

Effects of Computer Simulation Training on In Vivo Discrete Trial Teaching

Sigmund Eldevik · Iwona Ondire ·
J. Carl Hughes · Corinna F. Grindle ·
Tom Randell · Bob Remington

Published online: 12 July 2012
© Springer Science+Business Media, LLC 2012

Abstract Although Discrete-trial Teaching (DTT) is effective in teaching a many skills to children with autism, its proper implementation requires rigorous staff training. This study used an interactive computer simulation program (“DTkid”) to teach staff relevant DTT skills. Participants ($N = 12$) completed two sets of pre-tests either once ($n = 7$) or twice ($n = 5$) before brief DTkid training. These evaluated (a) simulated interactive teaching using DTkid and (b) in vivo teaching of three basic skills (receptive and expressive labeling; verbal imitation) to children with autism. Post-tests showed that DTkid training, rather than repeated testing, was significantly associated with improvements in staff’s ability to implement DTT both within the simulation and in vivo, and that the skills acquired showed both stimulus and response generalization.

Keywords Autism · Discrete-trial teaching · Staff training · Software simulation · DTkid

This research relied on software developed as the result of an award from the Economic and Social Research Council of Great Britain (RES-000-22-0142; Principal Investigator: Prof. Bob Remington).

S. Eldevik (✉)
Oslo and Akershus University College of Applied Sciences,
St. Olavs plass, Postbox 4, 0130 Oslo, Norway
e-mail: eldevik@online.no

I. Ondire · J. C. Hughes · C. F. Grindle
Bangor University, Bangor, UK

T. Randell · B. Remington
University of Southampton, Southampton, UK

Introduction

Discrete-trial Teaching (DTT) is one of the main teaching procedures used in Early Intensive Behavioral Intervention (EIBI) programs for children with autism. It is employed to teach language, social, and academic skills (Leaf and McEachin 1999; Lovaas and Smith 2003; Sturmey and Fitzer 2007). In DTT, the teaching process is broken down into three-term contingency units including a clearly defined instruction, response, and consequence, followed by a pause of no more than 5 s (inter-trial interval). If a learner responds incorrectly, prompts are given. To monitor children’s progress, staff typically record all trials presented as correct, incorrect, or prompted. Although DTT is highly effective when delivered correctly, a number of studies have shown that its utility may be dramatically reduced when staff who deliver it are not adequately trained (Allen and Warzak 2000; Smith and Lovaas 1998; Symes et al. 2006).

These data indicate the importance of developing efficient training procedures to ensure that DTT can be delivered with high levels of fidelity. Such training requires considerable time and resources, especially within larger scale organizations (e.g., Perry et al. 2008). For example, to provide the recommended standard EIBI program of 30 h per week, one child typically requires an intervention team of between two and five therapists (Green 1996). To provide enough staff, many service providers enroll graduate or undergraduate students as therapists for brief periods (i.e., between 2 and 6 months) in a training practicum. Moreover, most providers experience high levels of staff turnover, not least because the work is challenging, and, in many cases, levels of remuneration are relatively low.

High turnover necessarily increases both the demand for, and importance of, efficient and effective staff training.

Although some elements of such training, including grounding in the principles of Applied Behavior Analysis (ABA), can be delivered in a classroom setting, training that relates to actual implementation of teaching procedures must, at some point, be undertaken through real-life interaction with actual children with autism (i.e., *in vivo*; Joyce and Showers 2002). Such training may, however, be problematic, because children with autism can react negatively to new staff, especially if the latter are inexperienced and lack basic teaching and behavior management skills.

A number of researchers have focused on designing and evaluating methods of training staff to improve their implementation of DTT. For example, Koegel et al. (1977) have shown that teachers can acquire DTT skills through training procedures that include the use of training manuals, direct modeling of DTT procedures, and feedback on performance. Such training, however, was reported to have taken up to 25 h to complete. In a similar study, Ryan and Hemmes (2005) used a training package consisting of structured teaching sessions, videotaped instruction, role-playing, and *in vivo* training, supplemented by printed materials on DTT and other autism-related issues. Because of the importance placed by these authors on the acquisition of high levels of declarative knowledge (i.e., knowledge of the principles of behavior, developmental disabilities, professional behavior), training continued until participants achieved 100 % accuracy in written and oral quizzes. This process required between 25 and 35 sessions, each of between 1 and 2 h duration. The findings of both Koegel et al. (1977); Ryan and Hemmes (2005), therefore suggest that achievement of high standards in staff training can potentially raise direct resource costs and lead to short-term service disruption.

Improvement in DTT implementation has, however, also been achieved through other modes of training. Sarokoff and Sturmey (2004), for example, evaluated a training package consisting of instructions, feedback, rehearsal, and modeling. Correct implementation of DTT by all participants increased from an average of 45 % at baseline to 98 %, post training. The number of sessions required to achieve this outcome was not, however, reported and only three participants were trained. It should be noted, nevertheless, that similar training procedures can produce generalization of teaching skills across teaching programs and children (Sarokoff and Sturmey 2008). Furthermore, Arnal et al. (2007) reported use of a self-instructional manual to teach four staff to apply DTT across three tasks. Subsequent to a mean of 2.2 h exposure to the training manual, participants' mean accuracy in using DTT increased from 44 % at baseline to 67 %, but this improvement was measured in a teaching test with a role-playing confederate, rather than with a child with autism.

Although research has shown that traditional forms of training can increase the accuracy and consistency of DTT, it has also shown that such training is typically both costly and time consuming to implement. Moreover, familiarizing staff with the actual practice of DTT is difficult, not least because it requires time to supervise direct practice with children with autism and/or to role-play them. Additionally, the majority of the studies that have evaluated such procedures have not used direct measures of *in vivo* DTT implementation. These considerations, coupled with the high rates of staff turnover, underline the pressing need for a way of delivering practical DTT skills training without the costly involvement of trained clinicians or children with autism.

Randell et al. (2007) reported a potential solution to this problem through the use of DTkid, an interactive computer simulation developed as a training tool for tutors and carers of children with autism. Within DTkid, users interact with "SIMon", a virtual child with autism, and can either receive real-time onscreen feedback on their actions (i.e., "teaching mode") or simply have the accuracy with which they present discrete-trials evaluated within the program (i.e., "evaluation mode"). At present, both "object matching" and "receptive labeling" can be taught and evaluated within DTkid. Using undergraduate participants with no prior experience of behavioral interventions for autism, Randell et al. (2007) showed that DTkid training produced significant improvements in both declarative and procedural knowledge of DTT, in comparison with a control group who engaged in an unrelated interactive computer game for the same amount of time prior to assessment. Declarative knowledge was measured by the accuracy and confidence with which participants categorized video clips of correctly and incorrectly presented individual real-life DTT trials. Procedural knowledge (i.e., participants' ability actually to perform DTT) was assessed using DTkid in evaluation mode.

Although Randell et al.'s (2007) study provided some grounds for confidence that DTkid training increased procedural knowledge; it offered no evidence that the skills learned would transfer to *in vivo* teaching sessions with children with autism. Thus, the present study's primary research question was: Does DTkid training produce stimulus generalization of DTT skills? We hypothesized that simulation training would lead to an improvement in the capacity of staff to implement DTT *in vivo*. A second research question was: Does DTkid training also produce response generalization? We hypothesized that participants who had learned how to teach specific skills (e.g., object matching) using DTkid would show subsequent improvements in their ability to teach other skills (e.g., imitation). The final research question was: Does participants' declarative knowledge of DTT increase as a result of

DTkid training. Based on Randell et al.'s (2007) findings, we hypothesized that this would be the case.

Method

Participants

A convenience sample of 12 novice tutors (“staff participants”; 11 women, 1 man, $M_{age} = 23$ years, age range: 19–28 years) who were to be employed full time at a center providing behavioral interventions in the UK took part in the present research. None had previous knowledge of behavior analysis or experience in implementing DTT or any similar techniques. One held a Masters level degree in psychology and six had Bachelors level degrees in psychology or a related discipline. The remainder had high-school degrees and courses in general child care. All were informed of the nature of the study and gave informed consent prior to participation.

The novice tutors taught three boys with autism, aged 3, 5, and 9 years during the in vivo pre- and post-tests. The boys' diagnoses were made by a clinician independent of the present study. All children had been enrolled in intensive behavioral intervention programs at the center for between 1 and 3 years. None displayed challenging behavior that required any special procedures to be implemented. They were recruited to the study via a letter to their parents, who in turn signed an informed consent form allowing participation and videotaping for scoring purposes.

Setting

The center provided attending children with intensive behavioral intervention for 40 h per week (i.e., weekdays from 9 a.m. to 5 p.m.). The center was organized as a charity and placements funded by local authorities on an individual basis. As part of this provision, the two oldest children participating in the study also attended a local mainstream school for approximately 10 h a week, accompanied by center staff. All training sessions took place at the center.

DTkid training sessions were conducted in an office equipped with a desk, a chair, and a laptop on which DTkid software was installed. The in vivo implementation of DTT was carried out in a therapy room at the center. The room was equipped with desks, chairs and teaching materials required for the various teaching programs.

Apparatus

The DTkid simulation was developed by Randell et al. (2007) with the support of a research grant from the

Economic and Social Research Council of Great Britain. Its design was based on consultation with experienced ABA service providers who were regularly responsible for new staff training. Copyright of the application has been asserted, but it has not been marketed commercially. Those wishing to use DTkid for research purposes should contact the 5th or final author of the present study.

The DTkid software was installed on an Acer Aspire laptop with a 17 inch screen, running Windows XP Professional with peripheral mouse and headphones. A Sony DCR HC36 digital video camera positioned on a tripod was used to record all teaching sessions.

Measures

DTkid software in evaluation mode (see Randell et al. 2007: see Appendix 1) was used to assess participants' DTT competence in both object matching and receptive labeling training within the simulated teaching environment. The Evaluation of Therapeutic Effectiveness (ETE) Scoring Sheet (adapted from Koegel et al. 1977: see Appendix 2) was used to measure participants' competence in implementing DTT. This was done in vivo across three teaching programs, with a child with autism. A measure of declarative knowledge of DTT developed by Randell et al. (2007) was also obtained. This Video Observation Test required participants to view standardized clips of a teacher implementing DTT, categorize her performance in each as correct or incorrect, and report on their confidence in their judgments. Implementation details of all three measures appear in the Procedure section below.

Design and Analysis

The study employed a within-subjects pre-/post-test (AB) design for all 12 participants. For five participants, however, a second pre-test was also carried out prior to DTkid simulation training (i.e., AAB design) to control for the possible impact of repeated testing and exposure to testing materials. Thus, in the control condition one-way within-subjects analyses of variance could be conducted to check for changes between the three testing periods ($n = 5$), followed by within-subjects contrasts to compare both pre-test one with pre-test two and pre-test two with post-test. Paired *t* tests were used to check for changes between pre-test one and the post-test ($N = 12$; the training condition). These comparisons were made for all dependent variables including percentage correct scores on (a) measures of receptive labeling, expressive labeling, and verbal imitation obtained during an in vivo DTT test; (b) measures of object matching and receptive labeling obtained with DTkid in evaluation mode; and (c) declarative knowledge and confidence in scoring on the Video Observation Test.

For both pre-tests and the post-test, order of testing was randomized.

Procedure

All staff participants completed all the pre-tests and post-tests and training across 1 or 2 days (i.e., between 5 and 9 h) when they first joined the center. All pre-tests and the post-test were identical. The tests were all administered at the center and included the following measures. One set of testing took approximately 1 h to complete (i.e., in vivo DTT test: 15 min; DTkid: 25 min; Video Observation Test: 15 min).

In Vivo DTT Test

Participants completed three 2.5 min sessions with a student with autism, during each of which they were instructed by the experimenter to teach the child “to the best of their ability”. Three types of teaching program (receptive labeling, expressive labeling, and verbal imitation) were selected, one for each of the sessions. Center staff selected specific content within each program that the child had yet not mastered (i.e., which objects to use and which sounds or words to present for imitation). The participants were given standardized (scripted) instructions for each program. For the receptive and expressive language programs the following instruction was given: “The child is currently working on improving his vocabulary. Could you, to the best of your ability, teach him to point to (for receptive programs)/name (for expressive programs) this new item (usually a picture of a novel item).” For the verbal imitation program the following instruction was given: “The child is currently working on imitating sounds/words. Could you to the best of your ability teach him to imitate this sound/word.” Although all programs included were considered “basic”, only the receptive labeling program was part of the DTkid simulation training. Thus, the in vivo test provided the opportunity to assess both stimulus generalization (i.e., from DTkid to in vivo teaching) and response generalization (from a familiar to a novel DTT program) in relation to teaching skills acquired using DTkid. All sessions were videotaped for scoring using the ETE scoring sheet (see Appendix 2) which evaluates five different components of DTT: instructions (5 items), prompting (3 items), shaping (1 item), consequences (8 items), and discrete-trial structure (5 items). Participants’ performance in each of the three 2.5 min programs was scored in 30 s intervals from the video recordings. In each interval, an item was scored as correct only if all trials in that interval met the relevant ETE criteria. Items not observed in the interval were categorized as not applicable. On the basis of these data, a percentage score reflecting the

correct application of DTT components within each program was calculated, along with a mean score for each test.

The second author scored the video tapes. Two Master level students in behavior analysis from the University of Bangor were recruited to check reliability of scoring. Before the reliability testing was conducted, the students were trained to an 85 % inter-observer agreement criterion with one of the first two authors. Approximately 20 % of the sessions were randomly chosen and checked for reliability: All were above the 85 % criterion using the formula: $\text{agreements}/(\text{agreements} + \text{disagreement}) \times 100$. The students were blind to whether they scored a pre-test or a post-test session.

DTkid Evaluation Mode

DTkid did not provide any onscreen feedback to users relating to their performance, but simply recorded trials as either correct or incorrect, based on set criteria (see Appendix 1 of Randell et al. 2007). These data were used to calculate the percentage of discrete-trials performed correctly. During DTkid pre-tests, participants first received brief verbal instructions and a demonstration of the use of DTkid from the experimenter (first or second author), including how to use the mouse to select the various training stimuli, and how to select teaching instructions and feedback from the set menus (see Randell et al. 2007, for further details). No verbal instructions concerning the implementation of DTT were provided. Following familiarization with the simulation software, participants were given the opportunity to complete three practice trials with the experimenter present to help resolve any practical issues. Next, each participant worked with DTkid in both object matching and receptive labeling modes, each for 10 min.

Video Observation Test

Identical to Randell et al.’s (2007), measure of participants’ declarative knowledge of DTT, three practice and 24 test video clips (between 11 and 18 s long) were presented, each of a teacher implementing DTT in a program designed to teach object matching to a child with autism (role-played by a typically developing 8 years old boy). Eleven clips showed the teacher presenting a discrete-trial correctly, and 16 clips showed the teacher presenting a discrete-trial incorrectly (see Randell et al. (2007) for a full description of error-types illustrated). Onscreen instructions informed participants to use answer booklets provided to judge each discrete-trial presented as either correct or incorrect, and, as appropriate, to add a brief description error-type. Participants were additionally asked to indicate their confidence in each judgment made on a scale ranging from 0 to 10 (i.e., from “not at all confident” to “very confident”).

Progress through the Video observation test was self-paced. Participants first observed and rated all practice clips, immediately followed by the 24 test clips. Participants indicated to the experimenter when they had completed the test.

Training

Following pre-test(s), participants learned to implement DTT through interaction with DTkid simulation in teaching mode. This involved working through two teaching programs (receptive labeling and object matching), each for 10 min. Participants were verbally instructed that the procedure for using DTkid in teaching mode was the same as in evaluation mode, except that, if they made a procedural error during training, they would receive corrective onscreen feedback (see Randell et al. 2007, for full details) that explained how to correct any mistake made and which would remain on the screen until the correction was made. Only when a trial had been performed correctly, or errors made in a trial had been corrected, would the next DTkid training trial be presented.

Post-test

The post-test was conducted directly following the training, normally on the same day or the next morning.

Results

Mean scores on all study measures for all 12 staff participants, together with test statistics for pre-post comparisons, are presented in Table 1. Four paired sample *t* tests

were conducted on the data from all participants to establish whether, as hypothesized, there were significant improvements in in vivo DTT implementation (as measured by the ETE) between pre-test one and the post-test. Average scores pooled across the three teaching programs improved significantly, $t(11) = 5.9, p < .001$, reflecting improvements in DTT performance on each of the programs considered individually; receptive labeling, $t(11) = 2.6, p = .023$; expressive labeling, $t(11) = 4.5, p < .001$; and verbal imitation, $t(11) = 2.8, p = .017$. Likewise, there was a significant improvement between pre-test one and the post-test in participants' DTkid evaluation mode performance (i.e., teaching within the simulation), $t(11) = 6.9, p < .001$, mirroring improved performance on both receptive labeling, $t(11) = 6.7, p < .001$, and object matching, $t(11) = 6.3, p < .001$. Finally, significant improvements were also seen on the two Video observation tests; declarative knowledge $t(11) = 2.8, p = .016$, and confidence in scoring $t(11) = 2.8, p = .019$.

To evaluate whether the changes reported above were the result of DTkid training, or repeated evaluation, the data from the five individuals who participated in the control condition were analyzed separately. Four one-way within-subject analyses of variance were conducted to assess changes in in vivo teaching between pre-test one, pre-test two, and the post-test. As shown in Table 2, there was a main effect of time on group average scores pooled across teaching programs for the in vivo teaching, and for each of the three individual programs (receptive labeling, expressive labeling and verbal imitation). Within-subjects contrasts on these scores showed, however, that there was no significant change between pre-tests one and two, $F(1, 4) = 1.55, p = .281, \eta^2 = .281$, but marked improvement occurred between pre-test two and the post-

Table 1 Mean pre-test and post-test scores for the full sample (n = 12)

| Outcome variable | Mean score (SD) | | Pre-test 1 versus post-test (n = 12) | | |
|-------------------------------|-----------------|-------------|--------------------------------------|----------|------------------|
| | Pre-test 1 | Post-test | <i>t</i> | <i>p</i> | Cohen's <i>d</i> |
| In vivo teaching | | | | | |
| Pooled across programs | 61.0 (14.6) | 77.0 (12.5) | 5.85 | <.001 | 1.23 |
| Receptive labeling | 62.3 (19.2) | 75.8 (21.6) | 2.64 | .023 | 0.65 |
| Expressive labeling | 56.9 (19.7) | 80.9 (12.6) | 4.50 | <.001 | 1.55 |
| Verbal imitation | 63.6 (16.7) | 74.5 (14.3) | 2.82 | .017 | 0.73 |
| DTkid evaluation mode | | | | | |
| Pooled across programs | 20.4 (14.1) | 62.9 (21.9) | 6.88 | <.001 | 2.40 |
| Object matching | 18.4 (16.8) | 60.5 (24.1) | 6.25 | <.001 | 2.10 |
| Receptive labeling | 22.4 (14.3) | 65.4 (22.4) | 6.66 | <.001 | 2.39 |
| Video observation test | | | | | |
| Declarative knowledge | 71.2 (11.8) | 79.5 (8.5) | 2.84 | .016 | 0.84 |
| Confidence in scoring | 7.4 (1.1) | 8.2 (1.3) | 2.75 | .019 | 1.00 |

Table 2 Mean scores and comparisons at each time point for the control participants ($n = 5$)

| Outcome variables | Mean score (SD) | | | Pre-test 1 versus post-test (n = 5) | | | Pre-test 1 versus pre-test 2 (n = 5) | | | Pre-test 2 versus post-test (n = 5) | | |
|-------------------------------|-----------------|-------------|-------------|-------------------------------------|----------|------------|--------------------------------------|-------------------|------------|-------------------------------------|----------|------------|
| | Pre-test 1 | Pre-test 2 | Post-test | <i>F</i> | <i>p</i> | ηp^2 | <i>F</i> | <i>p</i> | ηp^2 | <i>F</i> | <i>p</i> | ηp^2 |
| In vivo teaching | | | | | | | | | | | | |
| Pooled across programs | 61.8 (15.5) | 67.6 (13.7) | 82.1 (8.7) | 10.80 | .005 | 0.730 | 1.6 | .281 | 0.280 | 8.67 | .042 | 0.684 |
| Receptive labeling | 64.3 (16.8) | 74.2 (7.9) | 85.5 (8.3) | 6.07 | .025 | 0.603 | 1.5 | .283 | 0.277 | 4.15 | .111 | 0.509 |
| Expressive labeling | 58.5 (24.5) | 75.9 (15.3) | 84.8 (3.8) | 6.09 | .025 | 0.604 | 6.9 | .058 | 0.633 | 2.04 | .226 | 0.338 |
| Verbal imitation | 62.4 (20.0) | 52.8 (24.7) | 76.0 (17.9) | 6.00 | .026 | 0.600 | 14.7 | .019 ^a | 0.786 | 7.12 | .056 | 0.640 |
| DTkid evaluation mode | | | | | | | | | | | | |
| Pooled across programs | 19.7 (14.4) | 21.8 (19.2) | 49.7 (27.5) | 11.27 | .005 | 0.738 | 0.4 | .581 | 0.082 | 9.38 | .038 | 0.701 |
| Object matching | 13.5 (12.0) | 19.6 (16.2) | 44.2 (26.8) | 10.93 | .005 | 0.732 | 2.3 | .202 | 0.720 | 10.26 | .033 | 0.720 |
| Receptive labeling | 26.0 (17.5) | 24.0 (23.3) | 55.3 (29.9) | 8.23 | .011 | 0.673 | 1.1 | .735 | 0.032 | 7.62 | .051 | 0.656 |
| Video observation test | | | | | | | | | | | | |
| Declarative knowledge | 74.2 (10.8) | 75.0 (7.2) | 82.5 (4.5) | 2.72 | .126 | 0.405 | | | | | | |
| Confidence in scoring | 8.0 (1.1) | 8.7 (0.7) | 8.7 (1.0) | 2.80 | .120 | 0.411 | | | | | | |

The comparison between pre-test 1 and pre-test 2 constitutes the control for repeated exposures to the tests

^a Pre-test 1 > pre-test 2

test, $F(1, 4) = 8.67$, $p = .042$, $\eta p^2 = .684$. This pattern of results is reflected in the analyses of the individual measures although some variability was observed.

Results of the five control participants' interactions with DTkid in evaluation mode showed the same pattern as for the in vivo teaching (i.e., no changes between pre-test one and pre-test two, but a significant improvement in the post-test). Again, one-way within-subjects analyses of variance were conducted to assess changes in performance with DTkid in evaluation mode between pre-test one, pre-test two, and post-test. There was a main effect of time on the pooled scores and on each of the two teaching programs tested (receptive labeling and object matching). Within-subjects contrasts on the pooled scores showed that there was no significant change between pre-tests one and two; $F(1, 4) = 0.359$, $p = .581$, $\eta p^2 = .082$, but marked improvement occurred between pre-test two and the post-test, $F(1, 4) = 9.38$, $p = .038$, $\eta p^2 = .701$. This pattern of results was also seen in each of the two DTkid teaching programs analyzed separately.

Because no significant changes were detected for the overall effect of time on either declarative knowledge $F(1, 4) = 2.72$, $p = .126$, $\eta p^2 = .405$, or confidence in scoring, $F(1, 4) = 2.80$, $p = .120$, $\eta p^2 = .411$ on the Video observation test for the five control participants, within-subject contrasts on these measures were not conducted.

To establish that the improvements seen for the whole group were not the result of repeated testing, the same analyses were performed for the subgroup of seven participants who did not receive a second pre-test (see Table 3). For these participants, pooled scores across the

three in vivo teaching programs improved significantly consequent to DTkid training, $t(6) = 3.6$, $p = .011$, as did pooled scores on the two DTkid evaluation mode tests, $t(6) = 6.6$, $p = .001$. Enhancements in performance on each of the DTkid programs were also significant, but, despite improvements on each the in vivo programs (generally with medium to large effect sizes) only improvement on the expressive labeling program, which had the lowest baseline score, reached significance, $t(6) = 3.4$, $p = .015$. Improvements in the Video observation test measures also approached significance $t(6) = 2.3$, $p = .061$ for declarative knowledge and $t(6) = 2.1$, $p = .084$ for confidence in scoring.

Discussion

The present study was designed to evaluate the utility of the DTkid simulation program to train staff to implement DTT in vivo with children with autism. Results indicated that simulation training on two teaching programs resulted in significant improvement in the accuracy with which participants' implemented DTT within DTkid in evaluation mode, and, crucially, when working directly with a child with autism in a real-world teaching setting. These latter, in vivo, improvements were observed across three programs, even though only one of these (receptive labeling) had been specifically practiced using the DTkid simulation. In keeping with previous findings (Randell et al. 2007), results also indicated significant increases in participants' declarative knowledge of DTT and their confidence in that

Table 3 Mean pre-test and post-test scores for the participants that only had one pre-test ($n = 7$)

| Outcome variables | Mean score (SD) | | Pre-test 1 versus post-test ($n = 7$) | | |
|-------------------------------|-----------------|-------------|---|----------|------------------|
| | Pre-test 1 | Post-test | <i>t</i> | <i>p</i> | Cohen's <i>d</i> |
| In vivo teaching | | | | | |
| Pooled across programs | 60.4 (15.2) | 73.4 (14.2) | 3.60 | .011 | 0.90 |
| Receptive labeling | 60.9 (22.0) | 68.9 (25.9) | 1.01 | .352 | 0.34 |
| Expressive labeling | 55.8 (17.5) | 78.0 (16.2) | 3.38 | .015 | 1.39 |
| Verbal imitation | 64.5 (15.5) | 73.4 (12.6) | 2.04 | .088 | 0.67 |
| DTkid evaluation mode | | | | | |
| Pooled across programs | 20.9 (15.1) | 72.4 (11.1) | 6.56 | .001 | 4.00 |
| Object matching | 22.0 (19.7) | 72.2 (14.2) | 5.53 | .001 | 3.03 |
| Receptive labeling | 19.8 (12.4) | 72.5 (13.5) | 6.69 | .001 | 4.24 |
| Video observation test | | | | | |
| Declarative knowledge | 69.1 (12.9) | 77.3 (10.2) | 2.31 | .061 | 0.73 |
| Confidence in scoring | 6.9 (1.0) | 7.9 (1.4) | 2.07 | .084 | 1.00 |

knowledge consequent to DTkid training. The subgroup of participants that received only one pre-test showed similar post-test improvements to those seen in the overall group, whereas the subgroup of participants who took part in a second pre-test conducted as a control condition showed no such improvements in either DTkid evaluation mode or in vivo performance between pre-test one and pre-test two (i.e., prior to DTkid training), but did show an overall improvement at post-test. This pattern of results strongly suggests that the improvements observed following DTkid training were not simply the result of repeated testing.

The results observed therefore indicated that, as hypothesized, DTkid training produced both stimulus and response generalization: Participants learned the teaching skills involved in object matching and receptive labeling using DTkid, and improvements were seen when these skills were tested within the simulation. It is also of considerable interest that improvements in receptive labeling observed using DTkid in evaluation mode were mirrored by improvements in participants' ability to teach receptive labeling in vivo. This can be regarded as indicating the kind of stimulus generalization expected from any effective simulation training (e.g., pilot instruction in a flight simulator). Additionally, however, participants showed improvements in delivering two kinds of DTT programs that had not been specifically trained in the simulation—expressive labeling and verbal imitation. This type of response generalization indicates that, subsequent to DTkid training, participants were able to use the basic elements of DTT across novel situations. It thus appears that participants had acquired a broader response class of general teaching skills, learning, for example, how to present material, when to prompt, and how to administer

contingent consequences for the child's responding. Although no formal assessment of social validity was attempted, responses to the training program at debriefing were uniformly positive.

These findings therefore also extend those of previous research that used only DTkid in evaluation mode to measure the impact of simulation training (Randell et al. 2007) by showing that DTkid training can also result in the emergence of in vivo procedural knowledge (i.e., teaching skills in real-life settings).

As hypothesized, and keeping with the findings of Randell et al. (2007), significant increases in participants' declarative knowledge of DTT, and in their confidence in that knowledge, were observed. The magnitude of such increases was smaller in the present research, however, perhaps owing to the substantially higher pre-test scores obtained; 71.2 % correct, compared with 59.3 % for the control group participants in the Randell et al. (2007) study. This suggestion is also supported by the fact that post-test scores for the DTkid group reported by Randell et al. (2007) study were virtually identical to those observed in the present study; respectively, 80 and 79.5 % correct. The discrepancy in scores between those obtained by the control group in Randell et al. (2007) and the pre-test scores in the present study requires further explanation. It is possible that the latter scores were higher because participants recruited by the center where the work was based had chosen to work with children with autism. In other words, our sample was characteristic of individuals who would typically receive training in the use of instructional methods with children with autism.

In the light of the baseline differences in participant knowledge and confidence between Randell et al. (2007)

and the present study, it is worth noting that, although our participants were not complete novices, their skills were nevertheless enhanced by DTkid training. Mean pooled in vivo scores for participants prior to any training was 61 % correct (range: 56.9–63.6 %). Therefore, the significant changes in real life teaching seen reflected improvements from a relatively high baseline, and required only very brief DTkid training (20 min in total on two basic teaching programs).

Some limitations of the present study should be noted. Firstly, use of more comprehensive methods for evaluating the quality of participants' DTT performance would have been desirable. Although a revised version of scoring procedures developed by Koegel et al. (1977) was employed, this measure has not been standardized or validated. Although no alternative was available at the time that the research was carried out, this measure does have the advantage of being in widespread use among behavioral clinicians working in the field of autism. Secondly, control procedures for repeated testing were limited by practical considerations. Because of this, although the subgroup of participants that received repeated pre-tests served the same purpose as a waiting list control group, the times of the first and second tests were not aligned between groups in the manner typical of larger scale studies and the data obtained could not be analyzed as a between-group design. Thus, although changes in performance following interaction with DTkid were significant, inferences of a causal role for the simulation must be made with caution until replication using standard randomized control procedures has been achieved using a larger number of participants. Despite such limitation, the present results would nevertheless appear sufficient to justify investment in the further development of DTkid as a simulation methodology and its further evaluation in controlled trials.

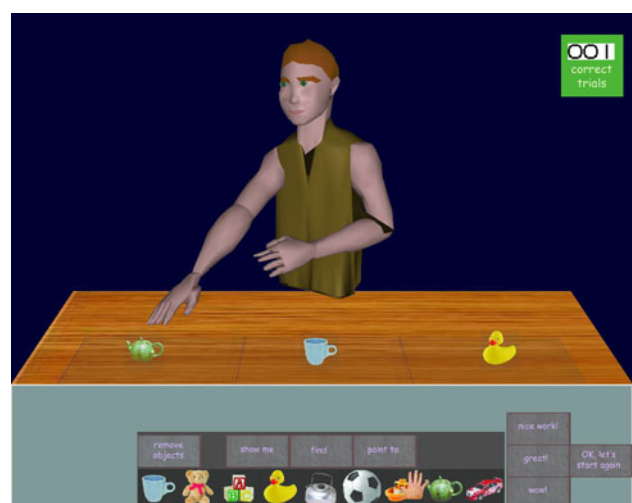
In summary, a simple simulation program, DTkid, has been shown to be an effective way of teaching inexperienced staff to conduct relatively simple programs with children who are familiar with DTT. DTkid's efficacy in teaching more advanced programs and procedures is unknown because the capabilities of the software are not yet sufficiently developed. In principle, however, it should be possible to use simulations to teach a novice therapist how to conduct a teaching situation with a child who is just starting on an early intervention program, how to establish a token economy system, how to pace teaching sessions, and how to handle challenging behaviors. As yet, however, DTkid's ability to teach these important skills awaits development.

In conclusion, the results of the present research indicate that the DTkid allows some of the key elements of DTT to be taught and evaluated effectively and efficiently via short sessions of computer simulation. This offers the potential to free up resources for service providers and to eliminate, or markedly reduce, the problems that arise when children with autism are taught by novice tutors. Because simulation training is unlikely ever to cover all aspects of staff education, however, it remains essential to identify which skills can best be taught by such means, and to consider how simulation can most effectively be used alongside other evidence-based techniques such as modeling, group training, lectures, and direct feedback. In this way the amount of staff training involving real children may be reduced, and available resources targeted towards the use of other methods to establish skills that cannot be taught using simulation techniques.

Acknowledgments We would like to express our appreciation The Sharland Foundation for supporting this research. We would also like to thank Rachel Lowe and Dimitra Tavoulari, Masters students in Behavior Analysis at Bangor University, UK, for helping with data collection and scoring.

Appendix 1

This screen shot of the DTkid simulation shows the child working on a receptive labeling program in evaluation mode. The teacher chooses from the panel below which objects to put on the table, what instruction to give and depending on the child's response what consequence to administer.



Appendix 2

The Evaluation of Therapeutic Effectiveness (ETE) Scoring Sheet (adapted from Koegel et al. 1977).

| Therapist: | | | |
|--------------------|-----------|------------------------------|--|
| Child: | | | |
| Date: | | | |
| # | Score | code/position/comment/target | |
| 1 | C I P | | |
| 2 | C I P | | |
| 3 | C I P | | |
| 4 | C I P | | |
| 5 | C I P | | |
| 6 | C I P | | |
| 7 | C I P | | |
| 8 | C I P | | |
| 9 | C I P | | |
| 10 | C I P | | |
| 11 | C I P | | |
| 12 | C I P | | |
| 13 | C I P | | |
| 14 | C I P | | |
| 15 | C I P | | |
| 16 | C I P | | |
| 17 | C I P | | |
| 18 | C I P | | |
| 19 | C I P | | |
| 20 | C I P | | |
| 21 | C I P | | |
| 22 | C I P | | |
| 23 | C I P | | |
| 24 | C I P | | |
| 25 | C I P | | |
| 26 | C I P | | |
| 27 | C I P | | |
| 28 | C I P | | |
| 29 | C I P | | |
| 30 | C I P | | |
| 31 | C I P | | |
| 32 | C I P | | |
| 33 | C I P | | |
| 34 | C I P | | |
| 35 | C I P | | |
| 36 | C I P | | |
| 37 | C I P | | |
| 38 | C I P | | |
| 39 | C I P | | |
| 40 | C I P | | |
| 41 | C I P | | |
| 42 | C I P | | |
| 43 | C I P | | |
| 44 | C I P | | |
| 45 | C I P | | |
| 46 | C I P | | |
| 47 | C I P | | |
| 48 | C I P | | |
| 49 | C I P | | |
| 50 | C I P | | |
| Coding - Behaviour | | | |
| 1 | | | |
| 2 | | | |
| 3 | | | |
| 4 | | | |
| 5 | | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--|---|-------|-------|-------|----|-----|---------------|--|--|--|--|----|----|----|----|-------|----------------|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| | <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%;">Pre</td> <td style="width: 15%;">Post1</td> <td style="width: 15%;">Post2</td> <td style="width: 15%;">IT</td> <td style="width: 15%;">Obs</td> </tr> <tr> <td colspan="5" style="text-align: center;">Session type?</td> </tr> <tr> <td>MT</td> <td>DT</td> <td>ET</td> <td>RR</td> <td>Other</td> </tr> <tr> <td colspan="5" style="text-align: center;">Type of trial?</td> </tr> </table> | Pre | Post1 | Post2 | IT | Obs | Session type? | | | | | MT | DT | ET | RR | Other | Type of trial? | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Pre | Post1 | Post2 | IT | Obs | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Session type? | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| MT | DT | ET | RR | Other | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Type of trial? | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>What is the SD?</p> <div style="border: 1px solid black; height: 20px; width: 100%;"></div> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>What is the target response?</p> <div style="border: 1px solid black; height: 20px; width: 100%;"></div> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>1. SD's</p> <p>a. Discriminable. Stands out. Distinct onset and offset. Does not occur together with name or verbal clutter. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>b. Appropriate. Teacher asks for the target. Does not mistake/only ask for wrong item. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>c. Consistent. Uses the predetermined SD's consistently. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>d. Uninterrupted. SD's are not interrupted by instructions to sit nicely or other behaviour from the teacher. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>e. Attending. The child should be attending to the teacher or the material when SD is issued. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>2. Prompts</p> <p>a. Effective. Must produce a correct response. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>b. Swift. Within 5 seconds and, if possible so that the child is not allowed time to make an incorrect response. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>c. Frequent. The child should not be allowed to make more than two incorrect responses before prompted. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>3. Shaping</p> <p>a. Approximation. Should only reinforce responses at least as good or closer approximations to target. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>4. Consequences</p> <p>a. Immediate. Within 3 seconds. Preferably within 0.5 seconds. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>b. Contingent. Reinforces only correct responses, not incorrect ones. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>c. Unambiguous. Praise should sound positive. Information should sound neutral. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>d. Consistent. All correct responses should be reinforced during acquisition. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>e. Effective tangibles (if used) Should be approached/readily taken by the child. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>f. Paired with praise. Praise should always come prior to the tangible reinforcer. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>g. Size. The amount/size of reinforcement is related to the child's effort. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>h. No response. If there is no response from the child in 5 seconds, teaching is continued. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>5. Structure</p> <p>a. Learn unit. The trial constitutes a discrete learn unit (SD-R-SR). 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>b. Ending positive. The session is ended with a correct response. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>c. Opportunity. After a prompted response, the child is given opportunity to do the task without prompt. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>d. Pace. Inter trial interval should not be more than 5 seconds (including onset from starting session). 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> <p>e. Variation. Except when when MT, no task should be repeated more than twice consecutively in the same way if performed correctly. 30-sec recording interval</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td><td style="width: 15%;"></td></tr> </table> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Overall average score | | | | | % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

References

- Allen, K. D., & Warzak, W. J. (2000). The problem of parental nonadherence in clinical behavior analysis: Effective treatment is not enough. *Journal of Applied Behavior Analysis, 33*, 373–391.
- Arnal, L., Fazzino, D., Martin, G. L., Yu, C. T., Keilback, L., & Starke, A. (2007). Instructing university students to conduct discrete-trials teaching with confederates simulating children with autism. *Developmental Disabilities Bulletin, 35*, 131–147.
- 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). Austin, TX: Pro-Ed.
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Koegel, R. L., Russo, D. C., & Rincover, A. (1977). Assessing and training teachers in the generalized use of behavior modification with autistic children. *Journal of Applied Behavior Analysis, 35*, 209–211.
- Leaf, R., & McEachin, J. (1999). *A work in progress: Behavior management strategies and a curriculum for intensive behavioral treatment of autism*. New York: DRL Books.
- Lovaas, O., & Smith, T. (2003). Early and intensive behavioral intervention in autism. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 325–340). New York, NY: Guilford Press.
- Perry, A., Cummings, A., Geier, J. D., Freeman, N. L., Hughes, S., LaRose, L., et al. (2008). Effectiveness of intensive behavioral intervention in a large, community-based program. *Research in Autism Spectrum Disorders, 2*, 621–642.
- Randell, T., Hall, M., Bizo, L., & Remington, B. (2007). DTkid: Interactive simulation software for training tutors of children with autism. *Journal of Autism and Developmental Disorders, 37*, 637–647.
- Ryan, C. L., & Hemmes, N. S. (2005). Post-training discrete-trial teaching performance by instructors of young children with autism in early intensive behavioral intervention. *The Behavior Analyst Today, 6*, 1–12.
- 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.
- 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.
- Smith, T., & Lovaas, O. I. (1998). Intensive and early behavioral intervention with autism: The UCLA young autism project. *Infants and Young Children, 10*, 67–78.
- Sturmey, P., & Fitzer, A. (2007). *Autism spectrum disorders: Applied behavior analysis, evidence and practice*. Austin, TX: Pro-Ed.
- 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.