Training Supervisors to Provide Performance Feedback Using Video Modeling with Voiceover Instructions

Natalie Shuler¹ · Regina A. Carroll²

Published online: 4 December 2018
© Association for Behavior Analysis International 2018

Abstract

Supervisors commonly use feedback to teach staff members to implement behavioral interventions. However, few studies have evaluated methods to teach supervisors to provide effective feedback. We used a multiple-baseline design to evaluate the use of video modeling with voice-over instruction to train 4 supervisors to provide performance feedback to a confederate therapist implementing a guided-compliance procedure. We assessed supervisors' accuracy with implementing 8 feedback component skills during scripted role-plays before and after video modeling. We also assessed the extent to which supervisors' skills generalized when providing feedback to a confederate therapist implementing novel behavioral procedures (i.e., discrete-trial training and mand training) and an actual therapist implementing the guided-compliance procedure with a child with autism. All supervisors mastered the feedback component skills following video modeling. Overall, the results of the current study suggest that video modeling may be an efficacious and efficient method to train supervisors.

Keywords Guided-compliance procedure · Performance feedback · Staff training · Supervision · Video modeling

Within human service agencies, such as early intensive behavioral intervention clinics, supervisors play a vital role in the quality of services provided to clients. Effective supervision practices require a supervisor to actively work to improve inadequate staff performance and to support and maintain quality staff performance (Reid, Parsons, & Green, 2012). One of the main responsibilities of a supervisor is to regularly

Research Highlights

- Video modeling with voice-over instruction led to improvements in supervisors' accuracy with implementing eight component skills of a performance feedback procedure.
- Video modeling with voice-over instructions may be an efficient alternative to multicomponent interventions that have been used in previous studies to teach supervisors to provide performance feedback.
- Assessing mastery of a supervisor's skills during scripted role-plays with confederates may be predictive of a supervisor's performance with an actual therapist.
- Additional training may be required for some supervisors to implement component skills of performance feedback at mastery level when providing feedback to a therapist implementing novel behavioral procedures.

✉ Regina A. Carroll
Regina.Carroll@unmc.edu

¹ West Virginia University, Morgantown, WV, USA

² Department of Psychology, University of Nebraska Medical Center's Munroe-Meyer Institute, 9012 Q Street, Omaha, NE 68127, USA

monitor the accuracy with which staff members implement interventions with clients (i.e., monitor treatment integrity; Peterson, Homer, & Wunderlich, 1982). Previous research has demonstrated that when staff members implement interventions commonly used in early intervention clinics with low integrity (e.g., discrete-trial teaching; DTT), it can have a negative impact on treatment outcomes for clients (Carroll, Kodak, & Fisher, 2013; Pence & St. Peter, 2015; Wilder, Atwell, & Wine, 2006). When staff members are implementing interventions with low integrity, it is necessary for supervisors to take steps to correct their performance.

Performance feedback, provided by a supervisor, can be an effective strategy for maintaining a staff member's accurate performance and for correcting a staff member's inaccurate performance. Effective performance feedback requires several components including providing (a) praise for the behaviors the staff member performed correctly, (b) a description of the behaviors the staff member performed incorrectly, (c) a rationale for changing behaviors performed incorrectly, (d) instructions for correct performance, (e) a demonstration of correct performance, (f) an opportunity for the staff member to practice correct performance, and (g) an opportunity for the staff member to ask questions (Behavior Analyst Certification Board, 2012; Parsons & Reid, 1995; Reid et al., 2012). Feedback is a common component across many evidence-

based staff-training procedures and has been identified as a necessary or sufficient component of staff-training procedures (Gil & Carter, 2016; Jerome, Kaplan, & Sturmey, 2014; Ward-Horner & Sturmey, 2012).

Despite the critical role that feedback plays in initial and ongoing staff training, employers do not routinely train supervisors to provide effective feedback to staff members. For example, DiGennaro Reed and Henley (2015) distributed a survey requesting information about staff and supervisory training and performance management procedures to 382 individuals who were currently Board Certified Behavior Analysts (BCBAs) or were seeking certification. Of the 382 respondents, 47% reported receiving feedback during initial training and 66% of respondents reported that performance feedback was used as a method of ongoing training. More than half of the respondents (77%) reported that they were responsible for supervising other staff members; however, a majority of these supervisors (66%) reported that their current place of employment did not provide any training on effective supervision practices. The results of this survey highlight the need for additional research to evaluate and disseminate methods for training supervisors to use evidence-based supervision practices. The dearth of literature on methods for teaching supervisors to provide performance feedback represents a significant problem because studies have shown that the accuracy of feedback can influence staff members' performance (e.g., Hirst, DiGennaro Reed, & Reed, 2013). By failing to train supervisors on how to provide performance feedback, we are taking a risk that supervisors will provide inaccurate or inadequate feedback, which may lead to staff members implementing behavioral interventions with low levels of treatment integrity.

At present, there are few studies that have evaluated methods to teach supervisory skills, including providing performance feedback to staff members (Durgin, Mahoney, Cox, Wheeltjens, & Poling, 2014; Green, Rollyson, Passante, & Reid, 2002; Jensen, Parsons, & Reid, 1998; Parsons & Reid, 1995). Parsons and Reid (1995) evaluated the use of a training package to teach supervisors in a residential facility for individuals with disabilities to provide feedback to staff members. The training package consisted of didactic instructions, role-playing, and direct observation with in vivo feedback. First, the experimenters taught the supervisors to accurately implement the teaching protocols used with clients in the facility. Next, the experimenters directly taught supervisors to monitor staff-teaching skills and to provide performance feedback. The results showed that supervisors' feedback skills improved following direct training and the supervisors' feedback was effective at maintaining the staff members' accurate performance. Although the training package was effective at teaching supervisors to provide feedback, it was time and resource intensive. Specifically, the training package required 8 h of classroom instruction, and it required a professional trainer

to observe and provide feedback to the supervisor in the natural environment. It may not always be practical for a professional to be present for an extended period of time to train supervisors. Thus, additional research is needed to identify efficient methods to train supervisory skills.

Video modeling is an effective staff-training procedure that has been shown to reduce the amount of time that a trainer needs to be present during staff training. Video modeling consists of having an individual watch a video that demonstrates the behavior that a viewer should imitate in an appropriate context. Video modeling with voice-over instructions has been used to teach staff members to implement a variety of interventions commonly used in early intervention clinics (e.g., Higgins, Luczynski, Carroll, Fisher, & Mudford, 2017; Vladescu, Carroll, Paden, & Kodak, 2012). The purpose of the present study was to evaluate the use of video modeling with voice-over instructions to train supervisors to provide performance feedback to a confederate therapist implementing a guided-compliance procedure (Wilder & Atwell, 2006). Following the video-modeling intervention, we assessed the extent to which supervisors' feedback skills occurred when providing feedback on novel behavioral procedures and to an actual therapist. Specifically, we assessed supervisors' accuracy with providing performance feedback to a confederate therapist implementing DTT and mand-training procedures and to an actual therapist implementing the guided-compliance procedure with a child with autism spectrum disorder (ASD). Additionally, we used data analysis procedures that allowed us to examine supervisors' accuracy with implementing individual component skills of the performance feedback procedures (similar to Higgins et al., 2017).

Method

Participants, Setting, and Materials

Four individuals who worked at a university-based early intervention clinic for children with ASD served as supervisors for this study. All supervisors worked at the clinic for a minimum of 20 h per week, receiving a tuition waiver and/or stipend for their work. All supervisors had at least 2 years of experience working with children with ASD. Supervisors 1, 2, and 4 were Caucasian females between the ages of 18 and 24 and were enrolled as full-time graduate students seeking a master's of arts in special education at the time of the study. As a part of this program, Supervisors 1, 2, and 4 were completing a course sequence and accruing hours to earn certification as BCBAs. Prior to the start of this study, Supervisors 1, 2, and 4 had each held a supervisor position in the clinic for approximately 1 year. As supervisors, they were responsible for assessing client skills, developing treatment programs, training therapists, and collecting and analyzing data related

to client and therapist performance. Supervisor 3 was a Caucasian male between the ages of 25 and 34 who had a bachelor's degree in psychology. Supervisor 3 served as data coordinator for the clinic. In this role, Supervisor 3 occasionally served as a therapist for clients and assisted with training therapists to implement behavioral procedures. None of the supervisors had received previous training on providing performance feedback.

During generalization probes, the supervisor provided feedback to an actual therapist who worked in the clinic. These therapists were junior or senior undergraduate psychology students who served on a volunteer basis and typically received course credit in professional field experience for their work. Each therapist worked directly with one or more clients, implementing interventions specific to the clients' goals.

We conducted all training sessions in a private conference room located in a university-based early intervention clinic. The conference room contained a table, four chairs, and the materials needed to conduct experimental sessions. Materials needed to conduct sessions included a tripod, a video camera, a laptop, data sheets, and protocols. Additionally, during all sessions the supervisor had access to a bin that contained two timers, two pens, a calculator, and several small toys (e.g., blocks, a car, a finger puppet).

Dependent Measures and Data Collection

The experimenter watched videos of all sessions and scored each supervisor's accuracy with implementing the eight component skills for providing performance feedback (see

Table 1). We summarized the data in two ways. First, for each session we calculated the supervisor's accuracy with implementing each component skill. We calculated the supervisor's percentage of accuracy by dividing the number of times the supervisor implemented a component skill accurately in a session by the total number of opportunities to implement that skill and multiplying by 100. We considered a component skill mastered during a session if the supervisor implemented that component skill accurately on 80% or more of the opportunities. We used these data to analyze supervisor performance across individual performance feedback component skills and for tailored training, if needed.

Second, we summarized the data as a percentage of mastered component skills in a session. We calculated the percentage of mastered component skills by dividing the number of skills that the supervisor implemented accurately on 80% or more opportunities by the total number of component skills. For example, if the supervisor implemented six of the eight feedback component skills accurately on 80% or more opportunities in a session (i.e., mastery), then the percentage of mastered component skills for that session would be 75%. Our mastery criterion for terminating video-modeling sessions was mastery of 88% of the component skills (i.e., seven out of eight skills) in a session. Our mastery criterion for terminating posttraining sessions was mastery of 88% of the component skills across two consecutive posttraining sessions. If any supervisor did not meet mastery following four posttraining sessions, we provided tailored training. Our mastery criterion for terminating tailored-training sessions was mastery of 88% of the component skills in a session.

Table 1 Performance feedback component skills

Dependent Measures	Operational Definition
1. Collects data accurately	The supervisor collects data on correct implementation of the procedure being observed, which aligns with the data of a trained observer with interobserver agreement of at least 80%.
2. Provides behavior-specific praise	The supervisor provides behavior-specific praise for each component skill that the therapist performed with accuracy above 80% during the session.
3. Describes incorrect performance	The supervisor describes each of the steps that the therapist performed with accuracy below 80% during the session.
4. Provides a rationale for change	The supervisor provides a rationale for changing ineffective performance for each of the steps that the therapist performed with accuracy below 80% during the session.
5. Provides instruction	The supervisor provides instructions for how to improve each of the steps that the therapist performed with accuracy below 80% during the session.
6. Provides a demonstration	The supervisor provides a model of correct implementation of each of the steps that the therapist performed with accuracy below 80% during the session.
7. Provides an opportunity for practice	The supervisor provides an opportunity for the therapist to practice each of the steps that the therapist performed with accuracy below 80% during the session. The supervisor has the therapist continue to practice until the therapist implements each step correctly.
8. Provides an opportunity for questions	The supervisor solicits questions from the therapist after providing feedback on correct or incorrect performance.

Additionally, we measured the duration of training required during the video-modeling, tailored-training (Supervisor 4 only), and posttraining sessions for each supervisor. We calculated the total duration of training in minutes for the following activities: (a) reviewing protocols and collecting treatment integrity data, (b) viewing the video model, (c) tailored training (Supervisor 4 only), and (d) providing feedback to a confederate. Observers collected data on the duration of training from the video and used the counter on the video to record the start and stop time for each activity. We operationally defined the start and stop time for each activity. For example, when collecting data on the duration of time for reviewing protocols and collecting treatment integrity data, the observer would record the start time when the experimenter handed the protocol to the supervisor and recorded the end time when the supervisor indicated to the experimenter that he or she had finished collecting data.

Interobserver Agreement (IOA) and Procedural Fidelity

A second observer watched the video and scored the supervisor's accuracy with implementing the eight component skills for providing performance feedback during an average of 54% (range 50%–57%) of the total sessions for each supervisor. We compared the primary and secondary observers' data and scored an agreement if both observers independently recorded the same response (e.g., both observers recorded that the supervisor implemented a component skill accurately during an opportunity) and a disagreement if observers recorded different responses (e.g., one observer recorded that the supervisor implemented a component skill accurately and the other recorded inaccurate implementation). When calculating IOA for the duration of training, we scored an agreement if both observers recorded the same time (within a 15-s window) for each activity within the session (e.g., collecting treatment integrity data, viewing the video model). We calculated the percentage of agreement between the observers by dividing the number of agreements by the total number of agreements plus disagreements and multiplying by 100. Mean agreement scores were 96% (range 93%–100%) for Supervisor 1, 93% (range 85%–97%) for Supervisor 2, 96% (range 82%–100%) for Supervisor 3, and 97% (range 94%–100%) for Supervisor 4.

The second observer also scored the experimenter's accuracy with implementing the experimental protocol (i.e., procedural fidelity) during an average of 54% (range 50%–57%) of the total sessions for each supervisor. During each session, the experimenter responses included (a) showing the video model, defined as showing the video model during video-modeling sessions only; (b) providing instructions to collect integrity data, defined as presenting the instructions as worded in the experimental protocol; (c) allowing the supervisor up to

10 min to review the protocols and operational definitions; (d) showing the session video that was predetermined prior to the start of the session; (e) withholding feedback on the supervisor's performance; (f) asking the supervisor the number of questions that were predetermined prior to the start of the session if the supervisor provided an opportunity to ask questions; and (g) engaging in the number of errors that were predetermined prior to the start of the session if the supervisor provided an opportunity to role-play. The secondary observer scored the experimenter's accuracy with implementing each step of the experimental protocol as either correct, incorrect, or not applicable. We calculated procedural fidelity by dividing the number of steps the experimenter implemented correctly by the number of steps implemented correctly plus the number of steps implemented incorrectly and multiplying by 100. Mean fidelity scores were 100% for Supervisors 1, 2, and 4 and 96% (range 86%–100%) for Supervisor 3.

Pretraining

In order for supervisors to provide accurate feedback, it is necessary for them to collect accurate data on participants' implementation of behavioral procedures. Thus, prior to the start of this study, supervisors were trained to collect data on therapists' accuracy with implementing the procedures used during the study (i.e., guided compliance, DTT, and mand training). As part of their regular duties as supervisors, they used the data sheets and operational definitions that we used in this study to assess therapists' accuracy with implementing behavioral procedures and to provide feedback. We assessed supervisors' accuracy with collecting data on the therapists' accuracy with implementing the guided-compliance, DTT, and mand-training procedures by having them collect data from videos depicting simulated sessions with two confederates (see the Simulated-Session Videos section). Supervisors' data were compared to data collected by a trained secondary observer. If agreement between the supervisor and the secondary observer was below 90%, then we conducted additional training on data collection (i.e., instructions, practice, and feedback). Direct training continued until the supervisor demonstrated that he or she could independently collect data on the therapists' accuracy with implementing the procedure with at least 90% agreement for two consecutive sessions. Due to experimenter error, Supervisor 1 completed the mand-training generalization probe in baseline prior to demonstrating that she could collect data on the therapist's accuracy with implementing the procedure with agreement at or above 90%. In addition to ensuring that supervisors could collect data accurately before participating in the current study, we also monitored each supervisor's accuracy with collecting data throughout the study (see data on Component Skill 1 in the Results section). With the exception of Supervisor 1's mand-training generalization probe mentioned previously, all

supervisors collected data with at least 80% accuracy throughout the study.

Simulated-Session Videos

We used simulated-session videos during pretraining, baseline, video-modeling, posttraining, and follow-up sessions and when assessing generalization to novel behavioral procedures. We used simulated sessions in order to give the supervisors the opportunity to provide feedback on a variety of steps of the procedures (i.e., guided-compliance, DTT, and mand-training procedures) and to limit exposure of actual therapists to low-quality feedback. The supervisors watched a video of two confederates (i.e., the first author role-playing as the therapist and a confederate role-playing as a child) during the implementation of either the guided-compliance, DTT, or mand-training procedures. The supervisor was instructed to collect data on the confederate therapist's accuracy with implementing each procedure. After collecting data, supervisors were instructed to provide feedback to the confederate therapist on her implementation of the procedure (see the Procedure section).

We created a total of 18 simulated-session videos for the guided-compliance procedure. During each video, a confederate therapist and child were present in a room with a table, two chairs, and leisure items (e.g., blocks, cars, puzzle). Each video lasted approximately 2 min (range 1.6–2.4 min). In each video, the therapist provided five instructions for the confederate child to follow (e.g., “Stand up,” “Pick up car,” “Stack

blocks”) using the 10 steps of the guided-compliance procedure (see Table 2). The confederate child engaged in predetermined responses during each trial. On each trial, the confederate child either (a) complied following the therapist's initial instruction, (b) complied following the therapist's model prompt, or (c) did not comply following the therapist's initial instruction or model prompt. In each video, the confederate child responded following each prompt (i.e., verbal, model, or physical prompt) at least once. The confederate child also engaged in zero to two instances of problem behavior (e.g., aggression or property destruction) in each video.

We varied the number and type of steps that the confederate therapist implemented correctly and incorrectly across videos. During each video, the confederate therapist implemented four to six steps of the guided-compliance procedure with accuracy above 80% and four to six steps with accuracy below 80%. For example, an error for *presents instruction once* may have included the therapist repeating the verbal instruction more than once in the absence of an additional prompt (e.g., model prompt). An error for *provides a model prompt* may have included the therapist presenting a model prompt too soon after presenting the initial instruction. The confederate therapist depicted each step of the guided-compliance procedure with accuracy below 80% in 5 to 10 videos.

During probes with novel behavioral procedures, the supervisor watched a video of a confederate therapist and child during the implementation of DTT or mand-training procedures. We created a total of seven simulated-session videos for the DTT procedure. During each video, a confederate

Table 2 Steps of the guided-compliance procedure

Dependent Measures	Operational Definition
1. Attends	The trainee is facing the child and within the child's line of sight when providing the instruction.
2. Presents clear instruction	The trainee presents a brief and clear instruction that is not phrased in the form of a question and does not include any unnecessary words or the child's name.
3. Presents one-step instruction	The therapist presents only one one-step instruction at a time.
4. Presents instruction once	The therapist presents the instruction only once in the absence of an additional prompt.
5. Provides a model prompt	If the child does not comply within 5 s (± 2 s) of the instruction, the therapist repeats the instruction while modeling compliance.
6. Provides a physical prompt	If the child does not comply within 5 s (± 2 s) of the model prompt, the therapist repeats the instruction while physically guiding the child to comply with the instruction.
7. Keeps the demand in place	The therapist does not present a new instruction until the child complies with the original instruction.
8. Provides praise following compliance	The therapist provides behavior-specific praise immediately (within 2 s) following compliance to the initial instruction or the model prompt.
9. Withholds praise for a physical prompt	The therapist does not provide praise following compliance with a physical prompt.
10. Withholds praise following problem behavior	The therapist does not provide praise following compliance if it occurs within 5 s of an actual or attempted instance of problem behavior (e.g., aggression, property destruction, or self-injurious behavior).

therapist and child were seated at a table with instructional materials (e.g., target cards, data sheet, token board) and leisure items (e.g., bubbles, tablet). The therapist presented five trials using the 10 steps of the DTT procedure (see Table 3). Each video lasted approximately 5 min (range 3.6–6.0 min). The confederate child engaged in predetermined responses during each trial. On one or more trials during each video, the confederate child (a) responded correctly, (b) responded incorrectly, or (c) did not respond following the therapist's instruction. The confederate child also engaged in zero to two instances of problem behavior (e.g., aggression) in each video. During each video, the therapist implemented five to six steps of the DTT procedure with accuracy above 80% and four to five steps with accuracy below 80%. The confederate therapist depicted each of the DTT steps with accuracy below 80% in two to six videos.

We created a total of seven simulated-session videos for the mand-training procedures. During each video, the confederate therapist and child were seated on the floor with leisure items (e.g., car, puzzle, book). The therapist presented five trials

using the 11 steps of the mand-training procedure (see Table 4). Each video lasted approximately 7 min (range 6.0–8.1 min). The confederate child engaged in predetermined responses during each trial. On each trial the confederate child either (a) independently requested the item, (b) requested the item following a nonspecific prompt (e.g., “What do you want?”), (c) requested the item following a model prompt, or (d) did not request the item. The confederate child also engaged in zero to two instances of problem behavior (e.g., aggression or property destruction) in each video. During each video, the confederate therapist implemented four to six steps of the mand-training procedure with accuracy above 80% and four to five steps with accuracy below 80%. The confederate therapist depicted 10 of the 11 steps of the mand-training procedure with accuracy below 80% in two to five videos (the confederate therapist depicted *provides a model prompt* with accuracy above 80% in all videos).

Due to experimenter error, two videos depicted fewer programmed errors than planned. In one video of the guided-compliance procedure, the confederate therapist implemented

Table 3 Steps of the DTT procedure

Dependent Measures		Operational Definition
1.	Establishes ready behavior	The therapist waits to present the instruction until the child is sitting with his/her bottom in the chair, is oriented toward the therapist or instructional material, and is not engaging in any disruptive movements with his/her hands and feet.
2.	Provides instruction	The therapist delivers the instruction as specified in the child-specific protocol, without any additional words, including the child's name.
3.	Delivers reinforcer	The therapist provides praise and a token immediately following a correct response to the initial instruction (within 1 s). If tangible reinforcement is removed for error-correction trials, then the therapist provides praise only for a correct response during an error-correction trial.
4.	Delivers prompt	The therapist delivers a model or physical prompt immediately following an error (within 1 s) or following no response within the scheduled prompt delay (± 2 s).
5.	Corrects errors	Following a model or physical prompt, the therapist removes instructional materials, turns away from the child for 1 s, and the re-presents the trial. The therapist continues to re-present the trial until the child responds correctly to the initial instruction or until the therapist has re-presented the trial five times without a correct response.
6.	Exchanges token	Once the child fills his/her token board, the therapist provides praise and immediate access to a preferred tangible item (within 2 s).
7.	Provides a reinforcement duration	The therapist lets the child play with the tangible item for the correct duration (± 5 s).
8.	Ignores problem behavior	The therapist attempts to block problem behavior (e.g., prevent the child from sweeping materials off the table). Following problem behavior, the therapist does not comment on the problem behavior. If problem behavior occurs during an intertrial interval, the therapist does not delay the onset of the next demand (i.e., a demand is presented within 2 s). If problem behavior occurs when a demand is in place, the therapist does not remove the demand.
9.	Has a 2-s intertrial interval	The therapist presents the next trial within 2 s (± 2 s) of the end of the last trial or the removal of a preferred item following a reinforcement interval.
10.	Collects data	The therapist collects data following the end of one trial and before the start of the next trial.

Table 4 Steps of the Mand-training procedure

Dependent Measures		Operational Definition
1.	Presents a choice trial	If the child does not initiate play with an item (i.e., picks it up independently) within 10 s ($\pm$ 5 s), the therapist holds up one or more items within the child's view.
2.	Presents an additional choice after no selection	If the child does not reach for any toy when a choice is provided within 10 s ($\pm$ 5 s), the therapist presents a different choice of toys.
3.	Provides brief access following a selection	The therapist allows the child to play with an item that he/she picked up independently or selected from a choice trial for 10 s ($\pm$ 5 s) and then removes it from the child's reach.
4.	Presents a choice after no interaction	If at any time the child stops interacting with a selected item or attempts to access items other than the item that the therapist is restricting access to, the therapist presents another choice of toys.
5.	Provides an opportunity for an independent mand	The therapist allows 10 s ($\pm$ 5 s) for an independent mand (i.e., provides no prompt).
6.	Provides a nonspecific prompt	The therapist provides a nonspecific prompt after 10 s ($\pm$ 5 s) with no mand (e.g., "What would you like?" or "What do you need?") and allows 10 s for an independent mand.
7.	Provides a model prompt	The therapist labels the item to provide a model after 10 s ($\pm$ 5 s) with no mand following a nonspecific prompt.
8.	Responds to errors	The therapist provides a model prompt following any error (i.e., a word that does not correspond to the item or an approximation of the word that is not listed on the child's approximation sheet).
9.	Provides reinforcement	The therapist provides immediate access to the requested item for 20 s ($\pm$ 5 s) following an acceptable approximation of the mand (including spontaneous mands) based on the child-specific definitions.
10.	Responds to problem behavior	The therapist does not provide the requested item or attention within 10 s ($\pm$ 5 s) of problem behaviors, as defined by the child-specific protocols, regardless of manding.
11.	Collects data	The therapist records any spontaneous and independent mands within 10 s ($\pm$ 5 s) of occurrence in the manner appropriate to child-specific protocols (i.e., premade paper data sheets or clicker tally).

only three steps of the guided-compliance procedure with accuracy below 80% (the video should have depicted the therapist implementing at least four steps of the guided-compliance procedure with accuracy below 80%). Both Supervisor 1 (Session 1) and Supervisor 2 (Session 6) were exposed to this video during baseline. Similarly, in one of the videos of the mand-training procedure, the confederate therapist only implemented three steps of the mand-training procedure with accuracy below 80%. Supervisor 2 was exposed to the video during baseline (Session 4). We removed both videos prior to completing sessions with Supervisors 3 and 4. For the purposes of replication, simulated-session scripts are available from the corresponding author.

Experimental Design and General Procedures

We used a multiple-baseline design across two supervisors to evaluate the effects of video modeling on supervisors' acquisition and maintenance of the eight component skills of performance feedback. Due to scheduling and resource

constraints, we were only able to run two supervisors concurrently. We conducted baseline and training sessions concurrently for Supervisors 1 and 2 and Supervisors 3 and 4. We assessed each supervisor's accuracy with implementing the eight component skills when providing feedback to a confederate on her implementation of a guided-compliance procedure. Also, we assessed the extent to which supervisors' accuracy with implementing the component skills of performance feedback occurred when providing feedback to an actual therapist implementing the guided-compliance procedure and a confederate therapist implementing DTT and mand-training procedures. We conducted one to three sessions per day, 2 to 3 days per week. Each session consisted of the experimenter (a) providing the supervisor with materials to collect data on the therapist's accuracy with implementing a behavioral procedure, (b) allowing time to review those materials, (c) playing a video of a therapist implementing the behavioral procedure, and (d) providing a prompt for the supervisor to provide feedback to the therapist.

Procedure

Baseline We included baseline sessions to assess each supervisor's accuracy with implementing the eight component skills of performance feedback prior to viewing the video model. At the start of the session, the experimenter provided the supervisor with a copy of operational definitions and a data sheet to collect data on the confederate therapist's accuracy with implementing the guided-compliance procedure. We gave the supervisors 10 min (or less if they indicated they were done) to review the protocols, operational definitions, and data sheet. Next, the supervisor watched a simulated-session video (i.e., video of a confederate therapist and child during the implementation of the procedure) and collected data on the confederate therapist's accuracy with implementing the guided-compliance procedure. We randomly rotated between simulated-session videos, with no supervisor seeing the same video more than once. Within 10 min of watching the video, the experimenter instructed the supervisor to try his or her best to provide feedback to the confederate therapist on her implementation of the procedure. While the supervisor provided feedback, the confederate therapist engaged in several predetermined responses including (a) asking zero to two questions if the supervisor provided an opportunity (e.g., the supervisor asked, "Do you have any questions?") and (b) making zero to two errors when practicing the implementation of the guided-compliance procedure if the supervisor provided an opportunity to practice. The experimenter determined these responses based on a random-number list generated for each supervisor prior to the start of the study. The experimenter did not answer questions or provide feedback on the supervisor's accuracy with implementing the component skills for providing performance feedback.

Video Modeling The purpose of this condition was to evaluate the effects of video modeling on a supervisor's accuracy with implementing the component skills for providing performance feedback. During video-modeling sessions, supervisors watched a 15-min video that included a model and voice-over instruction of each component skill for providing performance feedback. Prior to beginning the video, the experimenter provided the supervisor with a completed data sheet for a simulated session of the guided-compliance procedure. The experimenter instructed the supervisor to follow along as if he or she were collecting treatment integrity data while watching a video of the confederate therapist implementing the guided-compliance procedure with the confederate child. Then, for each component skill of performance feedback, a narrator provided instruction on how to implement the skill and showed a model of correct implementation of that skill. Each model depicted a confederate supervisor providing feedback to the confederate therapist (i.e., the first author) on her

implementation of the guided-compliance procedure during the simulated-session video shown at the beginning of the video model. Throughout the video model, the confederate supervisor provided feedback on the confederate therapist's implementation of each step of the guided-compliance procedure, resulting in 10 exemplars of performance feedback using the eight component skills.

Immediately after viewing the video, we conducted a simulated-feedback session using procedures identical to baseline. That is, the experimenter provided the supervisor with operational definitions and a data sheet and then gave the supervisor 10 min to review the materials. Next, the supervisor watched a simulated-session video of the guided-compliance procedure and collected data on the confederate therapist's accuracy with implementing the procedure. Within 10 min of watching the video, the supervisor provided feedback to the confederate therapist on her implementation of the procedure. While the supervisor provided feedback, the confederate engaged in predetermined responses using procedures identical to baseline. The experimenter did not answer any questions or provide feedback to the supervisor. Our mastery criterion for terminating video-modeling sessions was mastery (i.e., accurate implementation on 80% or more opportunities) of 88% of the skills (i.e., seven out of eight skills) in a session.

Posttraining Assessment We conducted the first posttraining session 1 to 4 days after a supervisor reached our mastery criterion for video modeling. The purpose of the posttraining assessment was to demonstrate that supervisors would continue to accurately implement the eight component skills for providing performance feedback when sessions were not conducted immediately after viewing the video model. We used procedures identical to baseline. Our mastery criterion for terminating posttraining sessions was mastery of 88% of the component skills across two consecutive training sessions.

Tailored Training (Supervisor 4 Only) After conducting four posttraining sessions with Supervisor 4, she did not reach our mastery criterion. So, we conducted tailored training. During tailored training, the experimenter used the eight performance feedback component skills to provide Supervisor 4 with feedback on her implementation of the performance feedback procedure. The experimenter's feedback was based on Supervisor 4's performance during the last three posttraining sessions. During those sessions, Supervisor 4 consistently implemented five of the eight feedback component skills at mastery level (i.e., implemented accurately on 80% or more opportunities). During the tailored-training session, the experimenter provided behavior-specific praise for skills that the supervisor implemented at mastery level during the last three sessions (e.g., the experimenter said, "You are perfect at providing instructions

for how to correctly implement steps that the therapist performed with less than 80% accuracy.”).

During the last three posttraining sessions, Supervisor 4 did not consistently implement three of the eight feedback component skills at mastery level (i.e., describes incorrect performance, provides a rationale for change, and provides an opportunity for questions). For these skills, the experimenter provided a description of the incorrect performance, a rationale for changing ineffective performance, and instructions on how to implement the skill correctly. Additionally, the experimenter modeled correct implementation of these skills and had the supervisor practice implementing the component skills through role-play. The experimenter provided specific feedback to the supervisor about her performance during practice and continued to practice until the supervisor implemented all component skills with 100% accuracy. Finally, the experimenter asked the supervisor if she had any questions about the feedback she was given. Immediately after tailored training (within 10 min), we conducted a simulated-feedback session using procedures identical to baseline. Our mastery criterion for terminating tailored-training sessions was implementation of 88% of the component skills with accuracy above 80% in a session. Following tailored training, we conducted an additional posttraining assessment with Supervisor 4.

Probes with Novel Behavioral Procedures and an Actual Therapist We assessed the extent to which supervisors’ accuracy with implementing the eight component skills of providing performance feedback occurred when providing feedback to a confederate implementing novel behavioral procedures and when providing feedback to an actual therapist. We assessed each supervisor’s accuracy with implementing the components of performance feedback during baseline and following posttraining sessions. We randomized presentation of the three probes for each supervisor. The experimenter did not answer any questions or provide feedback on the supervisor’s implementation of the component skills for providing performance feedback.

To assess supervisors’ performance when providing feedback on novel behavioral procedures, we assessed each supervisor’s accuracy with implementing the eight component skills when providing feedback to a confederate therapist on her implementation of DTT and mand-training procedures. At the start of the session, the experimenter provided the supervisor with a copy of a protocol, operational definitions, and a data sheet to collect data on the confederate therapist’s accuracy with implementing the procedure (i.e., either DTT or mand training). We gave the supervisors 10 min (or less if they indicated they were done) to review the materials. Next, the supervisor watched a simulated-session video and collected data on the therapist’s accuracy with implementing the procedure. Within 10 min of watching the video, the

experimenter instructed the supervisor to try his or her best to provide feedback on the confederate therapist’s implementation of the procedure. As in baseline, the confederate therapist engaged in predetermined responses, including asking questions and making errors when practicing the implementation of DTT or mand training.

During probes with an actual therapist, the supervisor met briefly with a therapist who was working with a child with ASD, handed the actual therapist a list of five tasks (e.g., stand up, stack blocks, sit in chair), and asked him or her to try his or her best to get the child he or she was working with to complete those five tasks. Either the supervisor or the experimenter filmed the therapist while he or she instructed the child with ASD to complete the five tasks. In general, the supervisor did not say anything to the therapist or answer any questions while he or she was being filmed. However, on a few occasions the therapist indicated that he or she was done when he or she had not yet finished all five tasks, so the supervisor had to remind the therapist which of the five tasks he or she still needed to complete with the child.

Immediately after filming the therapist working with his or her client, the experimenter provided the supervisor with materials needed to collect data on the therapist’s accuracy with implementing the procedure (i.e., protocol, operational definitions, and data sheet) and gave the supervisors 10 min (or less if they indicated they were done) to review them. The supervisor then watched the video of the actual therapist implementing the guided-compliance procedures and collected data on the therapist’s accuracy with implementing the procedure. Within 10 min of watching the video, the supervisor brought the therapist to the conference room and provided feedback on his or her implementation of the guided-compliance procedure during the session. The experimenter was not present while the supervisor provided feedback to the actual therapist.

Follow-up Probes We conducted follow-up probes 1 month after a supervisor reached our mastery criterion during video modeling or tailored training (Supervisor 4 only). The purpose of follow-up probes was to assess the extent to which supervisors’ accuracy with implementing the feedback component skills maintained over time. We conducted a follow-up probe using procedures identical to baseline, during which the supervisor provided feedback to a confederate therapist on her implementation of the guided-compliance procedure. We also conducted a follow-up probe where the supervisor provided feedback to an actual therapist on his or her implementation of the guided-compliance procedure using the procedures described previously for generalization probes with an actual therapist.

Social Validity After completing posttraining sessions, supervisors completed a social validity questionnaire to assess the

social acceptability of the procedures used in this study. The questionnaire was a modified version of the Treatment Acceptability Rating Form–Revised (TARF-R) and included 10 items (Reimers, Wacker, Cooper, & de Raad, 1992). The items addressed effectiveness of the procedures, disruptiveness of the training, and the supervisors' willingness to participate in training using the procedures again. We asked supervisors to indicate their level of agreement or disagreement with each item using a 6-point Likert-type scale with higher scores on an item indicating greater agreement with the statement and acceptability of the treatment (e.g., a score of 1 indicating *strongly disagree* and 6 indicating *strongly agree*) for a range of statements (e.g., "I think this training benefited me more than harmed me," and "I would recommend this training procedure to others."). We included two open-ended questions, asking which aspects of the intervention the supervisor found most and least acceptable.

Results

The top two panels of Fig. 1 show the results of Supervisor 1. The first panel shows the percentage of mastered component skills across baseline, video-modeling, posttraining, generalization, and follow-up sessions. The second panel shows Supervisor 1's accuracy with implementing each individual component skill during each session (see Table 1 for a list of each numbered component skill). This alternative data display allows visual inspection of which component skills were implemented at mastery level (i.e., implemented accurately during 80% or more opportunities) during each session that was depicted in the first panel. Black boxes indicate a component skill the supervisor implemented accurately during 100% of opportunities during the session, striped boxes indicate a component skill the supervisor implemented with accuracy between 80% and 99%, gray boxes indicate accuracy between 50% and 79%, and white boxes indicate accuracy below 50%.

During baseline, Supervisor 1 only implemented Component Skill 1 (collects data accurately) at mastery level when providing feedback to a confederate therapist on her implementation of the guided-compliance procedure and DTT and when providing feedback to an actual therapist implementing the guided-compliance procedure. Supervisor 1 did not implement any of the component skills at mastery level when providing feedback to a confederate therapist on her implementation of the mand-training procedure during baseline. After viewing the video model, Supervisor 1 implemented 100% of the component skills at mastery level. During posttraining sessions, Supervisor 1 continued to implement a high percentage of component skills at mastery level ($M = 84\%$) and reached our mastery criterion (i.e., mastery of 88% of component skills across two consecutive sessions) following four sessions. Following training, Supervisor 1

Fig. 1 Percentage of mastered component skills for Supervisor 1 (first panel) and Supervisor 2 (third panel) across baseline, video-modeling, posttraining, and follow-up sessions. Accuracy of individual component skills (as defined in Table 1) across sessions for Supervisor 1 (second panel) and Supervisor 2 (fourth panel)

implemented 75% of component skills at mastery level when providing feedback to a confederate therapist on her implementation of novel behavioral procedures and 88% of component skills at mastery level when providing feedback to an

actual therapist. During the 1-month follow-up probe, Supervisor 1 continued to implement 88% of component skills at mastery. Due to experimenter error for Supervisor 1, a 1-month follow-up probe was not conducted with an actual therapist. Table 5 depicts training time for each supervisor. Total time required to meet mastery was 127 min for Supervisor 1.

The bottom two panels of Fig. 1 show the results for Supervisor 2. During baseline, Supervisor 2 implemented Component Skills 1 (collects data accurately) and 8 (provides an opportunity for questions) at mastery level consistently when providing feedback to a confederate therapist on her implementation of the guided-compliance procedure. Supervisor 2 implemented 50% of component skills at mastery level when providing feedback to a confederate therapist on her implementation of DTT and 38% of component skills at mastery level when providing feedback to a confederate therapist on her implementation of mand-training procedures and an actual therapist on his implementation of the guided-compliance procedure. After viewing the video model, Supervisor 2 implemented a high percentage of component skills at mastery level, implementing all component skills at mastery level except Component Skill 3 (describes incorrect performance). During posttraining sessions, Supervisor 2 continued to implement all component skills at mastery level, except for Component Skill 3. Supervisor 2 reached our mastery criterion following two posttraining sessions. Following training, Supervisor 2 implemented 75% of component skills at mastery level when providing feedback to a confederate therapist on her implementation of novel behavioral procedures (i.e., DTT and mand-training procedures) and implemented 88% of component skills at mastery level when providing feedback to an actual therapist. During the 1-month follow-up probes, Supervisor 2 implemented 88% of component skills at mastery level when providing feedback to a confederate therapist and an actual therapist. Supervisor 2 required 72 min of training (see Table 5).

The top two panels of Fig. 2 depict the results of Supervisor 3. The first panel shows the percentage of mastered component skills across sessions. The second panel shows Supervisor 3's accuracy with implementing individual component skills during each session depicted in the first panel.

During baseline, Supervisor 3 only implemented Component Skills 1 (collects data accurately), 2 (provides behavior-specific praise), and 3 (describes incorrect performance) accurately when providing feedback to a confederate therapist implementing the guided-compliance and DTT procedures. Supervisor 3 only implemented Component Skill 1 and Component Skill 3 at mastery level when providing feedback to a confederate therapist implementing the mand-training procedures or an actual therapist implementing the guided-compliance procedures during baseline. The number of component skills that Supervisor 3 implemented accurately increased during video modeling; however, he had to view the video model three times before reaching our mastery criterion during video-modeling sessions. During posttraining sessions, Supervisor 3 continued to implement a high percentage of component skills at mastery level, requiring three posttraining sessions to reach our mastery criterion.

Following training, Supervisor 3 implemented 75% of component skills at mastery level when providing feedback to a confederate therapist on her implementation of DTT procedures and an actual therapist on his implementation of the guided-compliance procedure. Supervisor 3 implemented 50% of component skills at mastery level when providing feedback to a confederate therapist implementing the mand-training procedure. During the 1-month follow-up probes, Supervisor 3 implemented 75% of component skills at mastery level when providing feedback to a confederate and 63% of component skills at mastery level when providing feedback to an actual therapist. Supervisor 3 required 160 min of training to reach our mastery criterion (see Table 5).

The bottom two panels of Fig. 2 show the results of Supervisor 4. During baseline, Supervisor 4 implemented Component Skill 1 (collects data accurately) at mastery level consistently when providing feedback to a confederate therapist on her implementation of the guided-compliance, mand-training, and DTT procedures and when providing feedback to an actual therapist on her implementation of the guided-compliance procedures. After viewing the video model, Supervisor 4 implemented an increased number of component skills at mastery level but required two viewings of the video model to meet the mastery criterion for video-modeling sessions. Supervisor 4 did not meet our mastery criterion after

Table 5 Duration of training in minutes required for each activity during video-modeling, tailored-training, and posttraining sessions across supervisors

Training Activity	Supervisor			
	1	2	3	4
Reviewing protocols and collecting treatment integrity data	74	36	80	107
Viewing video model	15	15	45	30
Tailored training				20
Providing feedback to confederate	38	21	35	52
Total Time	127	72	160	209

Fig. 2 Percentage of mastered component skills for Supervisor 3 (first panel) and Supervisor 4 (third panel) across baseline, video-modeling, posttraining, and follow-up sessions. Accuracy of individual component skills (as defined in Table 1) across sessions for Supervisor 3 (second panel) and Supervisor 4 (fourth panel)

four posttraining sessions, so we completed tailored training on the skills that she did not implement at mastery level

consistently (i.e., describes incorrect performance, provides a rationale for change, and provides an opportunity for questions). During the tailored-training session, Supervisor 4 implemented every component skill at mastery level. During posttraining sessions, Supervisor 4 continued to implement a high percentage of component skills at mastery level and reached our mastery criterion after two posttraining sessions. Following training, Supervisor 4 implemented 88% of component skills at mastery level when providing feedback to a confederate therapist on her implementation of DTT and to an actual therapist on her implementation of the guided-compliance procedure. Supervisor 4 implemented 100% of component skills at mastery level when providing feedback to a confederate on her implementation of the mand-training procedures. During the 1-month follow-up probe, Supervisor 4 continued to implement 100% of component skills at mastery when providing feedback to a confederate therapist and 88% of component skills at mastery when providing feedback to an actual therapist. Supervisor 4 required 209 min of training to reach our mastery criterion (see Table 5).

Table 6 shows the responses of each supervisor to the social validity questionnaire. On a 6-point scale, with 1 indicating lower acceptability and 6 indicating higher acceptability of the intervention, the average rating across questions was 5.26 (range 4.3–6). Supervisors 1, 2, and 4 responded to each item with an acceptability score of 4 or higher (i.e., *slightly agree*, *agree*, or *strongly agree*). Supervisor 3 responded with acceptability scores within this range as well, with the exception of one item. Supervisor 3 responded with an acceptability score of 1 (i.e., *strongly disagree*) in response to the item “I did not see any strong disadvantage to participating in this training.” Overall, ratings of treatment acceptability were high, suggesting that the supervisors found the video-modeling intervention to be a socially acceptable intervention.

Discussion

The present study evaluated the effect of watching a video model with voice-over instructions on supervisors’ accuracy with implementing a performance feedback procedure. We found that video modeling alone was sufficient to produce mastery-level responding for three of four supervisors. For the remaining supervisor (Supervisor 4), her accuracy with implementing the performance feedback procedure increased above baseline levels after watching the video model; however, she did not reach our mastery criterion until we provided tailored training. Following 20 min of tailored training, Supervisor 4 immediately reached mastery-level responding with all eight performance feedback component skills. These findings replicate previous research, suggesting that when video modeling alone is not effective, brief tailored training is

Table 6 Supervisor responses on a 6-point likert-type scale for each item of the social validity questionnaire with higher scores indicating greater agreement with the statement

Question	M	Supervisor			
		1	2	3	4
1. I liked the training procedure used.	5.3	5	6	5	5
2. The skills I learned through participating in this training will make permanent changes in the way that I implement the behavioral procedure.	5.5	6	6	5	5
3. I did not see any strong disadvantage to participating in this training.	4.3	6	6	1	4
4. I think this training benefited me more than harmed me.	6.0	6	6	6	6
5. After training, I feel more confident in my ability to accurately implement the behavioral procedure.	5.0	4	6	5	5
6. I would be willing to participate in this type of training to learn additional skills.	5.0	5	6	5	4
7. I think that the training procedures were effective at teaching me to implement the behavioral procedure.	5.0	4	6	5	5
8. I think this type of training would be appropriate for teaching clinical skills to a variety of individuals.	5.5	6	6	6	4
9. I think the training procedures used would be suitable for most clinical settings.	5.5	6	6	5	5
10. I would recommend this training procedure to others.	5.5	6	6	5	5

usually sufficient to produce mastery-level responding (e.g., Lipschultz, Vladescu, Reeve, Reeve, & Dipsey, 2015).

In previous studies, researchers have used a package intervention to teach supervisors to provide performance feedback (e.g., Parsons & Reid, 1995). Though effective, these interventions can be time-consuming and require a trainer to be present. The results of the current study replicate previous research that has demonstrated that video modeling is an efficient procedure for training staff to implement a variety of skills (e.g., Lipschultz et al., 2015; Vladescu et al., 2012). In the current study, supervisors mastered the performance feedback component skills following an average of 2.4 h (range 1.2–3.5 h) of training. During training, a trainer only had to be present during the confederate role-plays and tailored-training session (Supervisor 4 only). Supervisors required on average 37 min (range 21–52 min) of role-play, and Supervisor 4 required 20 min of tailored training before reaching our mastery criterion. In general, the training time required in the current study was brief when compared to the training time required in previous studies teaching supervisory skills, which totaled between 4 and 8 h (Jensen et al., 1998; Parsons & Reid, 1995).

Following training, we evaluated the extent to which supervisors' implementation of the performance feedback component skills occurred when providing feedback to a confederate therapist implementing novel behavioral procedures or when providing feedback to an actual therapist implementing the guided-compliance procedure. Following training, each supervisor's accuracy with implementing the performance feedback procedure increased relative to baseline. In general, all four supervisors' responding was similar when providing feedback to a confederate therapist and an actual therapist. These results suggest that once a supervisor was able to provide performance feedback to a confederate therapist, these skills also occurred when the supervisor provided feedback to an actual therapist in the absence of any additional training.

Similarly, three out of the four supervisors accurately implemented the feedback procedure when providing feedback to a confederate therapist on her implementation of DTT and mand-training procedures. Although Supervisor 3 implemented a higher percentage of component skills at mastery level relative to baseline when providing feedback on the mand-training procedures, he only implemented 63% of the component skills at mastery level. Thus, for some supervisors it might be necessary to provide additional training before a supervisor is able to provide feedback to a therapist implementing novel behavioral procedures.

Supervisors rated the social validity of the intervention fairly high, with particularly strong agreement with the statement "I think this training benefited me more than harmed me." For Supervisors 1, 2, and 4, the ratings ranged from *slightly agree* to *strongly agree* for statements such as "I think the training procedures used would be suitable for most clinical settings" and "I would recommend this training procedure to others." Supervisor 3 provided ratings similar to Supervisors 1, 2, and 4 (i.e., ranging from *slightly agree* to *strongly agree*), with the exception of one rating. Supervisor 3 indicated that he *strongly disagreed* in response to the statement "I did not see any strong disadvantage to participating in this training." At the end of the survey, supervisors were able to respond to an open-ended question regarding aspects of the training procedure that they found most and least acceptable. In this section, Supervisor 3 reported that the aspect of the study he found least acceptable was the simulated role-plays. He stated that he would have preferred conducting feedback sessions with an actual therapist.

We chose to conduct simulated sessions with a confederate therapist during training for two reasons. First, using a confederate therapist ensured that we did not expose the actual therapists in the clinic to low-quality feedback any more than necessary. Past research has suggested that exposure to

inaccurate feedback can be detrimental to acquisition of skills (Hirst et al., 2013), and so limiting exposure to low-quality feedback may be important. Second, using simulated sessions allowed us to expose the supervisor to a variety of therapist and child responses. Each simulated-session video had four to five steps that were implemented with accuracy above 80% and four to five steps that were implemented with accuracy below 80%, ensuring that the supervisor had several opportunities to implement each component skill of performance feedback. When conducting generalization probes with an actual therapist, they correctly implemented between three and five steps and incorrectly implemented between one and five steps of the procedure. Thus, the number of errors we programmed in our simulated sessions were consistent with what actual therapists were doing in the clinic.

Additionally, by varying the errors that the simulated therapist made, we ensured that the supervisor could provide feedback on a variety of errors. During probes with the actual therapist, there was little variability in child responding. If we had chosen to use actual-therapist sessions for training, this lack of variability would limit the errors that the supervisor was exposed to. For example, none of the children ever engaged in problem behavior during our probes with an actual therapist, so a supervisor would not have had the opportunity to provide feedback to a therapist who did not withhold praise following problem behavior. All supervisors implemented the performance feedback procedures with similar accuracy when providing feedback to an actual therapist, suggesting that conducting training with a confederate during simulated role-plays did not hinder the acquisition of the performance feedback skills.

There were some potential limitations to the current study that should be noted. First, during baseline sessions, we did not provide the supervisors with written instructions for how to implement the performance feedback procedure. In many staff-training studies, it is common to include written instructions in baseline to demonstrate that written instructions alone are not sufficient to produce high levels of performance. Future research should include instructions in baseline in order to demonstrate that more resource-intensive interventions, such as video modeling with voice-over instructions, are necessary to produce mastery-level performance.

Second, supervisors could reach our mastery criterion while still making errors. Specifically, a supervisor had to implement 88% (i.e., seven of eight) of component skills accurately during 80% or more opportunities to meet our mastery criterion. Our mastery criterion was consistent with those used in previous research studies evaluating methods for training supervisors to provide performance feedback. For example, Parsons and Reid (1995) used a mastery criterion of 80% of six total components correct across two consecutive sessions. Additionally, supervisors had a low number of opportunities to perform each component skill of performance feedback in a session (i.e., between one and five opportunities). So

supervisors could make at most one error (and for some skills zero errors) and still meet our mastery criterion for a given skill. Thus, we felt that an 80% mastery criterion for a given component skill was a stringent mastery criterion. However, additional research is needed to determine the optimal mastery criterion when teaching supervisors to provide performance feedback.

A third potential limitation is that we did not collect data on qualitative aspects of the feedback. Although the feedback needed to be accurate to be scored as correct, we did not assess clarity of the feedback. Anecdotally, for some supervisors, qualitative aspects of feedback improved after viewing the video model. For other supervisors, they began including the eight component skills of performance feedback (e.g., providing behavior specific praise and describing incorrect performance), but the feedback was not always clear and consumable, even following training. Future research should consider evaluating the qualitative aspects of feedback. For example, after receiving feedback, a therapist could be asked to rate the feedback in terms helpfulness or clarity.

A final limitation of the current study is that we evaluated supervisor feedback but did not evaluate the influence of that feedback on therapist responding or, more importantly, child outcomes. Studies have demonstrated that quality feedback can improve treatment integrity (e.g., Jerome et al., 2014) and that treatment integrity failures can be detrimental to child outcomes (e.g., Wilder et al., 2006). Though that was not the purpose of the current study, to truly evaluate the effectiveness of this intervention, future research needs to evaluate the extent to which teaching supervisors to provide feedback results in improved therapist performance and child outcomes. Ideally, a study should be conducted that includes measures of performance across each level (i.e., supervisor implementation of performance feedback, therapist implementation of intervention, and child's responding within the intervention).

In addition to addressing the limitations of the present study, future research should evaluate the maintenance of the component skills of performance feedback over time. Specifically, studies should evaluate what additional supports must be put in place to ensure that supervisors continue to provide feedback with high levels of accuracy. Delayed outcomes, such as improved performance of therapists or child improvements, may not be sufficient to maintain accurate implementation of the component skills. Thus, it may be necessary for clinics to establish reward systems to increase the likelihood that supervisors will provide frequent and accurate feedback to therapists. In the current study, we taught supervisors how to implement eight performance feedback component skills. We based our selection of component skills off of those skills outlined by the Behavior Analyst Certification Board and previous research (Behavior Analyst Certification Board, 2012; Parsons & Reid, 1995). However, it is not clear if all eight feedback component skills are necessary to lead to changes in therapist performance.

Additional research should conduct component analyses to examine the specific component feedback skills that are necessary to improve therapist performance.

In conclusion, this study serves as a first step in research looking at effective methods for training supervisory skills by demonstrating that video modeling can be effective at increasing supervisors' accuracy with implementing performance feedback procedures. There is still much to be done to evaluate the impact of training supervisors to provide performance feedback. Only by conducting research that examines the therapist and child outcomes can we evaluate what level of accuracy (i.e., inclusion of components rather than correctness) is required to affect behavior change and what aspects of feedback (e.g., providing behavior-specific praise, providing a demonstration of correct implementation) are effective at improving therapists' treatment integrity. Additional studies should replicate the use of video modeling to train supervisors to provide performance feedback and extend the current study by examining the impact of the training on broader clinical issues, such as treatment integrity and child outcomes.

Author Note Natalie Shuler, Department of Psychology, West Virginia University; Regina A. Carroll, Department of Psychology, University of Nebraska Medical Center's Munroe-Meyer Institute.

We would like to thank Jennifer Owsiany and Victoria DiSciullo for their assistance with creating videos for this study. This study was based on a master's thesis submitted by the first author, under the supervision of the second author, to the graduate school at West Virginia University in partial fulfillment for the requirements of a MS in psychology.

Funding The authors received no financial support for the research, authorship, and/or publication of this article.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflicts of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Behavior Analyst Certification Board. (2012). *Supervisor training curriculum outline*. Retrieved from https://www.bacb.com/wp-content/uploads/2017/09/supervisor_curriculum.pdf
- Carroll, R. A., Kodak, T., & Fisher, W. W. (2013). An evaluation of programmed treatment-integrity errors during discrete-trial instruction. *Journal of Applied Behavior Analysis*, 46, 379–394. <https://doi.org/10.1002/jaba.49>.
- DiGennaro Reed, F. D., & Henley, A. J. (2015). A survey of staff training and performance management practices: The good, the bad, and the ugly. *Behavior Analysis in Practice*, 8, 16–26. <https://doi.org/10.1007/s40617-015-0044-5>.
- Durgin, A., Mahoney, A., Cox, C., Wheetjens, B. J., & Poling, A. (2014). Using task clarification and feedback to improve staff performance in a nongovernmental organization. *Journal of Organizational Behavior Management*, 34, 122–143. <https://doi.org/10.1080/01608061.2014.914007>.
- Gil, P. J., & Carter, S. L. (2016). Graphic feedback, performance feedback, and goal setting increased staff compliance with a data collection task at a large residential facility. *Journal of Organizational Behavior Management*, 36, 56–70. <https://doi.org/10.1080/01608061.2016.1152207>.
- Green, C. W., Rollyson, J. H., Passante, S. C., & Reid, D. H. (2002). Maintaining proficient supervisor performance with direct support personnel: An analysis of two management approaches. *Journal of Applied Behavior Analysis*, 35, 205–208. <https://doi.org/10.1901/jaba.2002.35-205>.
- Higgins, W. J., Luczynski, K. C., Carroll, R. A., Fisher, W. W., & Mudford, O. C. (2017). Evaluation of a telehealth training package to remotely train staff to conduct a preference assessment. *Journal of Applied Behavior Analysis*, 50, 238–251. <https://doi.org/10.1002/jaba.370>.
- Hirst, J. M., DiGennaro Reed, F. D., & Reed, D. D. (2013). Effects of varying feedback accuracy on task acquisition: A computerized translational study. *Journal of Behavioral Education*, 22, 1–15. <https://doi.org/10.1007/s10864-012-9162-0>.
- Jensen, J., Parsons, M., & Reid, D. (1998). Supervisory training for teachers: Multiple, long-term effects in an education program for adults with severe disabilities. *Research in Developmental Disabilities*, 19, 449–463. [https://doi.org/10.1016/s0891-4222\(98\)00017-1](https://doi.org/10.1016/s0891-4222(98)00017-1).
- Jerome, J., Kaplan, H., & Sturmey, P. (2014). The effects of in-service training alone and in-service training with feedback on data collection accuracy for direct-care staff working with individuals with intellectual disabilities. *Research in Developmental Disabilities*, 35, 529–536. <https://doi.org/10.1016/j.ridd.2013.11.009>.
- Lipschultz, J. L., Vladescu, J. C., Reeve, K. F., Reeve, S. A., & Dipsey, C. R. (2015). Using video modeling with voiceover instruction to train staff to conduct stimulus preference assessments. *Journal of Developmental and Physical Disabilities*, 27, 505–532. <https://doi.org/10.1007/s10882-015-9434-4>.
- Parsons, M. B., & Reid, D. H. (1995). Training residential supervisors to provide feedback for maintaining staff teaching skills with people who have severe disabilities. *Journal of Applied Behavior Analysis*, 28, 317–322. <https://doi.org/10.1901/jaba.1995.28-317>.
- Pence, S. T., Peter, S., & C, C. (2015). Evaluation of treatment integrity errors on mand acquisition. *Journal of Applied Behavior Analysis*, 48, 575–589. <https://doi.org/10.1002/jaba.238>.
- Peterson, L., Homer, A. L., & Wonderlich, S. A. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis*, 15, 477–492. <https://doi.org/10.1901/jaba.1982.15-477>.
- Reid, D. H., Parsons, M. B., & Green, C. W. (2012). *The supervisor's guidebook: Evidence-based strategies for promoting work quality and enjoyment among human services staff*. Morganton, NC: Habilitative Management Consultants.
- Reimers, T. M., Wacker, D. P., Cooper, L. J., & de Raad, A. O. (1992). Acceptability of behavioral treatments for children: Analog and naturalistic evaluations by parents. *School Psychology Review*, 21, 628–643.
- Vladescu, J. C., Carroll, R., Paden, A., & Kodak, T. M. (2012). The effects of video modeling with voiceover instruction on accurate implementation of discrete-trial instruction. *Journal of Applied*

- Behavior Analysis*, 45, 419–423. <https://doi.org/10.1901/jaba.2012.45-419>.
- Ward-Horner, J., & Sturmey, P. (2012). Component analysis of behavior skills training in functional analysis. *Behavioral Interventions*, 27, 75–92. <https://doi.org/10.1002/bin.1339>.
- Wilder, D. A., & Atwell, J. (2006). Evaluation of a guided-compliance procedure to reduce noncompliance among preschool children. *Behavioral Interventions*, 21, 265–272. <https://doi.org/10.1002/bin.222>.
- Wilder, D. A., Atwell, J., & Wine, B. (2006). The effects of varying levels of treatment integrity on child compliance during treatment with a three-step prompting procedure. *Journal of Applied Behavior Analysis*, 39, 369–373. <https://doi.org/10.1901/jaba.2006.144-05>.