Journal of
Applied Behavior Analysis
JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2015, **48,** 38–51 NUMBER 1 (SPRING)

TEACHING CHILDREN WITH AUTISM TO DISCRIMINATE THE REINFORCED AND NONREINFORCED RESPONSES OF OTHERS: IMPLICATIONS FOR OBSERVATIONAL LEARNING

Jaime Ann DeQuinzio and Bridget A. Taylor

ALPINE LEARNING GROUP

We taught 4 participants with autism to discriminate between the reinforced and nonreinforced responses of an adult model and evaluated the effectiveness of this intervention using a multiple baseline design. During baseline, participants were simply exposed to adult models' correct and incorrect responses and the respective consequences of each. During discrimination training, in the presence of target pictures, we taught participants to imitate the reinforced responses of an adult model and to say "I don't know" when an adult model's response was not reinforced. Test sessions were conducted after baseline, discrimination training, and generalization sessions to measure responding to target pictures in the absence of the model, prompts, and reinforcement. All 4 participants showed acquisition in the discrimination of reinforced and nonreinforced responses of the adult model during test sessions. Generalization to stimuli not associated with training was variable across the 4 participants. Implications for teaching observational learning responses to children with autism are discussed.

Key words: observational learning, autism, discrimination of contingencies

---

Observational learning has been defined as the learning of new responses that occurs as a result of observing the responding of a model and the consequences that this responding produces (Catania, 1998). The potential benefits of observational learning cannot be understated; if one can learn by observing the responses of others, it may reduce instruction time, maximize learning opportunities in mainstream learning environments, and permit the acquisition of novel information without explicit instruction (Delgado & Greer, 2009; Taylor & DeQuinzio, 2012). As several researchers have noted, however, observational learning may require some specific prerequisite skills, such as attention to a model, generalized imitation, and the demonstration of complex discriminations (Deguchi, 1984; Greer et al., 2006; Masia & Chase, 1997). In other words, successful observational learning is predicated on more than simple observation.

There is some evidence that children with autism may show deficits in the prerequisite skills necessary for observational learning (i.e., attending, imitation, and demonstrating complex discriminations). For example, studies have shown that many children with autism demonstrate poor visual attending, including inconsistent eye contact, failure to follow eye gaze, and failure to orient toward toys and materials (Donnelly, Luyben, & Zan, 2009; Leekam, Hunnisett, & Moore, 1998). In addition, a review of 21 studies on imitation of children with autism concluded that imitative deficits were apparent in some younger children (Williams, Whiten, & Singh, 2004). Further, researchers have documented that individuals with autism often have difficulty making auditory and visual discriminations (e.g., Plaisted, O'Riordan, & Baron-Cohen, 1998; Schreibman, 1975). Given these deficits, children with autism might

Portions of these data were presented at the annual conference of the Association for Behavior Analysis International in Minneapolis, Minnesota. We thank the teachers, students, and families of the Alpine Learning Group who participated in this study and Sheiba Koshy for her assistance with organizing sessions and data collection.

Address correspondence to Jaime A. DeQuinzio, Alpine Learning Group, 777 Paramus Road, Paramus, New Jersey 07652 (e-mail: Jdequinzio@alpinelearninggroup.org).

doi: 10.1002/jaba.192

experience difficulty learning by observing the responses of others and the resulting consequences associated with these responses.

Early studies have demonstrated that children with autism consistently fail to learn by simply observing the responses of others or by watching others engage in instructional interactions with teachers (Lovaas, Koegel, Simmons, & Long, 1973; Varni, Lovaas, Koegel, & Everett, 1979). For example, Varni et al. (1979) found that children with autism performed significantly worse than their age-matched typical peers on observational learning tasks. Pretests were conducted to determine a baseline level of responding to one-word instructions. During observational learning sessions, participants with autism observed the teacher provide instructions to an adult model. The adult model then engaged in the correct response, and the teacher provided reinforcement for that response. Finally, posttests were conducted in which the teacher presented the same instructions to the participants with autism in the absence of the model. Results showed that children with autism engaged in fewer of the observed responses than did the typical children. The researchers argued that these deficits may be related to a failure to discriminate or attend to relevant stimuli during the observation sessions. In other words, the participants missed opportunities to learn by observing others because of these attending and discrimination deficits.

Because of the early work of Bandura and colleagues (Bandura, 1965; Bandura & McDonald, 1963; Bandura, Ross, & Ross, 1963), it has become widely known that observational learning is an important part of human development with respect to the learning of new responses. Furthermore, researchers have argued that if one can indeed learn by observing the responses of others, it may reduce instruction time, maximize learning opportunities in mainstream learning environments, and permit the acquisition of novel information without explicit instruction (Greer et al., 2006; Ledford, Gast, Luscre, & Ayres, 2008). These benefits of observational learning combined with the prevalence of observational learning deficits among children with autism is precisely why research in the area of observational learning is essential. However, to date only a handful of studies have examined procedures to teach or assess observational learning in this population. In fact, studies of observational learning have primarily focused on individuals with general developmental disabilities (Christensen, Lignugaris-Kraft, & Fiechtl, 1996; Werts, Caldwell, & Wolery, 1996; Wolery, Ault, Gast, Doyle, & Griffen, 1991). These studies evaluated a variety of procedures, such as having participants observe competent models perform responses (e.g., Werts et al., 1996) or having participants observe instruction between a model and a teacher (e.g., Christensen et al., 1996).

In one of the earliest studies to incorporate children with autism, Egel, Richman, and Koegel (1981) examined the effects of observing peer models perform the correct responses on the learning of discrimination tasks. Modeled responses were followed by praise and reinforcement from a teacher. In a baseline assessment of the skills, the model was not present and the teacher praised and prompted the participants' correct responses. During intervention, participants were verbally instructed to attend to the materials and the model. Immediately after each response from the model, the teacher provided the same instruction to the participant. The teacher praised correct responses and manually guided imitative responses if necessary. Participants' responding was assessed again without the model present. Results indicated that all participants learned color, shape, and word discriminations during the peer-modeling condition, and responding was maintained when the model was removed.

Building on the Egel et al. (1981) study, Charlop, Shreibman, and Tryon (1983) compared the relative impacts of peer modeling and traditional trial-and-error instruction on the acquisition of a receptive labeling task. During baseline, no model was present. During intervention, the participant observed a peer perform responses and receive an edible item and praise for responding correctly. Although the model was trained to respond with 100% accuracy before the study, if he or she responded incorrectly, the teacher said "no." In addition, if participants' attending appeared to wander, they were told by the adult to "pay attention." During posttests, each participant was tested on the receptive labeling tasks that the peer had just modeled. Results indicated that peer modeling resulted in greater generalization across settings and instructors and greater maintenance 3 days after the completion of the study than instruction involving trial and error.

In another study that involved participants with autism, Rehfeldt, Latimore, and Stromer (2003) examined the effects of observational learning on the development of stimulus classes. Participants were given direct training on conditional discriminations (i.e., matching dictated word to corresponding picture and matching dictated word to printed word) with sets of stimuli from three categories (electrical appliances, occupations, and kitchen appliances). Every three trials, an adult peer model or sibling was taught, using manual prompts and reinforcement, to match dictated word to picture and dictated word to printed word using stimuli from two different categories (transportation and actions). During these sessions, participants were verbally instructed to attend to the model, and attending responses were praised. Results indicated that all participants demonstrated full stimulus classes by matching picture to printed word with the stimuli involved in direct training. However, they could not do so with the stimuli associated with observation of a model. In a second study using stimuli from the same larger superordinate category (electrical appliances, occupations, and kitchen appliances) for both the direct training and observation sessions, participants demonstrated full stimulus classes by matching picture to printed word with those involved in direct training and at least one of the model's sets of stimuli. The authors concluded that stimulus classes may be acquired more readily when the stimuli used in direct training and those used in the modeling sessions are from the same category rather than different categories.

These studies demonstrate that some individuals with a variety of developmental disabilities, including autism, can acquire novel responses as a result of observing the behavior of a model. What remains unclear, however, is what mechanisms are responsible for the acquisition of these responses that are acquired via observational learning. For example, across studies, it cannot be determined whether increased responding was attributable to prompts for attention toward the model, praise provided to the model, or opportunities to imitate the model during observation sessions. Further, none of these studies focused on the components of observational learning as a dependent measure; that is, they did not focus on direct teaching of observational learning responses.

Two recent studies have attempted to isolate component responses related to observational learning and to determine the effects of learning these responses on the emergence of new responses during observational learning sessions (Delgado & Greer, 2009; Taylor, DeQuinzio, & Stine, 2012). Taylor et al. (2012) evaluated a procedure for teaching children with autism to learn new sight words by watching a peer engage in an instructional interaction with a teacher. The authors argued that for observational learning to occur, children with autism might first need to learn to monitor the responses of a model. Monitoring responses consisted of imitating the model's response immediately after the model's production of the response and engaging in a matching response that indicates attention to relevant instructional stimuli. To test their hypothesis, the researchers trained monitoring responses with one set of stimuli and did not use monitoring training with another set of stimuli. In the training condition in which monitoring responses were taught, the participants acquired sight words faster than in a probe condition in which monitoring responses were not taught. These results strongly suggested that learning to monitor the responses of a model may have facilitated the subsequent observational learning of all three participants. As the authors noted, however, the study was critically limited by the fact that participants observed only the responses of competent models; in every trial, peers produced correct responses that were followed by reinforcement from the teacher. Observational learning, the authors acknowledged, involves more than simply attending to and imitating the correct responses of peers.

Delgado and Greer (2009) evaluated an intervention that taught children with autism to learn new sight words by monitoring both correct and incorrect responses of peers. In this study, monitoring consisted of counting the correct and incorrect responses of peers as indicated by teacher feedback to the peer. Specifically, participants were taught to choose a green block when a peer modeled a correct response (as indicated by the teacher's feedback) and to choose a red block when a peer modeled an incorrect response (as indicated by the teacher's feedback). The authors concluded that only when participants learned to monitor both the correct and incorrect responses of their peers during observation sessions did they show acquisition of new sight words. These results are promising because they indicate that children with autism can be taught to discriminate the correct and incorrect responses of others and that perhaps learning this discrimination is a fundamental precondition for observational learning to occur (Deguchi, 1984; Fryling, Johnston, & Hayes, 2011; Masia & Chase, 1997).

However promising these studies may have been, their results have yet to be replicated or expanded, and the research literature still lacks more rigorous examinations of the component responses associated with observational learning. With the preliminary results of Delgado and Greer (2009) in mind, we wondered if children with autism could learn to discriminate between the contingencies applied to a model by engaging in responses directly related to the instructional process. Rather than training participants to demonstrate discrimination of contingencies with red or green blocks, we wanted participants to demonstrate that discrimination of contingencies had occurred by imitating those responses of a model that were followed by reinforcement and not to imitate modeled responses that were not followed by reinforcement and to instead use an alternative response ("I don't know"). Specifically, we wanted to assess if participants could learn to say the correct names for pictures associated with reinforcement and to say "I don't know" for pictures associated with no reinforcement and a corrective statement made toward the model (e.g., "No that's wrong"). Thus, the purpose of the current study was to determine the extent to which children with autism could (a) learn to discriminate the reinforced and nonreinforced responses of an adult model by imitating the correct picture labels produced by a model and by saying "I don't know" when the adult model used an incorrect picture label, (b) to demonstrate these discriminations 1 and 10 min after observation sessions in the absence of the model, and (c) to engage in these observational learning skills in the presence of unknown pictures that had never been associated with discrimination training and reinforcement.

METHOD

Participants and Setting

The participants were four children with autism who attended a behaviorally based school for children with autism, had experience with token economies, and had a history of learning simple discriminations as well as match-to-sample conditional discriminations. In addition to one-to-one instruction, all of the participants were involved in small-group instruction for some portion of the school day during which they were instructed on various academic subjects such as reading, math, and spelling. All sessions took place in the participants' classrooms. Teachers who were familiar with the participants and trained in the use of applied behavior analysis as well as the procedures for this study conducted all sessions.

All four participants were chosen for this study because they all had goals related to observational learning included in their current individualized educational plans. Rose was 6 years old at the time of the study. Her instructional programs included handwriting, reading, spelling, basic math, drawing, answering questions, making comments, and social interactions with peers and adults. In addition, she was learning to brush her teeth, to prepare and eat her lunch, to cooperate with a haircut, and to tie her shoes. She also attended kindergarten in a regular education classroom for part of the day with the help of a personal teaching assistant. Her age-equivalent score on the Picture Peabody Vocabulary Test (PPVT) was 4 years 10 months. Her age-equivalent score of the Expressive Vocabulary Test (EVT) was 5 years 3 months.

Leah was 9 years old at the time of the study. Her instructional programs included basic math, reading, spelling, writing, answering questions, following complex instructions, completing tasks independently, and engaging in social interactions with peers and adults. In addition, she was learning self-help and leisure skills such as playing the piano, riding a bike, taking pictures, and using the restroom independently in the community. Her age-equivalent score on the PPVT was 4 years 9 months. Her age-equivalent score on the EVT was 4 years 4 months.

Mary was 12 years old at the time of the study. Her instructional programs included completing daily routines, answering and asking questions, math, describing objects, and commenting. In addition, she was learning self-help, leisure, community, and vocational skills such as braiding her hair, using a cellular phone, preparing meals, taking inventory, typing on a keyboard, purchasing items in the community, taking photographs, playing the guitar, running on the treadmill, and playing board games with peers. Her age-equivalent score on the PPVT was 5 years 5 months. Her age-equivalent score on the EVT was 5 years 4 months.

Mark was 7 years old at the time of the study. His instructional programs included answering questions and responding to comments, handwriting, reading comprehension, telling time, basic math and phonics skills, using correct verb tense, sustaining eye contact during social interactions, using polite language during interactions with others, and completing tasks independently. He was also learning to brush his teeth, to recognize and express emotions, to play a board game with peers, to ride a bike, and to type using a keyboard. His age-equivalent score on the PPVT was 6 years 3 months. His age-equivalent score on the EVT was 6 years 4 months.

Design

A multiple baseline design across participants was used to evaluate the effects of discrimination training on the acquisition of unknown picture labels and saying "I don't know" during training and generalization test sessions.

Dependent Measures

The main dependent measure was summarized as the percentage of correct responses during test sessions conducted after baseline, discrimination training, and generalization sessions. This measure included both correctly labeling pictures from the reinforced trials and saying "I don't know" when presented with pictures from the nonreinforced trials.

Test sessions were conducted 1 or 10 min after each baseline, discrimination training, and generalization session (described below). Test sessions were initially conducted 10 min after baseline, discrimination training, and generalization sessions; however, because Rose did not demonstrate the discrimination under these conditions, we implemented 1-min test sessions with all subjects and then increased to 10-min test sessions as responding increased. During the test sessions, the teacher presented the pictures used in the prior baseline, discrimination training, or generalization session. The six pictures were presented randomly three times each. The model was not present during these sessions. During each trial, the teacher sat across from the participant, held a picture in view, and said "What's this?" The teacher then recorded the participant's response without providing reinforcement or prompts and moved on to the next trial. A fixed-interval (FI) 30-s schedule of token reinforcement was provided for sitting quietly with hands down throughout the session.

Materials

A total of 12 pictures unknown to the participant were presented on index cards (7.6 cm by 12.7 cm). The pictures were identified as unknown to the participant from a pool of over 30 pictures during a preassessment and consisted of the following images: kidney, eclipse, compass, flash drive, tendon, tweezers, wishbone, toner, strainer, stick shift, caveman, and sea sponge. All 12 pictures were randomly assigned to be used during either discrimination training or generalization sessions, for a total of six pictures in each session type. Thus, the ratio of training to probe stimuli was one to one. One set of pictures was used during discrimination training (i.e., kidney, eclipse, compass, flash drive, tendon, and tweezers), and the other set was used during generalization sessions (i.e., wishbone, toner, strainer, stick shift, caveman, and sea sponge). These six pictures in the training and generalization sessions were further divided by randomly assigning them into two groups of three stimuli to be used on either reinforced or nonreinforced trials to teach and assess for the discrimination of reinforced versus nonreinforced responses. During discrimination training, the pictures used on reinforced trials were kidney, eclipse, and compass, and the pictures used on nonreinforced trials were flash drive, tendon, and tweezers. During generalization, the pictures used on reinforced trials were wishbone, toner, and strainer, and the pictures used on nonreinforced trials were stick shift, caveman, and sea sponge. In addition, on all nonreinforced trials, the incorrect responses provided by the adult model were based on unknown picture labels and were randomly arranged for each nonreinforced trial. The adult model responded with "blower," "lumber," or "dipstick" on nonreinforced trials during discrimination training and "oboe," "mummy," or "tunic" on nonreinforced trials during generalization. There was one exception with respect to the incorrect responses arranged for Mary; we modified the incorrect responses used by the model to match the incorrect labels that Mary was using during test sessions for the pictures in the nonreinforced category. For example, Mary always responded "hairclip" when presented with the picture of the tweezers during baseline and throughout discrimination training. Therefore, the model responded with the incorrect answer ("hairclip") on the nonreinforced trials.

Individualized motivational systems were used to reinforce on-task behavior. Reinforcers used during observation sessions, such as iPads, markers, and books, were chosen based on teacher reports of items that were commonly used as reinforcers in the classroom.

Procedure

Preassessment. To ensure that the pictures used were initially unknown to the participants, we conducted a preassessment. We presented over 30 pictures randomly, three times each and asked "What's this?" If the participant failed to label the picture on all three trials, we considered it unknown. We chose 12 target pictures to use in the study that were unknown to all three participants. Pictures were assigned to the discrimination training and generalization conditions, as described in the materials section above.

Baseline. We conducted baseline sessions to expose the participants to the reinforced and nonreinforced responses of the adult model so that we could evaluate performance during test sessions before implementation of discrimination training. The participant and adult model sat next to each other at a table, and the teacher sat across from them. An adult model was chosen instead of a peer so that we could control for the arrangement of correct and incorrect responses. A reinforcer was placed on the table in view of both the participant and adult model. The teacher held a picture between the two and said "What's this?" to the adult model. The adult model responded, and the teacher provided a programmed consequence depending on the trial type. There were two trial types that differed with respect to programmed consequences: reinforced trials and nonreinforced trials. The order of reinforced and nonreinforced trials was randomized. During reinforced trials, the adult model answered correctly and the teacher provided positive feedback (e.g., "That's right!" "Yes!" "You got it!" etc.) as well as a reinforcer (e.g., iPad). The teacher never used the picture label in the praise statement. During nonreinforced trials, the adult model answered incorrectly and the teacher provided negative feedback, (e.g., "I'm sorry, that's wrong," "Incorrect," "That's not right") and removed the reinforcer from the table. The teacher never provided the correct label to the adult model. An FI 30-s schedule of token reinforcement was also used to reinforce sitting quietly with hands down throughout the session.

Discrimination training. Training was conducted to teach participants to discriminate the reinforced and nonreinforced responses of the adult model. The participant and adult model sat next to each other at a table, and the teacher sat across from them. The arrangement and presentation of trials as well as the teacher's interaction with the adult model were similar to that of baseline, with one exception. During these sessions, the teacher implemented discrimination training with the participant after providing the consequence to the adult model.

After providing the consequence to the adult model on reinforced trials, the teacher immediately presented the picture to the participant and said "What's this?" If the participant imitated the adult model's response, the teacher provided praise (e.g., "Good, you said what she said when she got it right") and delivered the reinforcer to the participant. If the participant answered incorrectly or did not answer at all within 3 s, the teacher implemented an error-correction procedure. First the teacher stated a rule, "Say what she says when she gets it right," and then re-presented the picture to the adult model. Immediately after the modeled response and consequence, the teacher asked the participant to label the picture again. If the participant answered correctly on this error-correction trial, general praise, but not the reinforcer, was provided. If the participant still did not answer correctly, the trial was terminated and the teacher moved on to the next trial.

After providing the consequence to the adult model on nonreinforced trials, the teacher immediately presented the picture to the participant and said "What's this?" If the participant answered correctly by saying "I don't know," the teacher delivered praise and a reinforcer. If the participant did not answer correctly (i.e., did not say "I don't know"), the teacher stated the rule, "Say 'I don't know' when she gets it wrong," and then re-presented the picture to the adult model. Immediately after the modeled response and consequence, the teacher asked the participant to label the picture again. If the participant answered correctly on this error-correction trial by saying "I don't know," general praise, but not the reinforcer, was provided. If the participant still did not answer correctly, the trial was terminated and the teacher moved on to the next trial. An FI 30-s schedule of token reinforcement was also provided for sitting quietly with hands down throughout each session.

Generalization. To determine the extent to which the discrimination that had been learned during the discrimination training sessions would occur in the presence of stimuli not associated with discrimination training, we conducted generalization sessions. Generalization sessions were conducted approximately every three to five training sessions with a different set of pictures. These sessions were procedurally the same as baseline conditions.

Additional training with generalization stimuli. Because Rose, Leah, and Mark did not demonstrate the discrimination with the generalization stimuli, discrimination training was implemented with these stimuli beginning on Session 26 for Rose and Session 58 for both Leah and Mark. The same discrimination training procedure described above was used.

Interobserver Agreement and Treatment Integrity

Interobserver agreement was calculated for responses during test sessions conducted after baseline, discrimination training, and generalization. An agreement was counted if both the teacher and a second observer independently scored a response as correct or incorrect in the same trial. Interobserver agreement was calculated on a trial-by-trial basis by dividing the number of agreements by the number of agreements plus disagreements and converting the result to a percentage. For Rose, interobserver agreement was collected for 50% of the test sessions following baseline and was 100% for all observed sessions, for 55% of the test sessions following discrimination training with a mean of 99% (range, 93% to 100%), and for 69% of the test sessions following generalization and was 100%. For Leah, agreement was collected for 40% of test sessions following baseline and was 100%, for 80% of the test sessions following discrimination training with a mean of 99% (range, 93 to 100%), and for 43% of the test sessions following generalization and was 100%. For Mary, agreement was collected for 54% of the test sessions following baseline, 63% of the test sessions following discrimination training, and 78% of the test sessions following generalization, and was always 100%. For Mark, agreement was collected for 77% of test sessions following baseline, 72% of test sessions following discrimination training, and 50% of test sessions following generalization, and was always 100%.

Treatment integrity data were collected on the accurate implementation of the independent variable during discrimination training, generalization, and test sessions. An independent observer used a checklist that contained all of the procedures described above (environmental arrangement, accurate delivery of feedback based on trial and session type, use of error correction, schedules of reinforcement, and presentation of stimuli during test sessions). For Rose, integrity data were collected for 53% of sessions with a mean of 99% (range, 96% to 100%). For Leah, integrity data were collected for 37% of sessions with a mean of 98% (range, 89% to 100%). For Mary, integrity data were collected for 49% of sessions with a mean of 97% (range, 87% to 100%). For Mark, integrity data were collected for 59% of sessions and was 100%.

RESULTS

Figure 1 displays the percentage of correct responses during test sessions conducted after baseline, discrimination training, and generalization sessions. Initially, for all participants, test sessions were conducted 10 min after baseline sessions. For all four participants, responding correctly on test sessions during baseline was at or below 60%. Because the first four test sessions in the discrimination training condition did not result in an increase in responding over baseline for Rose, we decided to conduct the test sessions 1 min after discrimination training sessions. So we would have comparable baselines, we

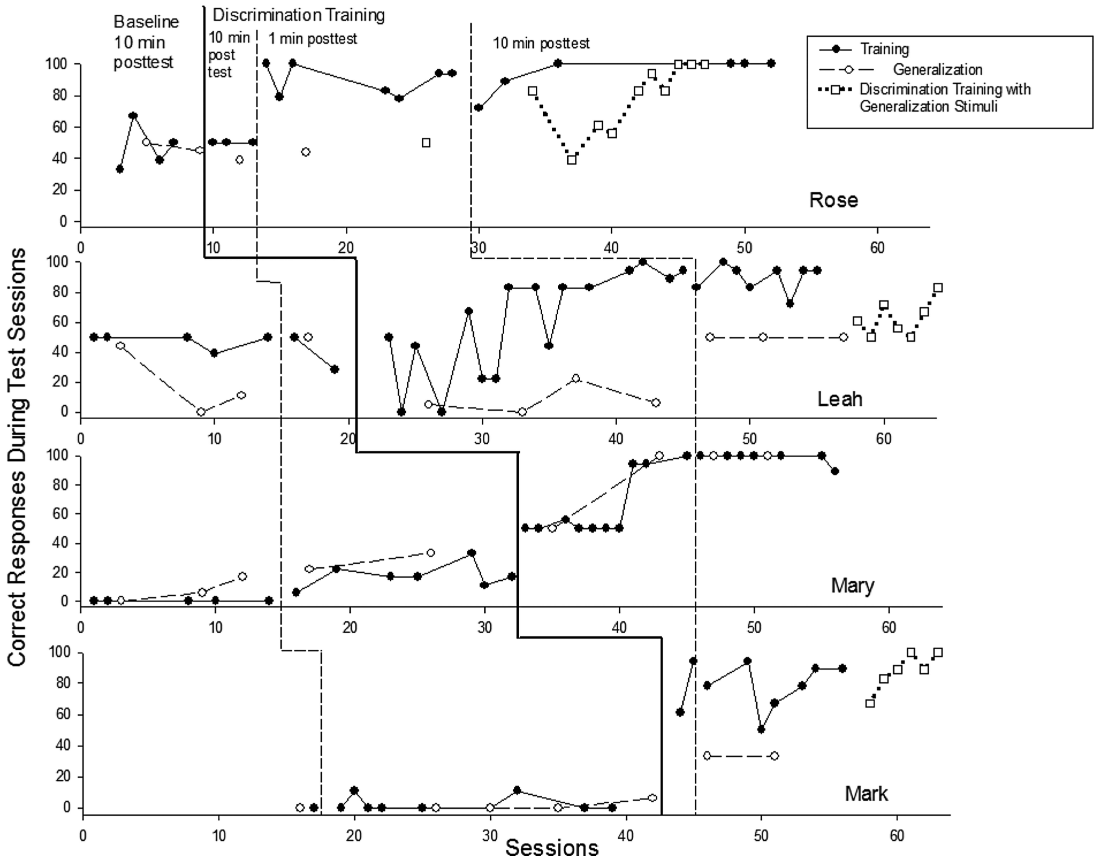

Figure 1. The percentage of correct responses during test sessions conducted after baseline, discrimination training, and generalization sessions for Rose, Leah, Mary, and Mark.

implemented the 1-min test sessions during baseline for the other participants as well. Only Mary showed a slight increase in correct responding during this 1-min baseline condition.

For Rose, correct responding increased to 100% during test sessions conducted 1 min after discrimination training. Correct responding decreased slightly but eventually increased to 100% again when the 10-min test session was reinstated. Correct responding during test sessions conducted 10 min after generalization sessions did not increase over baseline levels. Therefore, we implemented discrimination training during Session 26 of generalization, and responding during test sessions increased to 100% by Session 45. For Leah, correct responding increased to 100% during the test sessions conducted 1 min after discrimination training. Correct responding decreased slightly but eventually increased to 100% again when the 10-min test session was reinstated. Correct responding during test sessions conducted after generalization sessions did not increase over baseline levels for Leah either. Therefore, we implemented discrimination training during Session 58 of generalization, and responding increased to 94% by Session 67. For Mary, correct responding increased to 100% during test sessions conducted 1 min after discrimination training sessions and remained at 100% when the 10-min test was reinstated. A similar pattern of responding was observed for the test sessions conducted after generalization. For Mark, correct responding increased to 100% during test sessions conducted 1 min after discrimination training, dropped slightly when the 10-min test was reinstated, and remained at or above 80% for the remaining sessions. Correct responding during the test conducted after generalization sessions increased only slightly over baseline levels; for that reason, we implemented discrimination training for the generalization set of pictures on Session 58, and responding increased to 100% by Session 62.

Recall that during discrimination training, participants were asked to label the pictures immediately after the model's response. These data are not part of the experimental design; however, we did record responses to assess the extent to which participants were learning to make the discrimination within the training session. When modeled responses were reinforced, Rose responded with a correct picture label on 67% of the trials during the first session of discrimination training, and this responding remained high but slightly variable. When modeled responses were not reinforced, Rose responded correctly by saying "I don't know" on 100% of the trials across most of the sessions. Leah responded with a correct picture label on 78% of the trials during the first session of discrimination training. Responding on these trials was variable throughout discrimination training, ranging from 56% to 100%. When modeled responses were not reinforced, Leah responded correctly by saying "I don't know" on 44% of the trials during the first session of discrimination training. Similarly, responding on these trials was variable throughout discrimination training, ranging from 56% to 100%. Mary responded with a correct picture label on 89% of the trials when the responses of the model were reinforced during the first session of discrimination training. Responding correctly on these trials quickly increased to 100% and remained at this level for the remaining sessions. When modeled responses were not reinforced, Mary responded correctly on 89% of the trials during the first session of discrimination training, and correct responses quickly increased to 100% for the remaining sessions. Finally, for Mark, the percentage of trials during the first session of discrimination training in which he used a correct picture label and correctly said "I don't know" was 67% and 78%, respectively. For both trial types, correct responding increased to 100% with slight variability, ranging from 78% to 100% for both trial types.

DISCUSSION

Given the learning deficits of children with autism, it stands to reason that systematic instruction is needed to teach these children to discriminate among the contingencies of the responses applied to a model and to demonstrate this discrimination later. Similar to the Delgado and Greer (2009) study, we taught children with autism to discriminate reinforced responses from nonreinforced responses of others. Our study extended the Delgado and Greer study by teaching two practical responses that might be useful for future observational learning experiences: to imitate the responses of a model that were followed by positive reinforcement and to say "I don't know" in response to stimuli that were associated with modeled responses that were not reinforced. During discrimination training, all four participants learned to label unknown pictures by imitating the modeled responses that were reinforced by the teacher, and all learned to say "I don't know" in response to unknown pictures when the modeled responses were followed by negative feedback from the teacher. This discrimination was maintained 1 min and eventually 10 min after the training sessions for all participants.

One participant (Mary) demonstrated the discrimination during generalization tests with a different set of stimuli (i.e., different unknown labels) that were not associated with discrimination training. Three participants (Rose, Leah, and Mark) however, did not demonstrate the discrimination with the generalization stimuli until discrimination training was implemented with these stimuli. It may be the case that there were certain variables related to our instructional strategies that restricted generalization. For example, some of the stimuli present during discrimination training were not present during test sessions. Recall that the generalization stimuli (a different set of pictures) were never associated with discrimination training. The participants were never given the opportunity to imitate the reinforced responses and to say "I don't know" to the nonreinforced stimuli. In addition, the model and the consequences applied to the responses were not present in the test sessions. After we introduced these stimuli into discrimination training, however, all three participants learned to discriminate the contingencies applied to the model in test sessions 10 min after training.

Limited generalization of the discrimination to novel stimuli provides avenues for future research. First, in the present study, discrimination training was conducted with a total of only six stimuli: three reinforced and three nonreinforced responses. If more stimuli were used in the discrimination training sessions (e.g., 10 of each), generalization may have occurred. Second, we conducted training and generalization sessions separately. It might be the case that generalization could be enhanced by interspersing the generalization trials within the discrimination training sessions. Doing so may more closely associate the generalization stimuli with the discriminative stimuli used during discrimination training, leading to greater generalization. Third, to control for type of task, our study involved unknown picture labels to teach the observational learning responses. It is unclear if using different types of tasks within the discrimination training sessions (e.g., variety of academic tasks, a motor response, a vocal response, a social response) would have enhanced generalization. Fourth, for all participants, the total number of generalization sessions was disproportionately lower than the number of discrimination training sessions (i.e., approximately three discrimination training session to one generalization session). Increasing the number of generalization sessions to match the number of discrimination sessions, or interspersing generalization stimuli with testing stimuli, would provide more opportunities for the participant to observe the model and the consequences applied to the model's responses, potentially improving performance during generalization tests.

Another important factor was the timing of the test sessions. To date, there is no firm standard provided in the literature for when test sessions of observational learning for children with autism should take place. In fact, the research that has assessed observational learning using typically developing children implemented test sessions in which timing has varied from minutes to days (Meltzoff, Waismeyer, & Gopnik, 2012; Nadel et al., 2011). In our study, we chose 10 min for the test sessions based on prior research (Taylor, DeQuinzio, & Stine, 2012). However, despite being able to demonstrate the discrimination accurately during discrimination training, Rose was initially unable to do so during test sessions that occurred 10 min later. As a result, the time between discrimination training and test sessions was reduced to 1 min for Rose, and, in order to have comparable baselines, we changed the timing of test sessions in baseline for the remaining participants. For all participants, after performance was at 80% or above for the 1-min test sessions, the delay of the tests was changed to 10 min. Future studies could compare retention of these discriminated responses across varying intervals to identify a delay interval that is optimal for testing observational learning.

Ideally we would want participants to observe a model change a response when that response is followed by negative consequences. In our study, however, the model never changed the response based on teacher feedback on the nonreinforced trials. We did this to control for the arrangement of correct and incorrect responses. In other words, we scripted the responses of the teacher and the model according to trial type; each picture was always associated with a particular response and consequence. For example, the picture of kidney always produced the response "kidney" from the adult model and praise and a reinforcer from the teacher. Similarly, the picture of flash drive always produced the incorrect response "lumber," a negative statement (e.g., "No, that's wrong"), and no reinforcer. As a result of this arrangement, the participant never observed the model change her response based on the negative consequence from the teacher. It is unknown if doing so would have enhanced discrimination of the incorrect responses for the participant. Future studies could identify ways to control for trial type yet also have the model change the response as a result of the teacher's negative feedback. For example, the model could produce a different incorrect response when provided with negative feedback. That way, the participant would observe the model changing the response when provided with a negative consequence, which is arguably an essential component to observational learning (Want & Harris, 2001).

Mary's responses during baseline and test sessions may also inform future research. In baseline, she produced a specific response in the presence of each target picture, albeit incorrect. For example, she always said "hairclip" in the presence of a picture of a tweezers and "shirt" when presented with the picture of a tendon. These responses continued to be demonstrated during test sessions even though she was able to imitate correct responses and say "I don't know" accurately during discrimination training sessions. It could be the case that there were too many sessions of baseline in which she practiced these responses without feedback. Nevertheless, when procedures were adjusted to have the model use these same inaccurate labels that Mary used during baseline and test sessions, along with negative feedback and no reinforcer from the teacher, Mary changed her response in future test sessions, eventually reaching criteria. In future studies, participants' incorrect responses that are produced during baseline could potentially be used as the inaccurate label that is corrected by the teacher during discrimination training. In addition, a possible prerequisite for learning this discrimination is the ability to use the phrase "I don't know" when presented with unknown labels.

Finally, we did not use preference or reinforcer assessments to identify the items presented to the adult model as reinforcers for correct responses or to identify the items presented to the participants contingent on correct responding during discrimination training. Instead, we relied on teacher report to identify those items earned most frequently in the classroom. It may be argued that using items formally identified as reinforcers could have enhanced the effects of discrimination training; however, more research is required to fully understand the effects of using established reinforcers on observational learning.

In this study, we successfully taught children with autism to discriminate the consequences applied to the behavior of a model. This discrimination is perhaps the most complex component of observational learning, in that it requires the learner to sustain attention for long periods of time, to respond differentially to complex stimuli, and to demonstrate the discrimination later. It is difficult to train and test observational learning repertoires in a way that approximates the complexity of observational learning as it occurs in the natural environment. This study adds to the emerging body of research demonstrating that some children with autism can learn responses associated with observational learning and offers ideas for future research to refine procedures to improve the observational learning skills of children with autism.

REFERENCES

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology*, *1*, 589–595. doi: 10.1037/h0022070

Bandura, A., & McDonald, F. J. (1963). Influence of social reinforcement and the behavior of models in shaping children's moral judgment. *Journal of Abnormal and Social Psychology*, *67*, 274–281. doi: 10.1037/h0044714

Bandura, A., Ross, D., & Ross, S. A. (1963). Vicarious reinforcement and imitative learning. *Journal of Abnormal and Social Psychology*, *67*, 601–607. doi: 10.1037/h0045550

Catania, A. C. (1998). *Learning* (interim 4th ed.). Cornwall-on-Hudson, NY: Sloan.

Charlop, M. H., Schreibman, L., & Tryon, A. S. (1983). Learning through observation: The effects of peer modeling on acquisition and generalization in autistic children. *Journal of Abnormal Child Psychology*, *11*, 355–366. doi: 10.1007/BF00914244

Christensen, A. M., Lignugaris-Kraft, B., & Fiechtl, B. (1996). Teaching pairs of preschoolers with disabilities to seek adult assistance in response to simulated injuries: Acquisition and promotion of observational learning. *Education &Treatment of Children*, *19*, 3–18.

Deguchi, H. (1984). Observational learning from a radical-behavioristic viewpoint. *The Behavior Analyst*, *7*, 83–95.

Delgado, J. A. P., & Greer, R. D. (2009). The effects of peer monitoring training on the emergence of the capability to learn from observing instruction received by peers. *The Psychological Record*, *59*, 407–434.

Donnelly, J. L., Luyben, P. D., & Zan, C. S. (2009). Increasing eye contact toward learning materials in a toddler with autism. *Journal of Prevention & Intervention in the Community*, *37*, 170–176. doi: 10.1080/10852350902975926

Egel, A. L., Richman, G. S., & Koegel, R. L. (1981). Normal peer models and autistic children's learning. *Journal of Applied Behavior Analysis*, *14*, 3–12. doi: 10.1901/jaba.1981.14-3

Fryling, M. J., Johnston, C., & Hayes, L. J. (2011). Understanding observational learning: An interbehavioral approach. *The Analysis of Verbal Behavior*, *27*, 191–203.

Greer, R. D., Dudek-Singer, J., & Gautreaux, G. (2006). Observational learning. *International Journal of Psychology*, *41*, 486–499. doi: 10.1080/00207590500492435

Ledford, J. R., Gast, D. L., Luscre, D., & Ayres, K. M. (2008). Observational and incidental learning by children with autism during small group instruction. *J. ournal of Autism and Developmental Disorders*, *38*, 86–103. doi: 10.1007/s10803-007-0363-7

Leekam, S. R., Hunnisett, E., & Moore, C. (1998). Targets and cues: Gaze-following in children with autism. *Journal of Child Psychology and Psychiatry*, *39*, 951–962. doi: 10.1111/1469-7610.00398

Lovaas, O. I., Koegel, R., Simmons, J. Q., & Long, J. S. (1973). Some generalization and follow-up measures on autistic children in behavior therapy. *Journal of Applied Behavior Analysis*, *6*, 131–166. doi: 10.1901/jaba.1973.6-131

Masia, C. L., & Chase, P. N. (1997). Vicarious learning revisited: A contemporary behavior analytic interpretation. *Journal of Behavior Therapy and Experimental Psychiatry*, *28*, 41–51. doi: 10.1016/S0005-7916(96)00042-0

Meltzoff, A. N., Waismeyer, A., & Gopnik, A. (2012). Learning about causes from people: Observational causal learning in 24-month-old infants. *Developmental Psychology*, *48*, 1215–1228. doi: 10.1037/a0027440

Nadel, J., Aouka, N., Coulon, N., Gras-Vincendon, A., Canet, P., Fagard, J., & Bursztejn, C. (2011). Yes they can! An approach to observational learning in low-functioning children with autism. *Autism*, *15*, 421–435. doi: 10.1177/1362361310386508

Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced discrimination of novel, highly similar stimuli by adults with autism during a perceptual learning task. *Journal of Child Psychology and Psychiatry*, *39*, 765–775. doi: 10.1111/1469-7610.00375

Rehfeldt, R. A., Latimore, D., & Stromer, R. (2003). Observational learning and the formation of classes of reading skills by individuals with autism and other developmental disabilities. *Research in Developmental Disabilities*, *24*, 333–358. doi: 10.1016/S0891-4222(03)00059-3

Schreibman, L. (1975). Effects of within-stimulus and extra-stimulus prompting on discrimination learning in autistic children. *Journal of Applied Behavior Analysis*, *8*, 91–112. doi: 10.1901/jaba.1975.8-91

Taylor, B. A., & DeQuinzio, J. A. (2012). Observational learning and children with autism. *Behavior Modification*, *36*, 341–360. doi: 10.1177/0145445512443981

Taylor, B. A., DeQuinzio, J. A., & Stine, J. (2012). Increasing observational learning of children with autism: A preliminary analysis. *Journal of Applied Behavior Analysis*, *45*, 815–820. doi: 10.1901/jaba.2012.45-815

Varni, J. W., Lovaas, O. I., Koegel, R. L., & Everett, N. L. (1979). An analysis of observational learning in autistic and normal children. *Journal of Abnormal Child Psychology*, *7*, 31–43. doi: 10.1007/BF00924508

Want, S. C., & Harris, P. L. (2001). Learning from other people's mistakes: Causal understanding in learning to use a tool. *Child Development*, *72*, 431–443. doi: 10.1111/1467-8624.00288

Werts, M. G., Caldwell, N. K., & Wolery, M. (1996). Peer modeling of response chains: Observational learning by students with disabilities. *Journal of Applied Behavior Analysis*, *29*, 53–66. doi: 10.1901/jaba.1996.29-53

Williams, J. H. G., Whiten, A., & Singh, T. (2004). A systematic review of action imitation in autistic spectrum disorder. *Journal of Autism and Developmental Disorders*, *34*, 285–299. doi: 10.1023/b:jadd.0000029551.56735.3a

Wolery, M., Ault, M. J., Gast, D. L., Doyle, P. M., & Griffen, A. K. (1991). Teaching chained tasks in dyads: Acquisition of target and observational behaviors. *Journal of Special Education*, *25*, 198–220. doi: 10.1177/002246699102500205

Received November 6, 2013
Final acceptance July 30, 2014
Action Editor, Thomas Higbee