Journal of
Applied Behavior Analysis

JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2020, **53,** 449–464 NUMBER 1 (WINTER)

Enhancing the training integrity of human service staff using pyramidal behavioral skills training

TYLER G. ERATH, FLORENCE D. DIGENNARO REED, HUNTER W. SUNDERMEYER, DENYS BRAND, MATTHEW D. NOVAK AND MATT J. HARBISON

UNIVERSITY OF KANSAS

RICK SHEARS

COMMUNITY LIVING OPPORTUNITIES, INC.

This experiment used a pyramidal training model to evaluate the effects of behavioral skills training (BST), delivered in a 1-time group-training format, on the extent to which 25 human service staff implemented BST when training others how to implement behavioral procedures. Results indicated that (a) the training workshop increased BST integrity to mastery levels for the majority of participants with varying levels of education, organizational positions, and training experience, (b) the training effects generalized to teaching an untrained skill, and (c) high levels of BST integrity maintained at follow-up 4 to 6 weeks after training for all 3 participants with whom probes were conducted. Moreover, participants indicated high levels of satisfaction with both the training workshop and BST as a training procedure.

Key words: behavioral skills training, pyramidal training, staff training, training integrity

Empirically supported training can produce robust effects when used to teach human service staff how to implement a myriad of behavior-analytic procedures (Parsons & Reid, 1995; Reid, O'Kane, & Macurik, 2011; Roscoe, Fisher, Glover, & Volkert, 2006). For example, behavioral skills training (BST; Miltenberger, 2003; Parsons, Rollyson, & Reid, 2012)—a training package composed of instruction, modeling, rehearsal, and feedback—has been shown to be highly effective in training staff to implement discrete trial instruction (e.g., Geiger, LeBlanc, Hubik, Jenkins, & Carr, 2018; Sarokoff & Sturmey, 2004), guided compliance (Miles & Wilder, 2009), paired-stimulus preference assessments (e.g., Higgins, Luczynski, Carroll, Fisher, & Mudford, 2017; Lavie & Sturmey, 2002), and multiple phases of the picture exchange communication system (Martocchio & Rosales, 2016).

Although the literature contains several examples documenting the efficacy of BST as a training package, research also indicates less-than-optimal implementation of empirically supported training procedures by human service organizations in their own onboarding processes to train new staff (Reid, 1991). For example, DiGennaro Reed and Henley (2015) surveyed 382 individuals certified by or seeking certification from the Behavior Analyst Certification Board® on their self-reported experiences with staff training and supervision practices. Nearly 45% of respondents reported that preservice training was not provided to them prior to working in their position. Among the respondents for whom preservice training was provided, 81% reported that their training consisted of only verbal instructions. Few respondents reported receiving the full package of BST. Taken together, these results reveal systemic gaps in training integrity, wherein many newly hired staff are not provided with preservice training that incorporates the essential components of

Denys Brand is now affiliated with California State University, Sacramento.

Address correspondence to: Florence D. DiGennaro Reed, 4001 Dole Human Development Center, Department of Applied Behavioral Science, 1000 Sunnyside Avenue, University of Kansas, Lawrence, KS 66045. Email: fdreed@ku.edu

doi: 10.1002/jaba.608

empirically supported training. More broadly, self-reported degradations in the implementation of empirically supported training procedures may be indicative of a larger research-to-practice gap (Fixsen, Blasé, Duda, Naoom, & Van Dyke, 2010); although the extant research literature has identified evidence-based training practices (i.e., BST; Parsons et al., 2012), these practices are not consistently being applied with integrity by human service organizations.

Several potential barriers may contribute to the self-reported deficits in training integrity. Optimal training procedures may not be implemented by staff due to the total time or organizational resources required to successfully train new staff (DiGennaro Reed, Hirst, & Howard, 2013) and the need for an expert trainer to be present during all or part of initial training and to deliver performance feedback (Shapiro & Kazemi, 2017). Additionally, training may not be consistent with recommended practices due to competing contingencies in the work environment that prohibit trainers from effectively completing their training responsibilities (e.g., staffing shortages; Hewitt & Larson, 2007; Lerman, LeBlanc, & Valentino, 2015) or a lack of organizational contingencies for unsuccessful completion of job responsibilities (Budde, 1979). Thus, creating and disseminating training procedures to teach staff how to use BST and other empirically supported training procedures in their day-to-day training is of the utmost practical importance.

Various lines of recent research have sought to address these barriers. Prevention-based components, such as staff selection and preservice training, have demonstrated improvements in implementation integrity across various settings, clients, and skills (Fixsen et al., 2010; Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005). Technology-based training procedures are another line of research with a growing body of empirical support. For example, video modeling with voiceover instruction has been shown to increase staff implementation of discrete-trial instruction (Vladescu, Carroll, Paden, & Kodak, 2012), preference assessments (Deliperi, Vladescu, Reeve, Reeve, & DeBar, 2015; Delli Bovi, Vladescu, DeBar, Carroll, & Sarokoff, 2017), problem-solving interventions (Collins, Higbee, & Salzberg, 2009), and how to provide performance feedback (Shuler & Carroll, 2018). Recent research using self-instruction as a standalone training intervention has also indicated promising effects (Graff & Karsten, 2012; Hansard & Kazemi, 2018; Shapiro, Kazemi, Pogosjana, Rios, & Mendoza, 2016). Finally, researchers have documented that computer-based instruction may yield greater return on investment than BST (e.g., Geiger et al., 2018).

Another potential method for enhancing the resource efficiency of training while ensuring high levels of integrity is to integrate BST into a pyramidal training framework. Pyramidal training is a train-the-trainer approach that involves a skilled professional training a small group of staff who, in turn, train other staff within an organization (Parsons, Rollyson, & Reid, 2013; Pence, St. Peter, & Giles, 2014). Pyramidal training has been used in combination with BST to train staff to increase the integrity of preference assessments (Pence, St. Peter, & Tetreault, 2012), functional analysis conditions (Pence et al., 2014), and instructional skills (Page, Iwata, & Reid, 1982).

Researchers have shown promising evidence when using pyramidal BST to teach practitioners how to use BST to train others. In a multiple-probe design across participants, Parsons et al. (2013) used pyramidal training to teach seven teachers and two behavior technicians how to use BST. Then, the teachers and behavior technicians used BST to train other staff how to implement three behavior-change procedures (i.e., behavior-specific praise, least-to-most prompting, providing a two-item choice). Group-format BST increased the integrity of BST implementation for all participants when prompted to use the procedure to train other staff. Additionally, the training procedure

generalized to a novel skill in the staff's workplace.

Parsons et al. (2013) is the only published study to empirically evaluate ways to teach human service staff how to use BST to train others. Thus, there are boundary conditions that need to be extended, training variables that need to be more systematically analyzed, and areas to which results may not be generalizable. For example, all participants in Parsons et al. were teachers employed in a classroom-based setting for adults with intellectual and developmental disabilities; however, many human service organizations—especially those in community-based settings—may not have supervisory staff consistently present during all parts of the day. Also, wide variability exists in the levels of education of staff who work directly with clients. In a recent review of the staff training literature, Shapiro and Kazemi (2017) found that only 13.6% of participants reported a high school diploma as their highest level of education, confirming that relatively few published studies have been conducted with participants for whom a high school diploma is their highest level of education. Thus, additional research should be conducted with staff of diverse educational experiences and backgrounds to better represent the educational experiences of staff typically employed within this type of human service setting. Finally, Parsons et al. reported that participants took part in two training workshops, totaling approximately 2 hr. Thus, the literature could benefit from supplemental research that assesses ways to enhance the resource efficiency of a group training procedure to determine if one training workshop can produce equally high levels of BST implementation integrity, which would reduce the need for staff to attend two separate training workshops.

Given the paucity of research on pyramidal BST and the potentially significant impact on organizational training practices, the purpose of this investigation was to extend the literature on pyramidal BST and Parsons et al. (2013) by teaching staff with varying levels of education, organizational positions, and training experience how to use an empirically supported procedure to train others. Specifically, we employed a pyramidal training model to evaluate the effects of BST delivered in a one-time group-training format on the extent to which staff working in residential and community human service settings could be taught, in turn, to implement BST when training others how to implement a behavioral procedure.

METHOD

Participants and Setting

Twenty-five individuals working at a nonprofit agency that provides residential services to individuals with intellectual and developmental disabilities served as participants. All participants were recruited by their supervisors to take part in the study. Participants were selected by supervisors based on the criterion that they were interested in providing on-the-job training to newly hired staff within their group-home work site. Participants ranged in age from 23 to 59 years (M = 35.6 years) and had been employed by the organization for 6 months to 13 years (M = 2.4 years). At the time of the study, three participants had never provided in-home training for new employees. Eleven participants had experience providing in-home training for less than 1 year. Ten participants had experience providing in-home training for 1 or more years. Eleven participants reported their highest level of education was a high school diploma; two, eight, and three participants reported their highest level of education as an Associate's, Bachelor's, and Master's degree, respectively. One participant did not provide details about her experience. Table 1 provides demographic information about individual participants. Prior to taking part in the study and as part of their job requirements, all participants received brief training consisting of didactic instruction, rehearsal, and feedback on

how to reinforce desirable behavior. However, participants had not received training on how to train others to reinforce desirable behavior or training on how to train others using BST.

Training was conducted in a space located within the organization's main office building. The training room contained all necessary materials, including tables, chairs, two televisions, and video recording equipment. The room was divided into two sections. The *training section* was used to provide the group training and consisted of eight tables and 16 chairs that were oriented to face two televisions at the front of the room. Data collection for the experiment took place in the *data collection section* of the room, which consisted of the remaining two tables, four chairs, and video recording equipment. These materials were oriented to face the back of the room to minimize potential distractions during data collection trials. Although follow-up trials were in a different office suite located in the same building, the arrangement was like that described in the data collection section.

Table 1
Participant Education and Training Experience

Participant	Level of Education	Training Experience
Lisa	Bachelor's degree	< 6 months
Audrey	Master's degree	2 – 5 years
Macie	No response provided	No response provided
Laurie	Bachelor's degree	6 months – 1 year
Rustin	Bachelor's degree	1 – 2 years
Marty	Bachelor's degree	1 – 2 years
Maggie	High School Diploma	> 5 years
Ginger	High School Diploma	2 – 5 years
Thomas	Master's degree	< 6 months
Maynard	Bachelor's degree	None
Ani	Master's degree	> 5 years
Raymond	Bachelor's degree	6 months – 1 year
Frank	High School Diploma	2 – 5 years
Betty	Bachelor's degree	6 months – 1 year
Felicia	High School Diploma	None
Paul	High School Diploma	< 6 months
Athena	High School Diploma	6 months – 1 year
Lacey	High School Diploma	None
Katherine	Bachelor's degree	1 – 2 years
Blake	Associate's degree	1 – 2 years
Tiffany	High School Diploma	< 6 months
Teresa	High School Diploma	< 6 months
Holly	High School Diploma	> 5 years
Sylvia	High School Diploma	< 6 months
Jill	Associate's degree	< 6 months

Training Materials

Confederate staff. Two undergraduate research assistants served as confederate staff. Both individuals acted in the role of a newly hired direct support professional with no prior experience working with individuals with disabilities during all baseline and posttraining trials. Responses for the confederate staff were rehearsed prior to the study and standardized so each participant encountered 50% of trials with correct performances and 50% of trials with incorrect performances. A random number generator (randomnumbergenerator.com) was used to pseudo-randomize which trials and components contained performance errors. Research assistants received training that used a combination of instructions, modeling, rehearsal, and feedback on their performance.

Training scenarios. At the beginning of each trial, participants were presented with a 5-by-7-inch (12.7-by-17.8-cm) notecard that contained a written description of one training scenario. The purpose of the scenario was to provide the participants with a contrived practice opportunity to demonstrate how they would train staff. Each training scenario consisted of a single sentence (8 to 23 words), and described a situation in which a consumer engaged in an appropriate behavior that staff should reinforce (e.g., an individual independently put toothpaste on his toothbrush; an individual put her plate in the dishwasher without a prompt). Reinforcing desirable behavior involved (a) using a friendly tone of voice, and (b) delivering behavior-specific praise (i.e., expressing a positive statement and labeling the desirable behavior [e.g., "Good job putting toothpaste on your brush all by yourself!"]). Reinforcing desirable behavior was chosen as the target skill for participants to implement because the agency emphasized its importance in its service model and staff were familiar with how to implement the procedure; the latter

rationale permitted the study to focus on training staff how to implement BST and not the procedure to be taught using BST.

A new trial began each time a different notecard was provided to a participant. A trial concluded when participants either vocally indicated they were finished training or when no vocal response was provided for 10 s. This procedure of providing standardized training scenarios was used across all experimental conditions. We refrained from asking, "Are you all done?" or making similar statements to avoid providing accidental prompts to complete additional steps of BST.

Response Measurement

The dependent variable was participants' BST procedural integrity (i.e., the accuracy with which participants implemented BST). The five component steps required the trainer to (1) provide the trainee with instructions and a description of the target skill, (2) model the target skill for the trainee, (3) prompt the trainee to role play the target skill, (4) provide the trainee with positive and corrective feedback on performance, and (5) repeat steps 3 and 4 until the trainee performed the skill 100% correctly at least one time. Each new trial was viewed as an opportunity to implement the full BST package; thus, we asked participants to begin at Step 1 when a new notecard was presented (i.e., a new trial started).

Training integrity was calculated by dividing the number of training steps performed correctly during each trial by the total number of steps and converting this number into a percentage. The operational definitions for scoring each step as correct, incorrect or an omission are provided in Table 2. The mastery criterion was defined as a participant demonstrating 100% training integrity across two consecutive trials.

Interobserver Agreement and Procedural Integrity

A second observer collected data on training integrity from video recordings during at least 33% of trials for each experimental condition for each participant. An agreement was scored when both observers independently recorded a participant's implementation of a BST step in the same way (i.e., as correct, incorrect, or omission). A disagreement was scored when observers did not record a participant's implementation of a BST step identically (e.g., one observer recorded correct implementation of the step and the other observer recorded incorrect implementation or omission of the step). Interobserver agreement was evaluated on a step-by-step basis and was calculated by dividing the number of agreements by the number of agreements plus disagreements and converting this number into a percentage. Mean interobserver agreement across all participants was 97.5% (range, 80-100%) for baseline trials and 96.7% (range, 80-100%) for posttraining and posttraining plus feedback trials.

Procedural integrity of the independent variable was measured in vivo for 100% of trainings. An independent observer scored implementation using a procedural integrity checklist. The observer recorded whether the experimenter accurately provided vocal instructions, demonstrated all steps of the target skill, prompted participants to rehearse in dyads, and arranged opportunities for participants to receive peer feedback on their performance. Procedural integrity was calculated by dividing the number of correctly-implemented training components by the total number of training components and converting this number into a percentage. Procedural integrity was 100%.

Experimental Design and Procedure

The effects of a group training using BST were evaluated using a nonconcurrent multiple-baseline design across participants. Each participant attended a single professional development workshop during which all phases of the study were conducted except for follow-up probes. Four workshops were conducted with 2 to 10 participants attending each workshop.

Pretraining baseline. The purpose of this condition was to assess baseline levels of BST

Table 2
Behavioral Skills Training Component Checklist

1. Provide the trainee with instructions and a description of the target skill	
Correct	Describes all steps of target skill with 100% accuracy
Incorrect	(1) Describes some of the target skill steps, or (2) incorrectly describes one or more steps
Omission	Does not describe any steps of the target skill
2. Model the target skill for the trainee	
Correct	Models all steps of target skill with 100% accuracy
Incorrect	(1) Models only part/some of the steps, or (2) incorrectly models one or more steps
Omission	Does not provide a model of the target skill
3. Prompt the trainee to role play the target skill	
Correct	Prompts the trainee to practice
Incorrect	Incorrectly prompts the trainee to practice the target skill
Omission	Does not prompt trainee to practice
4. Provide trainee with performance feedback (i.e., positive and corrective feedback [FB])	
Correct	100% feedback accuracy (both positive and corrective, as needed)
Incorrect	(1) Provides only partial positive FB for correct performance, (2) provides only partial corrective FB for incorrect performance, (3) provides only positive FB and does not provide corrective FB for any steps completed incorrectly, or (4) provides only corrective FB and does not provide positive FB for any steps completed correctly
Omission	Participant does not provide the trainee with FB on performance
5. Repeat steps 3 and 4 until 100% correct performance at least one time	
Correct	Instructs the trainee to practice again until 100% accurate, or (2) does not instruct the trainee to practice again if 100% accurate
Incorrect	Prompts the trainee to practice when there was 100% accurate performance
Omission	Does not ask the trainee to re-practice when performance was less than 100%

integrity for each participant prior to training. To begin each trial, an experimenter provided the participant with a scenario notecard and gave the participant time to read the scenario. Next, the experimenter instructed the participant to train the confederate staff to reinforce the desirable behavior of the consumer described in the scenario. The experimenter stated the following instructions:

> To start, you are going to be asked to show how you would train a newly hired staff member to reinforce a desirable consumer behavior. You will do this with [insert research assistant name], who will be pretending to be a new staff member working in your home. Pretend [name] has never worked in the field and has only completed preservice training. While seated at the table, I will be handing you notecards. Each notecard contains an example scenario you can use to teach [name] how to reinforce a desirable consumer behavior. So, for this part of the training, please demonstrate how you would train [name] to reinforce a desirable consumer behavior.

During this condition, feedback was not provided, questions were not answered, and no additional instructions were given to participants. If the participant attempted to solicit additional information, the experimenter reread the instructions and told the participant to "Demonstrate how you would train a newly hired staff member to reinforce a desirable behavior based on the example scenario provided."

Training. We adopted a pyramidal training approach in which we trained a group of employees who were asked to then train confederate staff (Page et al., 1982). A single group training, lasting approximately 50 min (M = 48.5 min; range, 41–67 min), was conducted immediately following completion of baseline data collection for all participants. Training began with the primary instructor (i.e., the first or second author) conducting the didactic component of the

training. The instructor started the training by presenting the learning objectives, the importance of using an evidence-based approach to training, and the relation between training and various organizational outcomes. Next, the instructor provided a general description of BST and its use as an evidence-based approach.

The instructor followed this introduction by providing a detailed vocal description of BST accompanied by a PowerPoint presentation (i.e., vocal and written instructions). The procedure for training each component of BST consisted of presenting a definition, three modeled exemplars, and recommended practice guidelines for implementation. Modeled exemplars included demonstrations of the target skill in three different role-play scenarios. Each role play consisted of an experimenter playing the role of a trainer and another experimenter playing the role of a trainee. Each scenario began with a brief description that provided background information on the role play and the desirable behavior to be reinforced, followed by demonstrations of correct implementation of the training step. The practice guidelines consisted of supplemental information regarding the BST component and information related to real-world implementation based on recommended practice. For example, the recommended practice guidelines for modeling included vocal and written instructions where participants were instructed to (a) model the entire procedure, (b) standardize models across training episodes, (c) provide multiple examples, and (d) supplement modeling with other training procedures to maximize effectiveness. Questions from participants were answered throughout the training.

In the final part of training, the instructor prompted participants to rehearse in peer dyads and provide feedback to one another on their performance. This part of the training procedure consisted of one participant assuming the role of *trainer* and the other participant assuming the role of *staff*. The staff was instructed to respond to the trainer's instructions while simultaneously observing the trainer's performance. Staff then provided positive feedback for steps performed well and corrective feedback for steps that could be improved. When this process was complete for the first trainer, the roles were switched and the process was repeated. Training ended when both participants had played both roles and reported they were comfortable with the training procedure. We arranged practice and feedback in this manner because the allotted time for the workshop, its structure, and number of attendees made it challenging for the instructor to conduct 1:1 training and feedback.

Posttraining. The purpose of this condition was to assess BST integrity for each participant after taking part in the training workshop. Data collection occurred immediately after dyad rehearsal once each participant (a) played the role of both trainer and trainee and (b) vocally reported they were comfortable with the training procedure. Trials were conducted using the same procedure as the pretraining baseline.

Experimenter feedback. The experimenter provided brief performance feedback if a participant's BST integrity was less than 100% for three consecutive trials during posttraining. However, Blake and Teresa each received feedback after two consecutive trials below 100% due to experimenter error. The content included a vocal description of the step(s) a participant performed correctly and step(s) a participant performed incorrectly, as well as a description of how the step(s) should have been implemented. Performance feedback ended when a participant either demonstrated performance that met the mastery criterion or when the experimenters determined that more intensive training was needed outside of this experiment (i.e., visual analysis indicated acquisition was poor and not improving).

Generality of training effects. The experimenters conducted single-trial probes during baseline and posttraining to assess the generality of the training procedure with an untrained skill (i.e., teaching staff how to provide a choice). The procedure was similar to baseline, except that participants were asked to train confederate staff to

provide a choice to consumers. A choice consisted of presenting two or more options to a consumer and delivering the selected option to the consumer. All participants were provided with two pieces of individually wrapped bite-size candy to use as the teaching items for demonstrating how to provide a choice. Training integrity was calculated by dividing the number of training steps performed correctly by the total number of steps and converting this number into a percentage.

Follow-up. Given staff schedules, availability, and work locations (which spanned multiple counties within the state), follow-up was conducted with three participants 4 to 6 weeks after training. Each participant was asked to train a new confederate staff how to reinforce a desirable behavior and provide choice. Follow-up probes occurred in an office suite located within the same building as the training room.

Social Validity

When participants completed the study, they were asked to fill out a modified version of the Intervention Rating Profile-15 (IRP-15; Martens, Witt, Elliot, & Darveaux, 1985) to assess the social validity of the training procedures and utility of BST. The nine-item measure consisted of six items that asked participants to rate the acceptability of the training workshop and BST as a training procedure on a six-point Likert-type scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). Higher scores represent higher levels of participant satisfaction for all six items. The remaining three items were open-ended questions in which participants could provide qualitative feedback on the training workshop and BST as a training procedure. The experimenter asked participants to complete the questionnaires and place them in an unmarked manila envelope at the back of the training room when they concluded their sessions. The experimenter vocally stated that responses would be confidential and only used to inform training practices and future workshops conducted at the agency. The questionnaire did not include names or any identifiable information.

RESULTS

Figures 1–2, and 3 display the percentage of BST steps performed correctly (primary *y*-axis) and the error patterns for each BST component (secondary *y*-axis boxplots). Shaded boxes indicate correct implementation of the BST component; open boxes specify on which step implementation omission and/or commission errors occurred.

Figure 1 summarizes the data for the 10 participants (40%) who met the mastery criterion after the group training. During baseline trials and the baseline generality probe, mean BST integrity was 10% (range, 0–60%). Error analysis data indicate that participants did not correctly implement most of the BST components; if correct implementation of a BST component occurred, it typically involved step 2 (modeling the target skill). Across all participants who met the mastery criterion after group training, 75.4% and 24.6% of incorrectly implemented steps were omission errors and commission errors, respectively. Following the group training, mean BST integrity increased to 96% (range, 40–100%), and 9 of the 10 participants met the mastery criterion after the minimum number of trials to mastery (i.e., two trials). An error analysis revealed that participants correctly implemented all five BST steps, as indicated by the shaded boxes following the phase change. Additionally, BST integrity on the post-training probe assessing the generality of the training effects was 80% to 100% for all 10 participants. Finally, follow-up probes conducted with Maggie and Maynard indicated maintenance of the training effects at 4 to 6 weeks following training. BST integrity scores for both the primary dependent variable and the probe assessing the generality of the training effects were 100% for both participants.

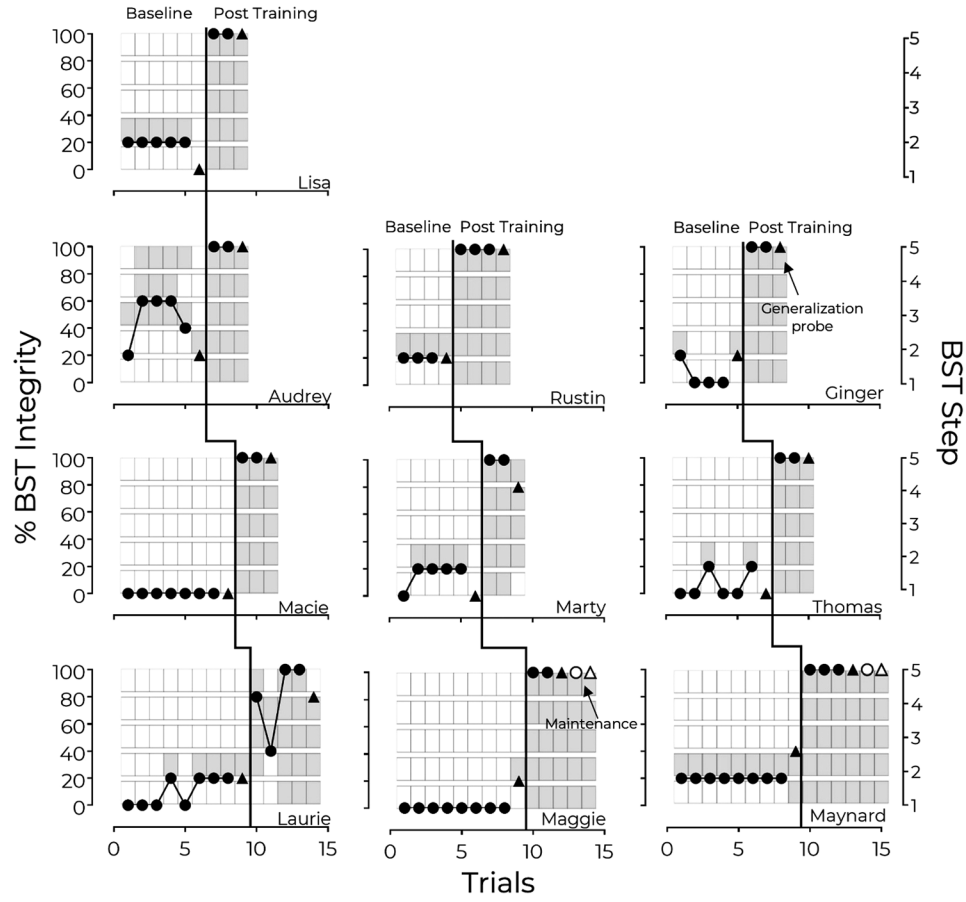

Figure 1. Percent BST integrity for (left *y*-axis) and BST steps implemented by (right *y*-axis) participants who achieved mastery after training. Shaded boxes indicate steps that were implemented correctly, and open boxes indicate steps that were either incorrectly implemented or omitted (right *y*-axis).

Figure 2 depicts the data for the 10 participants (40%) who required the group training plus supplemental experimenter feedback to achieve mastery levels of BST integrity. During baseline trials and the baseline generality probe, mean BST integrity was 5.4% (range, 0–20%). The error analysis indicated that eight participants did not correctly implement any of the BST components during this condition; the remaining two participants consistently implemented step 2 (modeling the target skill). For all participants who required supplemental experimenter feedback to reach the mastery criterion, 83.6% and 16.4% of incorrectly implemented steps were omission errors and commission errors, respectively. Following training, performance increased for all participants (M = 55%; range, 20–100%). Despite a posttraining increase in BST integrity, participants did not demonstrate mastery levels of performance. The most commonly missed BST step was step 1 (providing instructions on the target skill). Brief performance feedback was subsequently delivered, which increased performance to mastery levels (M = 85%; range, 20–100%). Integrity on the posttraining probe assessing the generality of the training effects was 100% for six participants, with the remaining four participants' scores between 40% and 80% integrity. Finally, follow-up probes

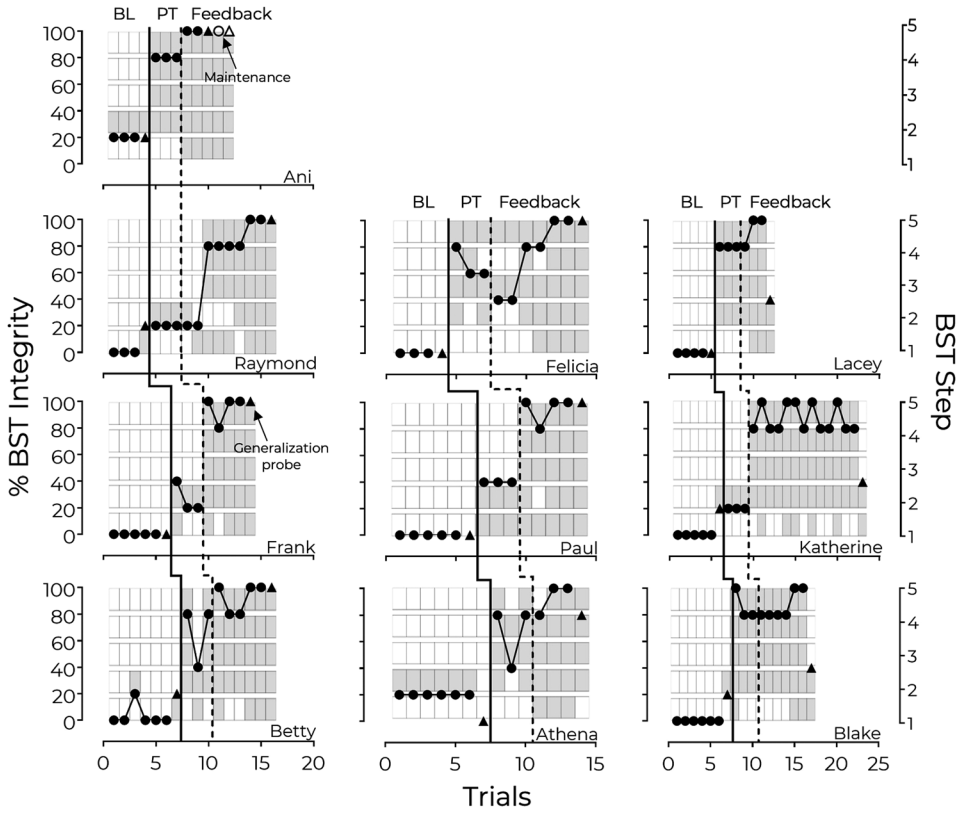

Figure 2. Percent BST integrity for (left *y*-axis) and BST steps implemented by (right *y*-axis) participants who required supplemental feedback following training. Shaded boxes indicate steps that were implemented correctly, and open boxes indicate steps that were either incorrectly implemented or omitted (right *y*-axis). BL = pretraining baseline; PT = posttraining.

conducted with Ani showed maintenance of the training effects, as demonstrated by BST integrity scores of 100% for both the primary dependent variable and the probe assessing generality.

Figure 3 displays the data for the remaining five participants (20%) who did not demonstrate mastery levels of BST integrity following the training workshop and brief performance feedback. Mean BST integrity during baseline trials and the baseline generality probe for participants in this group was 2.2% (range, 0–20%). Error analysis indicated that three participants did not correctly implement any of the BST components; the remaining two participants correctly implemented step 2 (modeling the target skill) intermittently on trials. Following the group-based training, performance was found to increase to near mastery levels for one participant (Teresa; M = 86.7%; range, 80-100%), with minimal to no increase in BST integrity for the other four participants (M = 8.3%; range, 0-20%). With the addition of brief performance feedback, mean BST integrity increased to 37.1% (range, 0–100%). No increase in BST integrity was found for two participants. For the remaining three participants (Teresa, Sylvia, and Jill), BST integrity did increase (range, 60-80%) above baseline, but not to mastery. Error analysis indicated that the most commonly missed BST step for these three participants was step 1 (providing instructions on the target skill).

Overall, participants found the group-training workshop to be an acceptable training modality and BST to be a useful way to train new staff (see

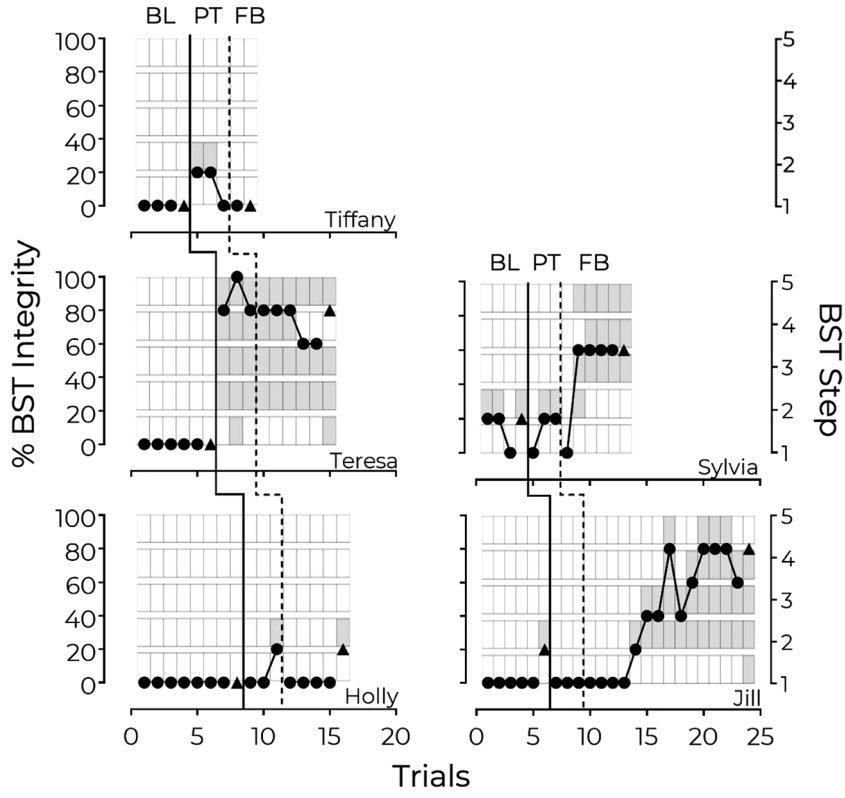

Figure 3. Percent BST integrity for (left *y*-axis) and BST steps implemented by (right *y*-axis) participants who did not meet mastery criterion. Shaded boxes indicate steps that were implemented correctly, and open boxes indicate steps that were either incorrectly implemented or omitted (right *y*-axis). BL = pretraining baseline; PT = posttraining; FB = performance feedback.

Table 3). Mean ratings for all six Likert-type questions were between 5 (*agree*) and 6 (*strongly agree*). With regard to the social validity of the procedures (i.e., the training workshop), participants indicated that (a) the group training was an acceptable format to help trainers accurately train new staff (M = 5.2; range, 4–6), (b) the training workshop was beneficial for teaching them how to train new staff (M = 5.1; range, 4–6), and (c) they liked the procedures used in the training workshop (M = 5.2; range, 4–6). With regard to the social validity of the outcomes (i.e., using BST to train staff), participants indicated that (a) BST is an acceptable way to train new residential staff (M = 5.4; range, 4–6), (b) most trainers would find BST beneficial for training new residential staff (M = 5.3; range, 4–6), and (c) they liked using BST to train new staff (M = 5.3; range, 3–6).

The most common open-ended response about the training workshop was that staff found the instructor-modeling and rehearsal components helpful in teaching them how to implement BST (n = 16). The most common open-ended response about BST as a training procedure was that staff found the rehearsal, feedback, and rerehearse until mastery components to be the most valuable additions for more effectively training new staff on the job (n = 11). Four respondents noted that requiring trainers to conduct practice opportunities and deliver feedback until the trainee reaches 100% accuracy could be challenging to implement when arranged during real-world training in the group homes.

Table 3
Modified IRP-15 Results

Question	Mean	Range
This training workshop would be an acceptable way to help trainers accurately train new staff.	5.2	4-6
Most trainers would find this training workshop beneficial for teaching them how to train new staff.	5.1	4-6
I like the procedures used today to assist me in training new staff.	5.2	4-6
Behavioral skills training is an acceptable way for trainers to train new staff.	5.4	4-6
Most trainers would find behavioral skills training beneficial for training new staff.	5.3	4-6
I like the procedures (behavioral skills training) that I will now use to train new staff.	5.3	3-6

Note. 1 (*strongly disagree*) to 6 (*strongly agree*)

DISCUSSION

The broad purpose of this experiment was to find a sustainable solution to help bridge the research-to-practice gap and deliver evidence-based procedures, such as BST, into the hands of everyday practitioners. To accomplish this purpose, we examined the effects of BST delivered in a group workshop format on the degree to which human service staff with varying levels of education and supervisory experience could be trained to implement BST with integrity. Most participants responded positively to the training procedure. Out of 25 participants, 10 participants demonstrated mastery levels of performance after taking part in the group-training workshop, and an additional 10 participants demonstrated mastery after the training workshop plus brief performance feedback. Follow-up probes conducted 4 to 6 weeks after training indicated maintenance of training effects for all three participants with whom probes were conducted.

Interestingly, five of 25 participants did not correctly implement BST despite receiving training and feedback. Their levels of education and experience were similar to other participants for whom training and feedback were effective, and they did not communicate concerns or questions during the experiment that might explain their performance. Unfortunately, the variables responsible for their poor performance are unknown. The finding that 20% of our participants did not demonstrate improvement to mastery underscores the importance of measuring staff behavior, including supervisory or training activities. These data provide some evidence that a desire to assume supervisory or training responsibilities does not necessarily ensure acquisition and competent performance of those skills.

The results of the current study have several important implications for the staff training literature and for behavior analysts working in human service settings. First, this study extends the literature on pyramidal training in general (Page et al., 1982; Pence et al., 2014), and pyramidal BST in particular (Parsons et al., 2013), by providing a methodological framework for experimentally evaluating BST implementation integrity as the targeted variable under investigation. Second, a portion of the findings also provide preliminary evidence that training containing *expert* provision of the full package of BST may not be necessary to teach all human service staff how to implement BST with high integrity. This suggestion is substantiated by the finding that 10 participants demonstrated mastery levels of BST implementation after the experimenters had only provided instructions and modeling. Participants experienced rehearsal and feedback in peer dyads, which did not involve input from the experimenters. Although similar results have been found in related studies involving staff (e.g., Moore & Fisher, 2007; Ward-Horner & Sturmey, 2010) and parents (Dogan et al., 2017) or utilizing technology (e.g., Geiger et al., 2018), to the best of our knowledge this is the first study to document this effect when teaching human service staff how to use BST to train others in a job-related skill.

Third, these findings provide additional support for and extend the boundary conditions of pyramidal BST as an effective and efficient

training procedure within human service settings (Parsons et al., 2013). Training delivered in a group-based format was used to teach staff with varying levels of education and supervisory experience how to use BST to train others. Moreover, each participant only took part in one training workshop lasting approximately 50 min. Thus, these results demonstrate that it is possible for human service organizations to teach a sizeable number of staff how to use BST, while taking into consideration organizational time and resources expended on the training process.

With regard to application, the training procedure used in the current study has several implications for behavior analysts and consultants operating in human service settings. First, this study demonstrates one possible approach for enhancing staff training practices in general, and the training repertoires of staff operating in any type of supervisory role in particular. Second, the training was delivered by behavioral consultants external to the organization, and specifically targeted teaching staff in or aspiring to residential supervisory roles how to use a performance- and competency-based training procedure (i.e., BST). Therefore, this procedure could be conceptualized as an organization's training system (Lerman et al., 2015) and one possible modality for maximizing the provision of consultative services due to its use of procedures to train staff a skill with a high level of practical utility and broad application.

There are several limitations that should be addressed in future research. First, more research is needed on the relation between staff implementation of BST and skill complexity. It is possible that we would have obtained different results if participants were asked to implement BST to train a more complex skill. Second, we conducted the study in a training room at the agency's main office and not in the group homes scattered across the state where participants worked. This decision may have limited the generality of the training effects to other untrained settings. We observed variability in the outcomes of our generalization probes to an untrained skill; it is possible that generalization to an untrained setting (i.e., group homes) would show similar variability. Therefore, future research may evaluate the effects of the training setting on acquisition, maintenance, and generalization of BST implementation integrity and identify ways to promote generalization across settings. Third, we trained participants how to use BST to teach a particular skill (i.e., reinforcing a desirable behavior). Consequently, there are limitations to the generality of the training effects to other skills not directly taught. However, probes to assess generality of the training effects to an untrained skill (i.e., providing choice) were quite encouraging, though variable. Moreover, follow-up probes also indicated maintenance of training effects for all three participants with whom probes were conducted. Although the results of this study provide supplemental evidence for the external validity of pyramidal BST as an efficacious training procedure (Parsons et al., 2013), future research on this topic is warranted to better understand how and in what ways human service organizations can maximize the provision of the training, as well as ways to enhance the generality of the training effects to on-the-job training provided by staff.

An important direction for future research would be to conduct an experimental analysis of the variables controlling BST skill acquisition. One approach could entail a component analysis of BST to determine which components are necessary within the pyramidal model for successful skill acquisition (Ward-Horner & Sturmey, 2010). For example, it may be that only modeling, practice, and feedback are necessary for acquisition, and trainers could omit instructions to enhance resource efficiency. Researchers could also evaluate a sequential approach to pyramidal training (Howard & DiGennaro Reed, 2014; Severtson & Carr, 2012). For example, future research could systematically evaluate the effects of less resource-intensive components of BST (i.e., instructions and modeling), followed by delivery of supplemental BST components

(i.e., role play and performance feedback) as necessary and for as many trials as needed.

Future research should also be conducted to further enhance the resource efficiency of the current training model by evaluating the role of technology-based, antecedent training modalities (e.g., video modeling and interactive computer training; Delli Bovi et al., 2017; Geiger et al., 2018; Hansard & Kazemi, 2018) on BST acquisition. Finally, experimenter error resulted in the early delivery of feedback to two participants (Blake and Teresa). We are unable to determine what effects, if any, this decision had on our interpretation of the results.

Despite these limitations, this study provides support for pyramidal BST as a resource efficient training model that can be used in human service settings to teach staff with varying levels of education, organizational positions, and training experience how to use an empirically supported procedure to provide training. For human service organizations to provide the highest quality of services, maximize consumer outcomes, and ensure new staff success, it is imperative to optimize the effectiveness and efficiency of all training practices, including preservice and on-the-job training (van Oorsouw, Embregts, Bosman, & Jahoda, 2009). Thus, implementation and integration of a pyramidal BST model into an organization's current training structure is one potential modality to enhance organizational training practices that may also help bridge the research-to-practice gap (Fixsen et al., 2010).

REFERENCES

Budde, J. F. (1979). *Measuring performance in human service settings: Planning, organization, and control*. New York: AMACOM.

Collins, S., Higbee, T. S., & Salzberg, C. L. (2009). The effects of video modeling on staff implementation of a problem-solving intervention with adults with developmental disabilities. *Journal of Applied Behavior Analysis*, *42*, 849–854. https://doi.org/10.1901/jaba.2009.42-849.

Deliperi, P., Vladescu, J. C., Reeve, K. F., Reeve, S. A., & DeBar, R. M. (2015). Training staff to implement a paired-stimulus preference assessment using video modeling with voiceover instruction. *Behavioral Interventions*, *30*, 314–332. https://doi.org/10.1002/bin.1421.

Delli Bovi, G. M., Vladescu, J. C., DeBar, R. M., Carroll, R. A., & Sarokoff, R. A. (2017). Using video modeling with voice-over instruction to train public school staff to implement a preference assessment. *Behavior Analysis in Practice*, *10*, 72–76. https://doi.org/10.1007/s40617-016-0135-y.

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.

DiGennaro Reed, F. D., Hirst, J. M., & Howard, V. J. (2013). Empirically supported staff selection, training, and management strategies. In D. D. Reed, F. D. DiGennaro Reed, & J. K. Luiselli (Eds.), *Handbook of crisis intervention for individuals with developmental disabilities* (pp. 71–86). New York: Springer.

Dogan, R. K., King, M. L., Fischetti, A. T., Lake, C. M., Matthews, T. L., & Warzak, W. J. (2017). Parent-implemented behavioral skills training of social skills. *Journal of Applied Behavior Analysis*, *50*, 805–818. https://doi.org/10.1002/jaba.411.

Fixsen, D. L., Blasé, K. A., Duda, M. A., Naoom, S. F., & Van Dyke, M. (2010). Implementation of evidence-based treatments for children and adolescents: Research findings and their implications for the future. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 435–450). New York: Guilford Press.

Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). Research on core implementation components. In D. L. Fixsen, S. F. Naoom, K. A. Blase, R. M. Friedman, & F. Wallace (Eds.), *Implementation research: A synthesis of the literature* (pp. 35–55). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231.

Geiger, K. B., LeBlanc, L. A., Hubik, K., Jenkins, S. R., & Carr, J. E. (2018). Live training versus e-learning to teach implementation of listener response programs. *Journal of Applied Behavior Analysis*, *51*, 220–235. https://doi.org/10.1002/jaba.444.

Graff, R. B., & Karsten, A. M. (2012). Evaluation of a self-instruction package for conducting stimulus preference assessments. *Journal of Applied Behavior Analysis*, *45*, 69–82. https://doi.org/10.1901/jaba.2012.45-69.

Hansard, C., & Kazemi, E. (2018). Evaluation of video self-instruction for implementing paired-stimulus preference assessments. *Journal of Applied Behavior Analysis*, *51*, 675–680. https://doi.org/10.1002/jaba.476.

Hewitt, A., & Larson, S. (2007). The direct support workforce in community supports to individuals with developmental disabilities: Issues, implications, and promising practices. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*, 178–187. https://doi.org/10.1002/mrdd.20151.

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.

Howard, V. J., & DiGennaro Reed, F. D. (2014). Training shelter volunteers to teach dog compliance. *Journal of Applied Behavior Analysis*, *47*, 344–359. https://doi.org/10.1002/jaba.120.

Lavie, T., & Sturmey, P. (2002). Training staff to conduct a paired-stimulus preference assessment. *Journal of Applied Behavior Analysis*, *35*, 209–211. https://doi.org/10.1901/jaba.2002.35-209.

Lerman, D. L., LeBlanc, L. A., & Valentino, A. L. (2015). Evidence-based application of staff and caregiver training procedures. In H. S. Roane, J. E. Ringdahl, & T. S. Falcomata (Eds.), *Clinical and organizational applications of applied behavior analysis* (pp. 321–351). San Diego, CA: Elsevier.

Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology: Research and Practice*, *16*, 191–198. https://doi.org/10.1037/0735-7028.16.2.191.

Martocchio, N., & Rosales, R. (2016). An evaluation of pyramidal training to teach implementation of the picture exchange communication system. *Behavioral Interventions*, *31*, 265–282. https://doi.org/10.1002/bin.1448.

Miles, N. I., & Wilder, D. A. (2009). The effects of behavioral skills training on caregiver implementation of guided compliance. *Journal of Applied Behavior Analysis*, *42*, 405–410. https://doi.org/10.1901/jaba.2009.42-405.

Miltenberger, R. G. (2003). *Behavior modification: Principles and procedures*. Belmont, CA: Wadsworth Publishing.

Moore, J. W., & Fisher, W. W. (2007). The effects of videotape modeling on staff acquisition of functional analysis methodology. *Journal of Applied Behavior Analysis*, *40*, 197–202. https://doi.org/10.1901/jaba.2007.24-06.

Page, T. J., Iwata, B. A., & Reid, D. H. (1982). Pyramidal training: A large-scale application with institutional staff. *Journal of Applied Behavior Analysis*, *15*, 335–351. https://doi.org/10.1901/jaba.1982.15-335.

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.

Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-based staff training: A guide for practitioners. *Behavior Analysis in Practice*, *5*, 2–11. https://doi.org/10.1007/BF03391819.

Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2013). Teaching practitioners to conduct behavioral skills training: A pyramidal approach for training multiple human service staff. *Behavior Analysis in Practice*, *6*, 4–16. https://doi.org/10.1007/BF03391798.

Pence, S. T., St. Peter, C. C., & Giles, A. F. (2014). Teacher acquisition of functional analysis methods using pyramidal training. *Journal of Behavioral Education*, *23*, 132–149. https://doi.org/10.1007/s10864-013-9182-4.

Pence, S. T., St. Peter, C. C., & Tetreault, A. S. (2012). Increasing accurate preference assessment implementation through pyramidal training. *Journal of Applied Behavior Analysis*, *45*, 345–359. https://doi.org/10.1901/jaba.2012.45-345.

Reid, D. H. (1991). Technological behavior analysis and societal impact: A human services perspective. *Journal of Applied Behavior Analysis*, *24*, 437–439. https://doi.org/10.1901/jaba.1991.24-437.

Reid, D. H., O'Kane, N. P., & Macurik, K. M. (2011). Staff training and management. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 281–294). New York: Guilford Press.

Roscoe, E. M., Fisher, W. W., Glover, A. C., & Volkert, V. M. (2006). Evaluating the relative effects of feedback and contingent money for staff training of stimulus preference assessments. *Journal of Applied Behavior Analysis*, *39*, 63–77. https://doi.org/10.1901/jaba.2006.7-05.

Sarokoff, R. A., & Sturmey, P. (2004). The effects of behavioral skills training on staff implementation of discrete-trial teaching. *Journal of Applied Behavior Analysis*, *37*, 535–538. https://doi.org/10.1901/jaba.2004.37-535.

Severtson, J. M., & Carr, J. E. (2012). Training novice instructors to implement errorless discrete-trial teaching: A sequential analysis. *Behavior Analysis in Practice*, *5*, 13–23. https://doi.org/10.1007/BF03391820.

Shapiro, M., & Kazemi, E. (2017). A review of training strategies to teach individuals implementation of behavioral interventions. *Journal of Organizational Behavior Management*, *37*, 32–62. https://doi.org/10.1080/01608061.2016.1267066.

Shapiro, M., Kazemi, E., Pogosjana, M., Rios, D., & Mendoza, M. (2016). Preference assessment training via self-instruction: A replication and extension. *Journal of Applied Behavior Analysis*, *49*, 794–808. https://doi.org/10.1002/jaba.339.

Shuler, N., & Carroll, R. A. (2018). Training supervisors to provide performance feedback using video modeling with voiceover instructions. *Behavior Analysis in*

Practice. https://doi, https://doi.org/10.1007/s40617-018-00314-5.

van Oorsouw, W. M., Embregts, P. J., Bosman, A. M., & Jahoda, A. (2009). Training staff serving clients with intellectual disabilities: A meta-analysis of aspects determining effectiveness. *Research in Developmental Disabilities*, *30*, 503–511. https://doi.org/10.1016/j.ridd.2008.07.011.

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. (2010). Component analyses using single-subject experimental designs: A review. *Journal of Applied Behavior Analysis*, *43*, 685–704. https://doi.org/10.1901/jaba.2010.43-685.

Received August 2, 2018
Final acceptance April 15, 2019
Action Editor, Corey Stocco