraprofessionals, but have many other competing responsibilities during the school day. Maintenance of effects suggests that paraprofessionals can continue high-fidelity implementation without additional training.

Second, observing a colleague receive feedback might improve paraprofessional implementation fidelity, but it does not consistently produce the desired outcome of sustained criterion-level fidelity. Of the three paraprofessionals who observed feedback, only one acquired criterion-level implementation fidelity for one of the prompting procedures through observed feedback alone. It is difficult to argue that small improvements during observed feedback might facilitate more rapid progress during direct feedback, because all participants acquired fidelity immediately after receiving direct feedback regardless of whether they had previously observed feedback. We were surprised that observed feedback did not have more potent effects, especially given that each pair of paraprofessionals were targeting very similar student skills. It seems that generalizing from a colleague's teaching to one's own teaching might have been a greater challenge for paraprofessionals than we anticipated. Specifically, paraprofessionals struggled with identifying their own implementation errors. Perhaps this was difficult because they were only exposed to a single exemplar of implementation (i.e., their colleague's implementation). Paraprofessionals might have been more successful at identifying their own implementation errors after observing multiple exemplars (Stokes & Baer, 1977).

Third, as hypothesized, there was no evidence that written directions alone had an effect on paraprofessional implementation fidelity. This was unsurprising, given that meta-analyses have shown that written directions are ineffective in isolation, but can make a positive contribution when combined with modeling and performance feedback (Brock et al., 2017). Our study corroborates these findings, because paraprofessionals made improvements only after they had access to both written directions and performance feedback.

Fourth, mixed patterns of student progress might be attributed to a number of factors. One student made progress that coincided with paraprofessional acquisition of criterion-level implementation fidelity. Other students began to demonstrate learning prior to paraprofessional training, suggesting that some students may respond to repeated instruction even if it is not evidence-based. Still other students did not make clear progress at any time in the study. In these cases, we offer several possible explanations. It is possible that the students would have benefited more from a different evidence-based approach (e.g., constant time delay, graduated guidance), and that the particular practices used in this study (i.e., simultaneous prompting

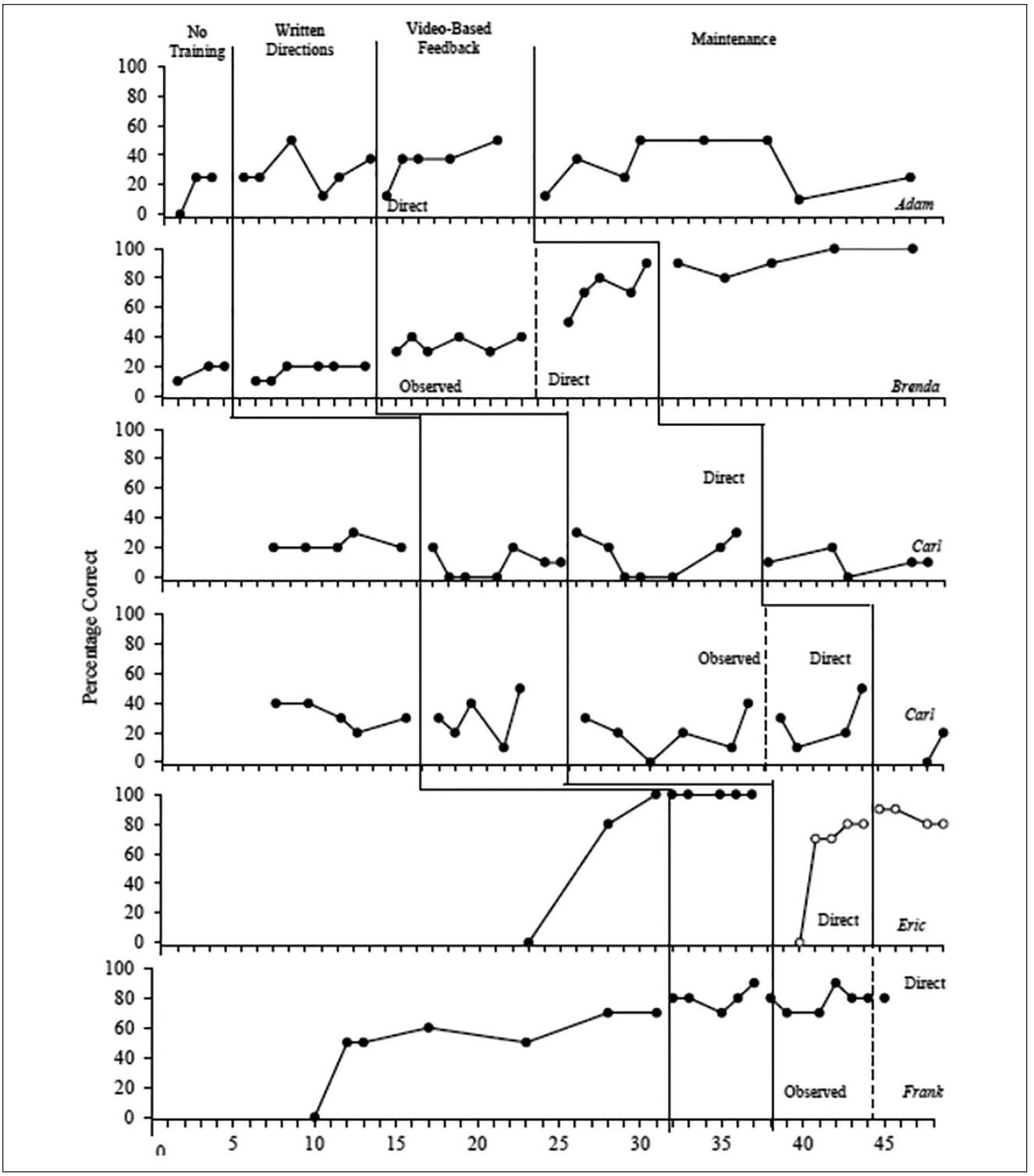

Figure 2. Student progress on individualized goals given paraprofessional instruction.

and LTM prompting) might not have been the best match for these particular students. Alternatively, instructional modifications such as more potent reinforcers or different instructional materials might have better promoted student progress. Students like Carl who exhibited challenging behavior may have made more progress if paraprofessionals were also trained to implement function-based behavior management strategies.

Table 3. Social Validity Questionnaire Ratings by Paraprofessional Participants.

Category/question	*M*	*SD*
Perception of skill after training		
How skilled are you in implementing simultaneous prompting?	4.83	0.41
How skilled are you in implementing least-to-most prompting?	4.83	0.41
How skilled are you in data collection?	4.67	0.81
Perception of training		
How effective was the training at enabling you to implement new strategies with your student?	4.83	0.41
How much do you think that the new strategies you learned contributed to your student's progress on his or her goal?	4.50	0.55
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?	4.83	0.41
How likely would you be to use the strategies that you learned in the future with a different student?	4.83	0.41
How likely would you be to participate in a similar training opportunity in the future?	5.00	0.00
How likely would you be to recommend a similar training opportunity to a colleague?	4.83	0.41

Note. Response options: 1 = *Not at all*, 2 = *Slightly*, 3 = *Somewhat*, 4 = *Moderately*, 5 = *Extremely*.

Fifth, highly variable performance in baseline condition demonstrates that untrained paraprofessionals do sometimes implement some steps associated with prompting procedures, but do not do so consistently or systematically without training. This is consistent with other studies that report highly variable paraprofessional implementation of systematic prompting procedures during a baseline condition (e.g., Leblanc et al., 2005).

Implications for Practice

Findings from this study have important implications for practice. Teachers and administrators can provide effective, low-intensity feedback to paraprofessionals by providing them with a written implementation checklist, directing them to video record their teaching, and providing delayed feedback at a time and place that are convenient. Teachers should follow their school's policies related to video recording for the purpose of professional development. We used inexpensive (<US$100) generic computer tablets to record implementation, but there is no reason why one could not use any video-recording technology that is already available in the classroom (e.g., camcorder, laptop with webcam). There may be some benefit to delivering training in a small-group format where paraprofessionals observe each other receiving feedback, but trainers should also deliver feedback directly to all participants to ensure criterion-level implementation fidelity. Although delayed video-based feedback allows teachers to observe the video and provide feedback more flexibly, it is still important for administrators to ensure that teachers have time during the school day (e.g., a prep period) to deliver feedback.

Limitations and Directions for Future Research

There are a number of limitations to this study that highlight directions for future research. First, we did not measure generalization, so it is unclear if paraprofessionals would have been able to implement practices with fidelity when targeting new students and situations. Furthermore, it is unclear how observed and direct feedback might differentially impact generalization. Researchers might design future studies that address these questions. Second, we demonstrated experimental control at the level of paraprofessional implementation fidelity, but not at the level of student progress. This highlights two additional training needs, including (a) the need to train paraprofessionals in a range of evidence-based practices so that they can select practices that are the best match for a particular student and goal and (b) the need to train paraprofessionals to make adjustments to their instruction to optimize student progress. Researchers might design future studies to identify effective ways to meet these needs. Third, because we only introduced performance feedback after the written directions condition, one cannot rule out the possibility that performance feedback is effective only after written directions. Finally, visual analysis of single-case design data does involve some subjectivity, and it is possible that other visual analysts may make somewhat different judgments.

Conclusion

In this study, we demonstrated that delayed, video-based performance feedback is an effective and efficient means to enable paraprofessionals to implement evidence-based systematic prompting procedures. In contrast with previous

recommendations to conduct live observations and deliver immediate feedback, this approach provides increased flexibility to teachers who are responsible for training and supervising paraprofessionals. This finding contributes to developing more feasible and effective models of paraprofessional training that have the potential to improve outcomes for students with severe disabilities.

Declaration of Conflicting Interests

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

Funding

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

ORCID iD

Scott A. Dueker https://orcid.org/0000-0002-4101-6338

References

Alpern, G. D. (2007). *Developmental profile 3 (DP-3)*. Western Psychological Services.

Brock, M. E., Biggs, E., Carter, E. W., Cattey, G., & Raley, K. (2016). Implementation and generalization of peer support arrangements for students with severe disabilities in inclusive classrooms. *The Journal of Special Education*, *49*, 221–232. https://doi.org/10.1177/0022466915594368

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 instruction 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., Seaman, R. L., & Downing, C. (2017). Promoting learning for a student with a severe disability through paraprofessional training. *Research and Practice for Persons with Severe Disabilities*, *42*, 211–224. https://doi.org/10.1177/1540796917729682

Carter, E. W., 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

French, N. K. (2001). Supervising paraprofessionals: A survey of teacher practices. *The Journal of Special Education*, *35*, 41–53. https://doi.org/10.1177/002246690103500105

Gast, D. L., & Ledford, J. R. (2014). *Single subject research methodology in behavioral sciences*. Routledge.

Giangreco, M. F., & Broer, S. M. (2005). Questionable utilization of paraprofessionals in inclusive schools: Are we addressing symptoms or causes? *Focus on Autism and Other Developmental Disabilities*, *20*, 10–26. https://doi.org/10.1177/10883576050200010201

Giangreco, M. F., Broer, S. M., & Edelman, S. W. (2001). Teacher engagement with students with disabilities: Differences between paraprofessional service delivery models. *Journal of the Association for Persons with Severe Handicaps*, *26*, 75–86.

Harrison, P., & Oakland, T. (2003). *Adaptive behavior assessment system (ABAS-II)*. The Psychological Corporation.

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

Leblanc, M. P., Ricciardi, J. N., & Luiselli, J. K. (2005). Improving discrete trial instruction by paraprofessional staff through an abbreviated performance feedback intervention. *Education and Treatment of Children*, *28*, 76–82.

Neitzel, J., & Wolery, M. (2009). *Steps for implementation: Least-to-most prompts*. National Professional Development Center on Autism Spectrum Disorders, Frank Porter Graham Child Development Institute, The University of North Carolina.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, *78*, 33–84. https://doi.org/10.3102/0034654307313793

Roid, G. H. (2003). *Stanford-Binet Intelligence Scale* (5th ed.). Riverside.

Scheeler, M. C., Morano, S., & Lee, D. L. (2016). Effects of immediate feedback using bug-in-ear with paraeducators working with students with autism. *Teacher Education and Special Education*, *41*, 24–38. https://doi.org/10.1177/0888406416666645

Sparrow, S. S., Cicchetti, D. V., Balla, D. A., & Doll, E. A. (2005). *Vineland adaptive behavior scale*. American Guidance Service.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *10*, 349–367. https://doi.org/10.1901/jaba.1977.10-349

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

Uitto, D. J., Chopra, R. V., & Stivers, J. (2016). *Council for exceptional children teacher education division position statement of training for paraeducators in special education*. tedcec.org

U.S. Department of Education. (2017). *IDEA Section 618 data products*. https://www2.ed.gov/programs/osepidea/618-data/index.html

Wechsler, D. (2003). *Wechsler intelligence scale for children–Fourth edition (WISC–IV)*. The Psychological Corporation.

Wermer, L., Brock, M. E., & Seaman, R. L. (2017). 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 Disorders*, *33*, 217–226. https://doi.org/10.1177/1088357617736052

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., & Kucharczyk SShultz, 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