Journal of
Applied Behavior Analysis

JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2018, **9999,** n/a–n/a NUMBER 9999 ()

# *LIVE TRAINING VERSUS E-LEARNING TO TEACH IMPLEMENTATION OF LISTENER RESPONSE PROGRAMS*

KANEEN B. GEIGER

AUBURN UNIVERSITY

LINDA A. LEBLANC, KATIE HUBIK AND SARAH R. JENKINS

TRUMPET BEHAVIORAL HEALTH

AND

JAMES E. CARR

BEHAVIOR ANALYST CERTIFICATION BOARD

Discrete-trial teaching is an effective teaching procedure that must be implemented with high integrity to produce optimal learning. Behavioral Skills Training (BST) has proven effective for staff training; however, BST is time and labor intensive. Computer-based instruction (CBI) programs may provide a more efficient and cost-effective alternative to live training if the CBI program is as effective as BST in producing accurate implementation. The current study compared CBI to BST to train novice undergraduate students to conduct discrete-trial teaching. Participants were randomly assigned to one of the two conditions and assessed prior to and after the completion of training. Results indicated that although both BST and CBI were effective at training participants to implement discrete-trial teaching, BST was slightly but significantly more effective whereas CBI quickly created a return on the investment of product development.

*Key words:* autism, behavioral skills training, computer-based instruction, conditional discrimination training, discrete trial teaching, e-learning, listener responding, return on investment, staff training

Applied behavior analysis (ABA) has consistently produced empirical evidence of significant improvements in overall intellectual and adaptive functioning for children with autism (Eldevik et al., 2009; Green, 1996; Rogers & Vismara, 2008). Discrete-trial teaching (DTT) is a commonly used ABA intervention technique for children with autism (Leaf & McEachin, 1999; Smith, 2001). A discrete trial lasts approximately 5 to 20 s, and trials are presented rapidly to maximize the learning opportunities in each teaching session (Smith, 2001). Each trial consists of five components: the discriminative stimulus ($S^D$), the prompt, the learner's response, the consequence, and the intertrial interval (Leaf & McEachin, 1999; Smith, 2001). A high level of precision is required to implement discrete-trial procedures well (LeBlanc, Gravina, & Carr, 2009), and many staff members and parents struggle to achieve high levels of procedural integrity (e.g., Johnson & Hastings, 2002; Symes,

This project served as the dissertation of the first author and was completed in partial fulfillment of the requirements for the PhD degree. This article does not represent an official position of the Behavior Analyst Certification Board.

Kaneen Geiger is now at Ivymount School. The author would like to thank Eric Fox for his assistance with programming the computer-based program and Cassandra Clark, Chelle Harbin, Leah Horner, Whitney McWherter, and Taylor Sweatt for their assistance with data collection. A version of the instructional module with certain flexible navigation features and hint features enabled is available at no charge at Foxylearning.com. A version of this program is now included in the curriculum at Trumpet Learning Center which is commercially available for training for Registered Behavioral Technicians.

Address correspondence to Kaneen Geiger. E-mail: kaneen.b.geiger@gmail.com

doi: 10.1002/jaba.444

Remington, Brown, & Hastings, 2006), which can limit the effectiveness of the teaching procedure.

Researchers have experimentally examined the effects of errors of commission during error correction for DTT and found higher treatment integrity resulted in better acquisition whereas lower treatment integrity resulted in poor acquisition (DiGennaro Reed, Reed, Baez, & Maguire, 2011; Grow et al., 2009). Incorrect implementation of discrete-trial procedures could lead to faulty stimulus control, false skill mastery, prompt dependence, and problem behavior, all of which can stall the learner's progress and waste time and money. Auditory–visual conditional discriminations involve the teacher presenting an auditory stimulus (e.g., "Dog") followed by an opportunity for the learner to select the corresponding stimulus (e.g., a picture of a dog) from an array of comparison stimuli (Green, 2001). Thus, for the learner to correctly respond, both the auditory and the visual stimuli must exert stimulus control over the selection response.

Many researchers have documented difficulties in teaching conditional discriminations (e.g., Harrison & Green, 1990; Johnson & Sidman, 1993; McIlvane, Dube, Kledaras, & Iennaco, 1990) to individuals with autism or intellectual disabilities. Optimal procedures for implementing conditional discrimination training have been described in the literature (Green, 2001; Grow & LeBlanc, 2013). These recommendations include having discrete trials with different target (i.e., sample) stimuli across trials, varying the position of the comparison stimuli across trials, minimizing inadvertent instructor cues, and using errorless teaching (i.e., most-to-least prompting) with frequent assessment probes to allow rapid prompt fading (Green, 2001; Libby, Weiss, Bancroft, & Ahearn, 2008; MacDuff, Krantz, & McClannahan, 2001). Hence, teaching conditional discriminations requires a high degree of precision to be implemented correctly.

Several variables can negatively impact the quality of DTT delivery. First, instructors for children with autism must be competent and fluent in delivery of DTT under stressful conditions (e.g., occurrence of problem behavior during instruction; Symes et al., 2006). Second, the typical levels of education and wages in early intensive behavioral intervention (EIBI) settings are low for the level of required instructional precision, which leads to high levels of staff turnover (LeBlanc et al., 2009). High staff turnover in EIBI settings can be costly due to the extensive training required to prepare a new staff member to effectively implement DTT with a child with autism and can lead to reduced access to therapeutic services for the child (Larson & Hewitt, 2005; Smith, 2001).

The most prominent evidence-based training package is Behavioral Skills Training (BST; Miltenberger, 2003). The BST package is a four-part training strategy that involves a) clear explicit instructions for the target behaviors, b) live or video modeling or demonstration, c) rehearsal or practice of the target behaviors, and d) feedback on the performance that occurred during rehearsal (Miltenberger, 2003). Three studies have examined the generalization of the effects of BST in teaching DTT (Lafasakis & Sturmey, 2007; Lerman, Tetreault, Hovanetz, Strobel, & Garro, 2008; Ward-Horner & Sturmey, 2008). Lefasakis and Sturmey (2007) used BST to train parents to implement DTT for motor imitation skills with their children. After training, the parents demonstrated improved implementation of DTT for motor imitation and those effects generalized to teaching other motor imitation targets and to a vocal imitation program. In addition, improvements in implementation were associated with increases in child correct responding on the target skills. Ward-Horner and Sturmey (2008) also demonstrated that the effects of training DTT for one program (i.e., gross motor imitation) generalized to another program (i.e., vocal imitation). Lerman et al. (2008) demonstrated that individuals trained to use three prompting procedures (i.e., least-to-most, most-to-least, and time delay) within DTT used the procedures effectively across additional targets and learners.

Although a large body of evidence supports the use of BST to teach staff or other caregivers to implement DTT with high accuracy (Parsons, Rollyson & Reid, 2012), the training package involves a considerable time investment for a trainer with at least some portion of the training requiring individual rehearsal and feedback opportunities (Crockett, Fleming, Doepke, & Stevens, 2007; Downs, Downs & Rau, 2008). Given the high staff turnover rate in early intervention sites, the BST method may not always be cost effective for training. In addition, BST requires the trainer be present to observe each staff member's rehearsals and provide feedback, which may not be practical for large or multisite agencies or when training caregivers in remote locations.

Interactive computer training, or computer-based instruction (CBI), involves the presentation of training material via a computer or Internet site and requires the learner to answer questions about the material or engage in some activity related to the material (Williams & Zahed, 1996). It is proposed to be an effective alternative when face-to-face instruction is not possible (LeBlanc et al., 2009). CBI has been demonstrated to be more effective than lecture (Williams & Zahed, 1996) and reading (Eckerman et al., 2002) with the benefits of being private and self-paced. Additionally, ongoing implementation of CBI is more efficient and cost-effective once the training materials are created, though there is significant initial time and expense to develop the training materials (Blanchard & Thacker, 2004). Thus, the CBI product must be used often enough to justify the initial development cost. Another potential disadvantage to CBI is that the responses required within the training platform (e.g., answering multiple choice questions) may not be similar to the responses required during implementation of the procedure that is being trained (e.g., presenting a stimulus array). Depending on the extent of the difference, high performance in the training platform may not generalize to implementation of a procedure with high procedural integrity. To maximize potential generalization, responding in the CBI platform should be as similar as possible to responding when implementing the procedure.

CBI has become a widely used method of staff training for various skills in a number of different disciplines (e.g., Desrochers, Clemmons, Grady, & Justice, 2001; Eckerman et al., 2002; Lambert, 1989) including behavior analysis (e.g., Nosik & Williams, 2011; Pollard, Higbee, Akers & Brodhead, 2014). Nosik and Williams (2011) used a CBI program to train four staff to implement DTT using least-to-most prompting to teach matching and receptive instructions. The CBI consisted of instructions, video models of exemplars, multiple choice questions, and scoring checklists for videos exemplars and non-exemplars of the DTT procedure. CBI improved participants' implementation of DTT to approximately 70% to 90% of steps correct. After scoring videos, participants' implementation of DTT further improved to approximately 90% to 100% of steps correct. More recently, Nosik, Williams, Garrido, and Lee (2013) compared the effects of BST and CBI training on six staff's implementation of DTT with an adult with autism. Participants trained with BST outperformed participants trained with CBI. In addition, participants trained with BST reported feeling very competent to perform the task, whereas CBI-trained participants reported feeling somewhat competent. These studies and others (Higbee et al., 2016; Pollard et al., 2014; Serna et al., 2016) illustrate the utility of interactive CBI programs (e.g., video clips, instructions, quizzes) to train people to implement various components of DTT.

The purpose of this study was to compare CBI to BST to train novice undergraduate students to conduct DTT for auditory–visual conditional discriminations. We chose to target auditory–visual conditional discrimination because it is one of the foundational skills taught in DTT and is important for language comprehension, and other pre-academic and academic skills taught in EIBI curricula (Green, 2001; Smith, 2001). This direct comparison extends the comparison of Nosik et al. (2013) to a new skill and with a group research design. In addition, this study evaluated the cost effectiveness of the two procedures and the potential return on investment of the development costs for the CBI module to inform applied practice implementation at scale.

## METHOD

### *Participants and Setting*

Fifty undergraduate students at a large public university in the Southeastern U.S. participated in this study. A power analysis was conducted which indicated that 50 participants would be needed to ensure a reasonable chance to detect a difference if a difference existed. Participants earned extra credit for their participation. Participants were excluded from the study if they had previous experience working in an EIBI setting; only one participant met this exclusionary criterion. Participants were assigned to experimental conditions using a random number generator that was constrained to randomize up to 50. Twenty-five of the participants experienced individually administered online CBI, whereas the other 25 participants experienced individually administered instruction via BST. There were no statistically significant differences between the groups on any demographics (e.g., age, number of previous experiences in developmental disabilities). See Table 1 for the statistics for these demographic variables.

Additionally, 10 staff from an ABA service agency participated in a follow-up evaluation (i.e., modified CBI curriculum) to assess the effects of the CBI training on implementation with a child with autism. At the time of the study, participants had been employed with the agency, on average, just over a year ($M$ = 12.7 months; range, 3-25 months) and had between 11 and 84 months of experience working with individuals with developmental disabilities ($M$ = 47.3 months). Table 1 provides additional demographic information on these participants.

### *Setting and Materials*

Training activities took place in a room on a college campus containing a computer with Internet access on a desk (for CBI), a screen projector (for BST), a table, and chairs. Probes took place in an adjacent room with a table, chairs, a data sheet (available from the first author upon request), stimuli, reinforcers, a video camera, and a research assistant. During the pretraining, posttraining, and postfeedback probes, the participants used a data sheet, a pen, three 7.6 cm x 7.6 cm picture card stimuli, and edible and tangible reinforcers.

### *Curriculum*

Participants were taught to use errorless teaching procedures to establish auditory–visual discriminations (i.e., listener discriminations). The curriculum consisted of 11 lessons presented either in CBI or live BST. The introductory lesson described the use of DTT procedures with children with autism, the basics of a discrete trial (i.e., antecedent, behavior, consequence, intertrial interval), and the importance of listener skills. The next lessons covered setting up the teaching environment, arranging the comparison array (i.e., three items presented equidistant apart), getting the learner's attention, presenting the conditional stimulus, and responding to the learner's correct responses. The next lesson in the module covered the use of most-to-least prompting for teaching trials and least-to-most prompts for

Table 1
Participant Background Information

| | CBI | BST | | | Extension |
|---|---|---|---|---|---|
| | Mean | Mean | *t*(48) | p | Mean |
| Age | 20.6 | 21.9 | 1.02 | 0.311 | 29.4 |
| # of Psychology Courses Completed | 3.2 | 2.3 | 1.54 | 0.130 | - |
| # of Prior Psychology Practica | 0.2 | 0.1 | 1.05 | 0.299 | - |
| # of Other Experiences in Developmental Disabilities | 0.1 | 0.1 | 0.00 | 1.000 | - |
| Months Employed with Agency | - | - | - | - | 12.7 |
| Months of Experience with Individuals with Disabilities | - | - | - | - | 47.3 |

assessment probes, consistent with the recommendations of Green (2001) and Grow and LeBlanc (2013). The final lesson covered teaching an entire block of trials (i.e., putting all prior information together across multiple trials and learner responses) and data collection. Nine of 11 lessons provided opportunities to practice and receive performance feedback (e.g., verbal, textual).

### *Procedures*

*Pretraining probes.* After completing a demographics form, each participant completed a pretraining probe in which he or she conducted 24 trials of DTT to teach auditory–visual conditional discriminations to an adult actor pretending to be a child with autism. The trials consisted of two fully counterbalanced stimulus array arrangements with rotated target presentation (Grow & LeBlanc, 2013). The number of trials was selected to provide equal opportunities to respond to correct independent responses and incorrect responses at all potential prompt levels including repeated incorrect responses and no response until full physical prompt. The experimenter provided the participant with the necessary materials to implement a listener responding program using DTT during a role-play with an actor. The actor followed a scripted sequence of correct and incorrect responses (sequence available upon request). The experimenter provided a neutral response (e.g., "Just do your best") if the participant inquired about how to implement the procedure. Feedback was not provided during or after the role-play. The experimenter collected data on the participant's accuracy of implementation in the probe. If the participant implemented DTT with 70% accuracy or higher in the pretraining probe, he or she would have been excused from the study; however, none scored above 70%. Participants were then randomly assigned to either experience CBI training or BST.

*CBI.* The CBI program (Geiger, Severtson, & LeBlanc, n.d.) was accessed from a computer with speakers located in the research lab space. The experimenter started the CBI program for the participant but did not interact with the participant (e.g., answer questions, provide feedback) while the program was running. The program contained narration, on-screen text, pictures, animations, and video exemplars and nonexemplars of the procedures. Some lessons contained active response opportunities pertinent to implementation of the procedures (e.g., click-and-drag to arrange stimuli, click to mark a virtual data sheet). See Supporting Information for sample screenshots from the CBI program. Each part of the conditional discrimination procedure was taught in a lesson in the CBI program. At the end of each lesson, the participants completed a quiz (e.g., multiple choice, true/false, scoring video samples) to assess their understanding of information presented in the CBI program. If the participant answered incorrectly, he or she was required to view the lesson again before retaking the quiz and moving on to the next module. This continued until the participant answered all quiz questions correctly. At the end of the program, the participant completed a cumulative quiz with multiple choice questions about the entire procedure. Once the participant completed all lessons and correctly answered the cumulative quiz questions, with at least 90% accuracy, the CBI training was complete. In the cumulative quiz, participants received feedback on incorrect answers and had the opportunity to repeat the quiz but did not have to return to instruction in the module. The navigation of the module was fixed and progressive (i.e., the participant could not choose to jump ahead).

*BST.* The experimenter delivered a live, individual, interactive BST session with a PowerPoint® presentation consisting of on-screen text, pictures, and video models of exemplars and nonexemplars of the procedure (i.e., instructions and modeling). The PowerPoint® was projected on a white wall in the research space. The content, text, pictures, video models, and teaching topics were identical to those presented in the CBI program. A detailed list of instructional components and general content for both programs are available from the first author upon request. The experimenter answered participant questions throughout the training. The experimenter provided rehearsal opportunities and performance feedback throughout the slideshow training. Instruction did not progress to a new module until the participant had completed the component correctly by performing the target skill three times consecutively with 100% accuracy. Once the entire presentation was completed, the participant rehearsed 12 trials of DTT with the experimenter in a role play. The experimenter followed a scripted response during role plays, which is available from the authors upon request. If the participant made an error, the experimenter immediately stopped the role play to provide verbal feedback and then allowed the participant to resume. Practice continued until the participant completed 12 teaching trials with at least 90% accuracy.

*Posttraining probes.* The procedures for the posttraining probe were identical to the pretraining probes described above but the script varied slightly so that different confederate responses occurred on different trials (e. g., partial physical required on target 2, trial 7 in one script; partial physical required on target 1, trial 5 in another script), but the proportions of types of responses remained equal.

*Feedback.* Performance feedback was provided immediately following the posttraining probe if a participant implemented the teaching session with less than 85% accuracy. The instructor provided details on the specific procedural steps implemented incorrectly, explained how each step should be completed, and modeled the correct implementation of the step. If the participant asked questions, the instructor answered them and provided additional models, if necessary. Once the participant received feedback, he or she was immediately allowed to complete a postfeedback probe, which was conducted similar to the pre- and posttraining probes but was a slight variant for the specific trials. The time from posttraining probe to postfeedback probe was usually 1-5 min. Postfeedback probes were not conducted if participants implemented the procedure with 85% accuracy or higher during the posttraining probe.

*Extension demonstration with children with autism.* After the initial comparison study was completed, a follow-up demonstration of the use of the CBI training was conducted in a provider agency for children with autism. Ten participants employed at the agency experienced a modified CBI curriculum. The data collection component of the CBI module was modified after the completion of the initial comparative study based on the error analysis from the prior 25 participants. The data collection procedure was simplified and the visuals for the instruction in this component were enhanced. These participants completed pretraining, posttraining, and postfeedback probes similar to those described above. These participants completed a follow-up probe with a child with autism during the child's regularly scheduled therapy hours. Follow-up probes occurred after participants met criterion in the posttraining or postfeedback probes.

*Acceptability questionnaire.* Participants completed an acceptability questionnaire after their participation was complete. The questionnaire consisted of 16 questions with Likert-type scaled responses (i.e., 1 = strongly disagree, 5 = strongly agree). The questionnaire was adapted from an established acceptability measure in the staff training literature (Rothwell & Sredl, 1997) and an acceptability measure currently used in an online computer-based training program (Fox, n.d.). The questionnaire items referenced the program content, presentation of the material, and the participant's overall impression of the training program. Training acceptability questionnaires are available from the authors upon request.

### *Experimental Design and Measurement*

A randomized two-group repeated measures design was used to compare the effectiveness of CBI and BST on participant implementation of a conditional discrimination procedure. Participants were randomly assigned, via a random number generator, to participate in either the CBI or BST training. The main dependent measure was the percentage of steps correctly completed during the pre- and posttraining probes, and the postfeedback probe, if applicable. Data were collected in vivo and from videos of these probes. Each step was scored independently (i.e., if a step was scored as incorrect for a trial, later steps in that trial could still be scored as correct). The percentage of steps completed correctly was calculated by dividing the number of steps completed correctly by the total number of steps and converting to a percentage.

Several secondary dependent measures were collected including a) material development and training time, b) the average amount of time it took participants to complete training, c) the acceptability of the training procedure, and d) the types of errors made in posttraining probes. Material development and training included the sum of the time spent a) writing the content for the program; b) creating the PowerPoint® slideshow for BST or creating the web-based program for CBI; c) participants' total training time in which they were interacting with the instructor in BST, excluding pre- and posttraining probes; and d) the total time during which the instructor provided feedback if the participant failed to meet the mastery criterion in the posttraining probe. The amount of time required to complete training was measured from the beginning of the training session to the end of the training session and did not include time used to conduct pretraining, posttraining, or postfeedback probes. For CBI, the web-based program recorded the duration of training from the time participants started the program until they completed the final question of the cumulative quiz. To measure training time of BST, the instructor started a stopwatch at the beginning of training and stopped the stopwatch after presenting the last lesson. Acceptability of the training procedure was measured via a training acceptability questionnaire.

The types of errors made in the posttraining probe were evaluated to determine if any one error was more common than others. Procedural steps assessed in the error analysis are presented in Table 2. The data on types of errors made were analyzed as the percentage of participants who made an error on any given step for each group. The percentage of participants who made an error was calculated by dividing the number of participants who made at least one error on a specific step by the total number of participants in the group and converting to a percentage. Finally, the average attempts until the mastery criterion was met in training were reported for each group but were not compared

Table 2
Percentage of Participants who Erred at Least Once on Skills Taught in CBI and BST during Posttraining Probes

| Task Analysis Steps | CBI *M* | BST *M* |
|---|---|---|
| Arranging the stimuli | 48% | 28% |
| Getting learner's attention* | 52% | 16% |
| Presenting the instruction | 40% | 24% |
| Waiting 3 s for response | 16% | 8% |
| Making no response and removing stimuli following learner errors** | 52% | 20% |
| Immediately providing prompt during teaching trials | 52% | 28% |
| Providing correct prompt level* | 80% | 28% |
| Providing prompt at next level of intrusiveness during error correction* | 76% | 32% |
| Providing praise and tangible reinforcer following learner's correct response | 76% | 60% |
| Recording data** | 100% | 80% |

*Note.* Statistically significant differences between groups are denoted by * (<.01) and ** (<.05).

statistically because the types of within-training performance (i.e., answering questions in CBI, performing the teaching procedure in BST) were not identical.

*Interobserver agreement.* A second trained observer who was not blind to condition collected interobserver agreement (IOA) data on participant's accuracy (the primary measure) for 31% of CBI participants and 60% of BST participants across all probes. The participants were selected for IOA using a random numbers generator. For participant accuracy, agreement was scored if both observers identified a procedural step as being completed correctly or incorrectly. Point-by-point agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and converting the ratio to a percentage. Mean IOA for CBI pretraining, posttraining, and postfeedback probes was 95% (range, 86%- 100%), 97% (range, 93%-100%), and 98% (range, 94%-100%), respectively. Mean IOA for BST pretraining, posttraining, and postfeedback probes was 96% (range, 75%-100%), 99% (range, 97%-100%), and 99%, respectively. IOA data were not collected for the secondary measures, including investigator's time investment, total training time, and acceptability of the training procedure.

*Experimenters' procedural integrity.* Procedural integrity was scored for the implementation of BST and probe sessions from videos by an independent, trained but not blinded observer. Procedural integrity was calculated as the percentage of steps correctly completed by the instructor during BST and when providing feedback after posttraining probes, and the percentage of steps correctly completed by the actor (i.e., researcher serving as a confederate client) in the pretraining, posttraining, and postfeedback probes. For BST, procedural integrity was measured with a checklist, in which the observer recorded whether the instructor provided correct instructions, rehearsal opportunities, and feedback for each module of the BST program. Experimenter procedural integrity for BST was measured for 80% of BST sessions. For all sessions measured, experimenter procedural integrity was 100%. For the posttraining and postfeedback probes, procedural integrity measures included the actor's presentation of learner responses based on a scripted learner response sequence and the use of corrective feedback if implementation accuracy was below 85%. Procedural integrity percentage was calculated by dividing the number of steps performed correctly by the total number of steps and multiplying by 100. Experimenter procedural integrity in the probes was measured for 40% of CBI participants and 52% of BST participants with participants selected via random number generator. Mean procedural integrity for the CBI probes was 95% (range, 86%-100%) and for the BST probes was 99% (range, 96%-100%).

*Data analysis.* A repeated-measures 2 x 2 ANOVA was used to compare the effects of CBI and BST on participants' accuracy during the pretraining and posttraining probes. A two-way, or 2 x 2, ANOVA determines how a response is affected by two factors with an estimate of the likelihood that differences in responding might be accounted for by chance. When each participant is assessed repeatedly prior to and after intervention, it is referred to as a repeated measures 2 x 2 ANOVA. Data from the postfeedback probe were not included in the repeated-measures ANOVA calculation as the sample size was much smaller. An independent samples *t*-test was used to compare learner time invested and acceptability of the training procedure. A chi-square test was used to compare the percentage of participants who made an error on specific steps during the posttraining probe.

## RESULTS

### *Effectiveness of the Training Procedures*

Figure 1 (top panel) displays participants' percentage of procedural steps completed accurately during the pretraining, posttraining, and postfeedback probes. All of the participants performed poorly during the pretraining probe, averaging 12% (range, 0%-41%) and 8% (range, 0%-44%) for CBI and BST, respectively. Participants' accuracy improved substantially after training, averaging 87% (range, 65%-97%) and 96% (range, 74%-100%) for CBI and BST, respectively. Both the effect sizes for CBI x probe condition ($d$ = 7.48) and BST x probe condition ($d$ = 11.56) were well above Cohen's (1988) convention for a large effect ($d$ = .80). The percentage of participants who met the mastery criterion of least 85% of steps completed accurately was 64% (16 out of 25) and 92% (23 out of 25) for CBI and BST, respectively. Finally, everyone who received feedback improved their scores in the postfeedback probe and met the mastery criterion after one feedback session. The mean percentage of total steps completed accurately in postfeedback for CBI was 95% (range, 88%-99%) and for BST was 99% (range, 98%-99%). A 2 x 2 repeated measures ANOVA showed a significant main effect of Probe Condition (i.e., pretraining vs. posttraining), $F(1, 48) = 2023.00$, $p < .01$

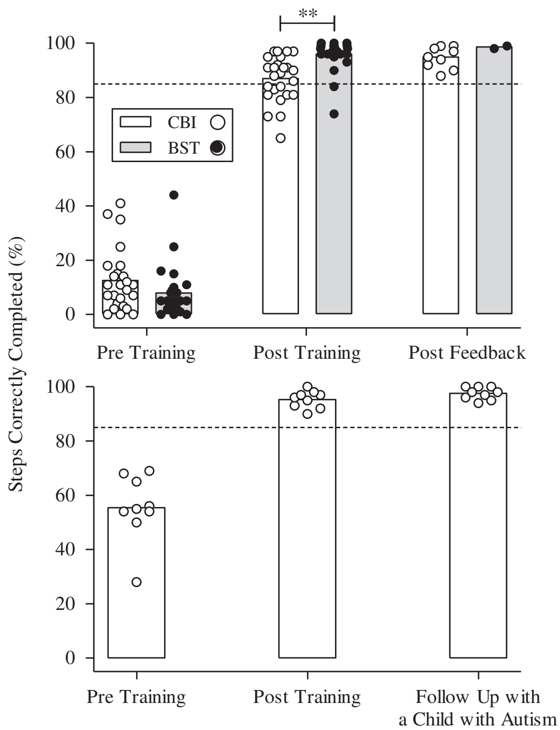


Figure 1. Participants' accuracy of implementation when teaching auditory–visual conditional discriminations to adult actors during the main study (top panel), and to adult actors and a child with autism in the extension with employees at a provider agency for children with autism (bottom panel). In the top panel, the asterisks denote a significant difference at the .01 level. The white bar represents the mean percentage score for participants who experienced CBI. The gray bar represents mean percentage correct for participants who experienced BST. The circles (white = CBI; black = BST) represent each participant's percentage of steps completed correctly.

and a significant interaction effect of Probe Condition x Training Type, $F(1, 48) = 13.96$, $p < .01$. These results indicate both CBI and BST were effective at improving participants' accuracy of implementation, though BST was significantly more effective than CBI.

Figure 1 (bottom panel) depicts participants' procedural integrity during the pretraining, posttraining, and postfeedback probes for the 10 CBI participants employed within the autism service agency. The percentage of procedural steps completed correctly during pretraining probes was low across participants, averaging 55% (range, 28%-69%). This average substantially increased to 95% (range, 90%-100%) during the posttraining probes. Given the integrity with which participants implemented the target skill, feedback and the postfeedback probe(s) were not provided. During the follow-up probe with the child with autism, participants' procedural integrity remained high and the average increased slightly to 97% (range, 94%-100%).

Table 2 displays the percentage of participants who made at least one error on a given procedural step during the posttraining probe. The steps with statistically significant differences between groups were 1) getting the learner's attention ($\chi^2$[1, $N$ = 50] = 7.22, $p$ < .01); 2) providing the correct prompt level ($\chi^2$[1, $N$ = 50] = 13.61, $p$ < .01); 3) not responding to learner errors and removing the stimuli ($\chi^2$[1, $N$ = 50] = 5.56, $p$ < .05); 4) providing the prompt at the next level of intrusiveness during error correction, ($\chi^2$[1, $N$ = 50] = 9.74, $p$ < .01); and 5) recording data ($\chi^2$[1, $N$ = 50] = 5.56, $p$ < .05). Recording data was the most common skill in which participants erred during both CBI and BST.

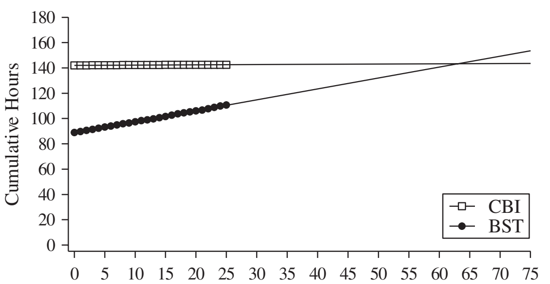


Figure 2. Time investment of the trainer for CBI and BST. The white squares represent total time invested to create the module and train the CBI group, and the black circles represent total time invested to create the training program and train the BST group. The solid lines are the trend lines for the data sets.

*Efficiency and Utility Analysis*

The two training programs were compared with respect to their efficiency as well as their effectiveness. The efficiency and cost-effectiveness of the programs were estimated by comparing the total average duration of instruction during CBI to that of BST and incorporating that information into a utility analysis with the initial time invested in preparation of the instructional materials. The mean duration of learner time invested for CBI was .99 hr (59.32 min) (range, .78-1.4 hr), whereas the mean duration of learner time invested for BST was .86 hr (51.76 min) (range, .67-1.13 hr). A $t$-test confirmed the difference in duration of learner time was significant $t(48) = 3.40$, $p < .01$.

Figure 3 displays the time investment of the trainer for developing materials, delivering instruction, and providing feedback to participants who failed to meet the mastery criterion for both the CBI and BST programs. The initial time investment for developing materials was 142 hr for CBI and 89 hr for BST. As additional participants were trained, the trainer did not invest any additional time during training for the CBI group but did invest an additional .5 hr to provide feedback to 10 participants, increasing the total time investment to 142.5 hr. For BST, the trainer invested a mean of .86 hr (51.8 min) to train each participant and an additional .14 hr (8.5 min) to provide feedback to two participants ($M$ = .07 hr), increasing the total time invested to 110.7 hr for training 25 participants. The total amount of time invested in BST was less than the time invested for CBI. However, if additional individuals were to be trained using the two methods, the projected breakeven point (i.e., equal investment) would be 62 participants in a given condition, assuming the same mean training duration for future trainees. This means that the trainer would have had to train 62 people in CBI to make the time investment per participant equal to that invested in training 62 people in BST. If the trainer had trained more than 62 people in each group, then CBI would have been more efficient than BST.

*Acceptability of the Training Procedures*

For all of the items, both CBI and BST were rated positively on the training acceptability questionnaire. Only two items resulted in statistically significant differences in participants' ratings. For the item indicating that the learning objectives were successfully achieved, the mean rating for CBI was 4.64 (range, 4-5) and the mean rating for BST was 4.88 (range, 4-5; $t(48) = 2.03$, $p < .05$). For the item indicating if the participant would have preferred to learn the content in the other training format the mean rating for CBI was 2.76 (range, 1-5), and the mean rating for BST was 1.68 (range, 1-5), which was significant at $t(48) = 3.74$, $p < .01$. These results indicate, on average, the CBI group might prefer to learn the content with a live instructor (i.e., the other condition), and the BST group would not want to learn with a computer-based program.

## DISCUSSION

The data suggest both CBI and BST were effective at improving participants' procedural implementation of DTT, using the errorless learning technique. These findings are consistent with the previous literature on using BST (e.g., Crockett et al., 2007; Downs et al., 2008; Lafasakis & Sturmey, 2007) and using CBI (e.g., Nosik & Williams, 2011; Pollard et al., 2014) to train staff to implement DTT. Similar to Nosik et al. (2013), BST was more effective than CBI, but both procedures resulted in a majority of participants immediately performing at a criterion of success. A very brief feedback session resulted in criterion level performance for the remaining participants.

Previous research on component analyses of BST has shown rehearsal and feedback are necessary for optimal effectiveness of the training package (e.g., Roscoe & Fisher, 2008; Sterling-Turner, Watson, & Moore, 2002). The BST program allowed participants to rehearse each step of the procedure live with an adult, conduct the entire procedure, and receive performance feedback. The CBI program provided some active response activities that simulated rehearsal with feedback (e.g., collecting data on a virtual data sheet), but did not provide active response activities live with an adult for every step in the procedure (e.g., providing prompts, reinforcing correct responses) or for conducing the entire procedure. Therefore, it is somewhat unsurprising that participants in the BST group performed better than participants in the CBI group. However, CBI participants' performance improved to criterion level after receiving feedback from the experimenter during the postfeedback probe. This finding suggests the difference in scores between the two groups could be due to the lack of rehearsal with feedback during portions of the CBI program. Future research could examine the use of CBI that includes more active response activities with feedback, to train participants to implement DTT (e.g., Higbee et al., 2016). In addition, the number of attempts during active response activities could be recorded. Because the active response activities were a simulation of rehearsal of the procedural step, data on performance in these activities may have been more informative than the data on performance in the multiple-choice quizzes. Future research could examine the possibility of collecting data on attempts to criterion in active response activities in CBI and if that correlates with performing the actual procedure with learners.

It is important to note that this study compared CBI to an optimized version of BST. The BST program used recommended practices, such as (1) conducting one-on-one teaching sessions provided by an expert instructor, (2) providing multiple models of both exemplars and nonexemplars, allowing frequent rehearsals of component skills (e.g., present the instruction, provide reinforcement) as well as the terminal skill (i.e., conduct a DTT teaching session), and (3) requiring the participant to rehearse until meeting a mastery criterion (Parsons et al., 2012). The version of BST used in this study is likely more optimal than typical live training using BST outside of university research settings. Although the quality of CBI would not change in a practice setting, common-practice BST or basic didactic training without rehearsal and feedback would likely be of lower quality than that studied here. CBI ensures every trainee has the same training experience (i.e., same number and type of opportunities to respond; Higbee et al., 2016) and eliminates the variability that can occur if multiple trainers deliver training.

This investigation examined procedures to teach college students and then staff to implement conditional discrimination training under reasonable conditions. A limitation of this study is that the confederate did not respond with problem behavior during the pretest and posttest probe conditions, which may occur in everyday practice. The primary purpose of the study was to directly compare the two training procedures' effects on the same repertoire rather than to produce the entire repertoire required by someone who might work with a child with autism. As the participants in the initial study were novices who had never implemented any form of DTT, it seemed reasonable to the researchers to test the newly taught skills without including significant problem behavior in the test probes. In the extension with staff and children with autism, problem behavior may or may not have occurred as this could not be programmed in advance. In BST, the trainee often learns to implement a skill to fluency under optimal conditions (i.e., the posttest probe) and then more stressful or challenging situations are introduced and practiced once mastery has occurred in the initial situation. Future studies employing single subject designs with staff in human service agencies might examine the effects when problem behavior is included in the posttest condition with a confederate or include training with children with autism who exhibit different levels of severity of problem behavior.

Many other important skills could be taught via CBI such as implementing naturalistic programming, implementing preference assessment procedures, and collecting data. Each of these skills could be targeted for training with CBI and it is worthwhile to directly examine several of them before the findings of this study are generalized too far. Other procedures might be easier or more difficult to teach and this could impact the social acceptability of the training procedure as well as the effectiveness of either BST or CBI. There might also be a significant difference in the time required for instructional design for some of these other training topics. For example, the module studied here is now part of a complete CBI curriculum designed to meet the requirements for the Registered Behavior Technician® credential. The module designed to teach implementation of multiple formats of preference assessments is twice as long in duration as the module evaluated here which might worsen the perceived experience of the learner. It is also worth examining the extent to which the entire initial training for implementation of EIBI can be conducted via CBI with reasonable outcomes. There are multiple commercially available curricula that provide a series of modules targeting different skill sets. It will be important to determine the effects of these CBI curricula with respect to the skills produced and the required amount of live supplemental instruction required to prepare novices for working with individuals with autism and other disabilities.

The procedural steps with substantially more errors in CBI than BST involved providing prompts (e.g., immediately provide prompt, correct prompt level, provide prompt at next level during error correction) and recording data. The *Providing Prompts* module in the CBI program had one active response activity associated with arranging the prompts according to their level of intrusiveness. The *Providing Prompts* quiz had a higher mean attempts to criterion compared to some other quizzes in the program, suggesting that more explicit models and active response activities could be added to further clarify this skill.

Although there was a significant difference in the level of errors in data collection between the two groups, both groups had a considerably high proportion of participants that made errors on this step. In BST, participants were provided an opportunity to rehearse data collection and receive feedback during the *Probe Trials and Teaching Trials* modules. In CBI, participants were provided simulated opportunities of rehearsal with feedback, by clicking on a virtual data sheet based on an ongoing video during the *Probe Trials and Teaching Trials* modules of the program. However, participants may have been required to complete too many steps to reliably perform data collection (i.e., identifying appropriate prompts from the probe to use in later teaching trials, identifying if the learner's response was correct or incorrect, implementing error correction based on the learner's selection, and immediately taking data). Therefore, procedural integrity may be improved if the data collection system required participants to record only the prompt level required to get a correct response in probes and whether each teaching trial was correct or incorrect. The CBI module was changed, and participants in the extension experienced training with this modification. All participants were able to meet criterion during posttraining probes and during implementation with a child with autism.

The procedural step of *presenting the instruction* accounted for a very low number of errors for both the CBI and BST groups. The rehearsal data that show a low mean attempts to criterion ($M$ = 3.20) in BST for presenting the instruction correspond to the high performance on this step in the posttraining probe. However, in CBI the *presenting the instruction* quiz had one of the highest mean attempts to criterion of all the quizzes. This discrepancy between quiz score and implementation of the procedural step in the posttraining probe suggests the quiz may be assessing skills that are not necessary for correct implementation and/or one or more quiz questions is unnecessarily difficult. One quiz question involves identifying that the instructor is accidently motioning towards the correct item by looking more at that item and leaning her head towards it. This may be a difficult discrimination for participants to make from a video example, and identifying this mistake may not be necessary for participants to correctly provide the instruction during DTT. Future research on using this CBI program might examine removing or changing this question to see if performance on the quiz improves and results in similarly low levels of errors in providing the instruction when participants implement DTT.

CBI required slightly more learner time investment ($M$ = 59 min) than BST ($M$ = 52 min) and substantially more initial time investment for creating the CBI program than the ongoing instructor time investment for BST for training 25 participants. This finding is consistent with the literature, which suggests the main limitation of CBI is the initial time and financial investment to create the materials (Blanchard & Thacker, 2004). However, the utility analysis showed that if the programs were used to train more than 62 people each, then CBI would be more efficient because of the recurring need to invest the trainer time with live BST. These findings suggest creating in-house CBI training programs may not be cost-effective unless the agency is training a large number of staff and/or are training people who are in geographically remote locations when live BST is not possible (LeBlanc et al., 2009). Future studies comparing staff training procedures could use a similar utility analysis to fully examine the return on the initial investment required to create the training materials.

In sum, both BST and CBI have previously been demonstrated to be effective for training staff in many different skills, including the implementation of DTT (e.g., Pollard et al., 2014; Sarokoff & Sturmey, 2008; Serna et al., 2016; Ward-Horner & Sturmey, 2008). This study extended Nosik et al. (2013) by offering an additional comparison of CBI to BST with an evaluation of the return on investment. Although BST was slightly more effective and efficient than CBI, minimal performance feedback closed the gap and would generate a return on the initial development investment after fewer than 70 staff were trained. Additionally, this study provided further evidence that CBI was effective at training participants in a complex DTT procedure—auditory–visual conditional discrimination training—using the errorless teaching prompting method. This study provides further support that CBI may offer an acceptable staff training alternative if optimized BST is not possible or not feasible.

## REFERENCES

Blanchard, P. M., & Thacker, J. W. (2004). *Effective training: Systems, strategies, and practices* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Crockett, J. L., Fleming, R. K., Doepke, K. J., & Stevens, J. S. (2007). Parent training: Acquisition and generalization of discrete trial teaching skills with parents of children with autism. *Research in Developmental Disabilities, 28,* 23-36. https://doi.org/10.1016/j.ridd.2005.10.003

Desrochers, M. N., Clemmons, T., Grady, M., & Justice, B. (2001). An evaluation of simulations in developmental disabilities (SIDD): Instructional software that provides practice in behavioral assessment and treatment decisions. *Journal of Technology in Human Services, 17*, 15-27. https://doi.org/10.1300/J017v17n04_02

DiGennaro Reed, F. D., Reed, D. D., Baez, C. N., & Maguire, H. (2011). A parametric analysis of errors of commission during discrete-trial training. *Journal of Applied Behavior Analysis, 44*, 611-615. https://doi.org/10.1901/jaba.2011.44-611

Downs, A., Downs, R. C., & Rau, K. (2008). Effects of training and feedback on Discrete Trial Teaching skills and student performance. *Research in Developmental Disabilities, 29,* 235-246. https://doi.org/10.1016/j.ridd.2007.05.001

Eckerman, D. A., Lundeen, C. A., Steele, A., Fercho, H., Ammerman, T., & Anger, W. K. (2002). Interactive training versus reading to teach respiratory protection. *Journal of Occupational Health Psychology, 7*, 313-323. https://doi.org/10.1037/1076-8998.7.4.313

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology, 38*, 439-450. https://doi.org/10.1080/15374410902851739

Fox, E. (n.d.). *Survey for an introduction to relational frame theory.* Retrieved from http://foxylearning.com/tutorials/rft

Geiger, K., Severtson, J. M., & LeBlanc, L. A. (n.d.). *Teaching listener responding to children with autism.* Retrieved from https://foxylearning.com/tutorials/tlr.

Green, G. (1996). Early behavioral intervention for autism: What does research tell us? In C. Maurice, G. Green, & S. C. Luce (Eds.), *Behavioral intervention for young children with autism: A manual for parents and professionals* (pp. 29-44). Southborough, MA: New England Center for Autism.

Green, G. (2001). Behavior analytic instruction for learners with autism: Advances in stimulus control technology. *Focus on Autism and Other Developmental Disabilities, 16*, 72-85. https://doi.org/10.1177/108835760101600203

Grow, L. L., Carr, J. E., Gunby, K. V., Charania, S. M., Gonsalves, L., Ktaech, I. A., & Kisamore, A. N. (2009). Deviations from prescribed prompting procedures: Implications for treatment integrity. *Journal of Behavioral Education, 18*, 142-156. https://doi.org/10.1007/s10864-009-9085-6

Grow, L., & LeBlanc, L. (2013). Teaching receptive language skills: Recommendations for instructors. *Behavior Analysis in Practice, 6(1),* 56-75. https://doi.org/10.1007/BF03391791

Harrison, R. J., & Green, G. (1990). Development of conditional and equivalence relations without differential consequences. *Journal of the Experimental Analysis of Behavior, 54*, 225-237. https://doi.org/10.1901/jeab.1990.54-225

Higbee, T. S., Aporta, A. P., Resende, A., Nogueira, M., Goyos, C., & Pollard, J. S. (2016). Interactive computer training to teach discrete-trial instruction to undergraduates and special educators in Brazil: A replication and extension. *Journal of Applied Behavior Analysis,* 4, 780-793. https://doi.org/10.1002/jaba.329

Johnson, C., & Sidman, M. (1993). Conditional discrimination and equivalence relations: Control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 59*, 333-347. https://doi.org/10.1901/jeab.1993.59-333

Johnson, E., & Hastings, R. P. (2002). Facilitating factors and barriers to the implementation of intensive home-based behavioural intervention for young children with autism. *Child: Care, Health and Development, 28*, 123-129. https://doi.org/10.1046/j.1365-2214.2002.00251.x

Lafasakis, M., & Sturmey, P. (2007). Training parent implementation of discrete-trial teaching: Effects on generalization of parent teaching and child correct responding. *Journal of Applied Behavior Analysis, 40*, 685-689. https://doi.org/10.1901/jaba.2007.685-689

Lambert, M. E. (1989). Using computer simulations in behavior therapy training. *Computers in Human Services, 5*, 1-12. https://doi.org/10.1300/J407v05n03_01

Larson, S. A., & Hewitt, A. S. (Eds.) (2005). *Staff recruitment, retention, and training strategies for community human services organizations*. Baltimore, MD: Brookes Publishing.

Leaf, R., & McEachin, J. (Eds.) (1999). *A work in progress: Behavior management strategies and curriculum for intensive behavioral treatment of autism*. New York, NY: DRL Books.

LeBlanc, L. A., Gravina, N., & Carr, J. E. (2009). Training issues unique to autism spectrum disorders. In J. Matson (Ed.), *Applied behavior analysis for children with autism spectrum disorders* (pp. 225-235). New York, NY: Springer.

Lerman, D. C., Tetreault, A., Hovanetz, A., Strobel, M., & Garro, J. (2008). Further evaluation of a brief, intensive teacher-training model. *Journal of Applied Behavior Analysis, 41*, 243-248. https://doi.org/10.1901/jaba.2008.41-243

Libby, M. E., Weiss, J. S., Bancroft, S., & Ahearn, W. H. (2008). A comparison of most-to-least and least-to-most prompting on the acquisition of solitary play skills. *Behavior Analysis in Practice, 1*, 37-43. https://doi.org/10.1007/BF03391719

MacDuff, G. S., Krantz, P. J., & McClannahan, L. E. (2001). Prompts and prompt-fading strategies for people with autism. In C. Maurice, G. Green, & R. Foxx (Eds.), *Making a difference: Behavioral intervention for autism* (pp. 37-50). Austin, TX: Pro-Ed.

McIlvane, W. J., Dube, W. V., Kledaras, J. B., & Iennaco, F. M. (1990). Teaching relational discrimination to individuals with mental retardation: Some problems and possible solutions. *American Journal on Mental Retardation, 95*, 283-296.

Miltenberger, R. G. (2003). *Behavior modification: Principles and procedures* (3rd ed.). Belmont, CA: Wadsworth Publishing.

Nosik, M. R., & Williams, W. L. (2011). Component evaluation of a computer based format for teaching discrete trial and backward chaining. *Research in Developmental Disabilities, 32*, 1694-1702. https://doi.org/10.1016/j.ridd.2011.02.022

Nosik, M. R., Williams, W. L., Garrido, N., & Lee, S. (2013). Comparison of computer based instruction to behavior skills training for teaching staff implementation of discrete-trial instruction with an adult with autism. *Research in Developmental Disabilities, 34*, 461-468. https://doi.org/10.1016/j.ridd.2012.08.01

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

Pollard, J. S., Higbee, T. S., Akers, J. S., & Brodhead, M. T. (2014). An evaluation of interactive computer training to teach instructions to implement discrete trials with children with autism. *Journal of Applied Behavior Analysis, 47*, 765-776. https://doi.org/10.1002/jaba.152

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology, 37*, 8-38. https://doi.org/10.1080/15374410701817808

Roscoe, E. M., & Fisher, W. W. (2008). Evaluation of an efficient method for training staff to implement stimulus preference assessments. *Journal of Applied Behavioral Analysis, 41*, 249-254. https://doi.org/10.1901/jaba.2008.41-249

Rothwell, W., & Sredl, H. (1997). *The ASTD reference guide to professional training roles and competencies* (Vol. *2*). Amherst, MA: HRD Press.

Sarokoff, R. A., & Sturmey, P. (2008). The effects of instructions, rehearsal, modeling and feedback on acquisition and generalization of staff use of discrete trial teaching and student correct responses. *Research in Autism Spectrum Disorders, 2*, 125-136. https://doi.org/10.1016/j.rasd.2007.04.002

Serna, R. W., Foran, M. M., Cooke, C., Hurd, K. E., Tello, A. J., Vangapalli, R., & Hamad, C. D. (2016). Teaching discrete trial training: The effects of asynchronous computer-based instruction on live implementation. *Journal of Special Education Technology, 31*, 39-49. https://doi.org/10.1177/0162643416633334

Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities, 16*, 86-92. https://doi.org/10.1177/108835760101600204

Sterling-Turner, H. E., Watson, T. S., & Moore, J. W. (2002). The effects of direct training and treatment integrity on treatment outcomes in school consultation. *School Psychology Quarterly, 17*, 47-77. https://doi.org/10.1521/scpq.17.1.47.19906

Symes, M. D., Remington, B., Brown, T., & Hastings, R. P. (2006). Early intensive behavioral intervention for children with autism: Therapists' perspectives on achieving procedural fidelity. *Research* *in Developmental Disabilities, 27*, 30-42. https://doi.org/10.1016/j.ridd.2004.07.007

Ward-Horner, J., & Sturmey, P. (2008). The effects of general-case training and behavioral skills training on the generalization of parents' use of discrete-trial teaching, child correct responses, and child maladaptive behavior. *Behavioral Interventions, 23,* 271-284. https://doi.org/10.1002/bin.268

Williams, T. C., & Zahed, H. (1996). Computer-based training versus traditional lecture: Effect on learning and retention. *Journal of Business and Psychology, 11*, 297-310. https://doi.org/10.1007/BF02193865

*Received October 18, 2016*
*Final acceptance May 22, 2017*
*Action Editor, Joshua Jessel*

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website.