Journal of
Applied Behavior Analysis

TEACHING OBSERVATIONAL LEARNING TO CHILDREN WITH AUTISM

JACQUELYN MACDONALD AND WILLIAM H. AHEARN

WESTERN NEW ENGLAND UNIVERSITY, THE NEW ENGLAND CENTER FOR CHILDREN

Observational learning (OL) is critical for the acquisition of social skills and may be an important skill for learning in traditional educational settings. Although OL occurs during early childhood in the typically developing population, research suggests that it may be limited in children diagnosed with autism spectrum disorder (ASD). The purpose of the present study was to develop an assessment to test for the presence of OL across a variety of tasks. If OL was deficient, we sought to teach it by training specific skills. Six participants who had been diagnosed with ASD demonstrated deficits in OL. After an initial assessment, a multiple-probe design across OL tasks showed that training produced acquisition of these skills across multiple exemplars. After training, 5 of the 6 participants engaged in OL across multiple tasks and task variations, demonstrating generalization. For 1 participant, generalization of performance did not occur across tasks but did occur within task variations.

Key words: autism, generalization, observational learning

Learning by watching others is an important way by which people learn about their environment (Miller & Dollard, 1941; Taylor & DeQuinzio, 2012; Weiss & Harris, 2001). Varni, Lovaas, Koegel, and Everett (1979) suggested that children with autism spectrum disorder (ASD) often do not learn via observation. Researchers have speculated that certain skills must be present within the repertoire of the individual for observational learning (OL) to occur (Browder, Schoen, & Lentz, 1986; Greer, Dudek-Singer, & Gautreaux, 2006; Taylor & DeQuinzio, 2012; Weiss & Harris, 2001). Moreover, it is likely that children with ASD present with significant deficits in at least some of the skills necessary to engage in OL. These include attending (Patten & Watson, 2011), imitation (Williams, Whiten, & Singh, 2004), delayed imitation (Garcia, 1976), and consequence discrimination (Pereira Delgado & Greer, 2009). Although OL is critical to social and academic development, there is little behavior-analytic research on how to teach this skill.

Historically, there has been some confusion regarding the terms *observational learning* and *generalized imitation*. Generalized imitation occurs when some imitative responses occur following a model of a response without a previous history of that response (Birnbrauer, 1976). An example of generalized imitation includes an observer engaging in a novel response (e.g., clapping his or her hands at the end of a song during a recital) after observation of the model clapping hands in that situation. This behavior presumably occurs, in part, due to the history of reinforcement for emitting behavior similar to that of the model. These skills are also part of OL; however, OL also involves discrimination of contingencies for the model's responding. Therefore, *observational learning* is defined as differential responding based on the observed response and its corresponding consequence. At

This paper was submitted to the Department of Psychology in the School of Arts and Sciences at Western New England University by the first author under the supervision of the second author in partial fulfillment of the requirements for the doctoral degree in behavior analysis. We thank Jason Bourret, Rachel Thompson, and Amanda Karsten for their feedback on earlier versions of this manuscript.

Correspondence concerning this paper should be addressed to Jacquelyn MacDonald, The New England Center for Children, 33 Turnpike Road, Southborough, Massachusetts 01772 (e-mail: jmacdonald@necc.org).

doi: 10.1002/jaba.257

the onset of observation, the observer attends to the social aspects of the situation, including the model, and any discriminative stimuli that might effectively occasion the behavior of the observer. Some form of the behavior must also already be present within the repertoire of the individual (i.e., he or she must be capable of engaging in the modeled response) to replicate the modeled act. Finally, the observer attends to the consequences that follow the modeled action and then responds differentially based on the action and the consequence. An example of OL would be when a student does not correctly point to a written word in the presence of a picture associated with the written word. The student observes another student receive instruction by a teacher followed by consequences (in the form of positive or corrective feedback). Then, because of observing this performance and the consequences that follow, the student correctly points to the word associated with the picture in the absence of direct instruction.

Observational learning requires the participant to engage in several responses during and after the modeled action. In a review of the literature, Taylor and DeQuinzio (2012) suggested that children with ASD have deficits in many of these responses. They posited that teaching sustained attention to peer models, generalized imitation of a peer's vocal and motor responses, and discrimination of consequences could facilitate the emergence of OL in the natural environment. To date, however, no method exists for assessing the presence of OL and the skills necessary to engage in OL, although three studies have specifically examined teaching OL.

Researchers have tried to teach OL through peer-yoked contingencies, peer monitoring, and teaching a differential observing response. Each study focused on teaching attending to discriminative stimuli, consequences related to the OL task, or both. Taylor, DeQuinzio, and Stine (2012) assessed the effects of teaching a monitoring response on acquisition of differentially responding to words. During the training condition, participants were taught to imitate vocally the word read by the peer and to match the vocal response to the written word on the matching board. During the exposure condition, participants only observed the peer's reading response. Test conditions occurred 10 min after each condition to assess acquisition of the previously observed reading words. Results suggested that all participants acquired expressive identification of words (i.e., tacts) with and without monitoring. Some limitations of this study were that participants did not have the opportunity to observe differential consequences with a subsequent opportunity to respond, the absence of the assessment of OL before training, and the absence of the assessment of generalization of the established repertoire.

Rothstein and Gautreaux (2007) also sought to teach academic OL repertoires to three students with emotional and behavioral disorders using a peer-yoked contingency. Participants were taught to respond correctly with a set of stimuli (e.g., names of trees, presidents, musical instruments) before they observed another student learn to respond correctly in the presence of a different set of stimuli. After training, higher levels of correct responding occurred; however, when the participants were presented with novel stimuli, correct responding was not reliably demonstrated. Observation of differential consequences for responding was also not included.

Pereira Delgado and Greer (2009) investigated whether OL emerged after the monitoring of correct or incorrect responses emitted by peers under conditions that included observation of consequences following modeled responses. Students observed their peers vocally respond to a written word and then indicated whether the peer response was correct or incorrect by choosing a colored block. The experimenter provided reinforcement to the observer for correct block choice (i.e., choosing the block that represented the target student's response). Feedback was provided for incorrect monitoring.

Although responding increased from baseline to posttreatment, these results do not strongly demonstrate that acquisition of OL was solely the result of the peer-monitoring procedure. Multiple-probe sessions were conducted after the peer-monitoring procedure, so it is possible that repeated practice during probe sessions produced correct responding. Probes of novel OL tasks would have helped to clarify whether OL was established.

The purpose of the present study was to extend the literature by addressing these limitations. First, our aim was to develop an assessment to test for the presence of OL across a variety of tasks including academic and leisure activities. When the performance of OL was deficient, we taught specific skills (i.e., attending, imitation, delayed imitation, consequence discrimination, as suggested by Taylor & DeQuinzio, 2012) to produce that performance. Then, generalization of OL was assessed following training on different tasks, variations on that type of task, or both. Our end goal was to produce generalized performance of OL using multiple exemplars (i.e., a variety of tasks and variations on each type of task).

METHOD

Participants

Six participants, who received educational and clinical services at a private school for children with ASD and related disorders, were included in this study. They ranged in age from 8 to 21 years old. All participants engaged in a point response before the start of the experiment but did not reliably engage in an imitative repertoire. Specifically, all participants imitated simple actions but did not imitate actions with objects. George was a 16-year-old boy who had been diagnosed with pervasive developmental disorder and mitochondrial cell disorder. He communicated primarily with an iPad loaded with TouchChat programming. He used simple sentences to request preferred items and activities. He identified familiar objects and pictures receptively but had a limited expressive vocabulary (i.e., five words). George did not receive a score on the Peabody Picture Vocabulary Test (4th ed., Form B; PPVT-4B) due to repeated responding errors in the first set of the assessment. Fred was an 8-year-old boy who had been diagnosed with ASD. He communicated primarily through vocal speech and natural gestures (e.g., pointing and waving). He could identify familiar objects and all of the letters of the alphabet both expressively and receptively. Fred received a score of 2 years 7 months on the PPVT-4B. Emma was a 14-year-old girl who had been diagnosed with ASD. She communicated primarily using an iPod touch with TouchChat programming. In addition, she used vocal approximations (one- to two-word phrases) to express needed or desired items or activities, but these phrases could be difficult to understand by unfamiliar persons. Emma could receptively identify a variety of familiar objects as well as one-step directions. She did not receive a score on the PPVT-4B due to repeated errors on the first part of the assessment. Teddy was a 10-year-old boy who had been diagnosed with ASD. He communicated primarily through vocal speech using one- to two-word utterances and natural gestures (e.g., pointing and waving). He could identify a variety of familiar objects expressively and receptively as well as follow many two-step directions. He did not receive a score on the PPVT-4B due to repeated errors on the first part of the assessment. Sally was a 21-year-old woman who had been diagnosed with ASD. She was vocally verbal and communicated using short sentences and phrases to make requests, to indicate acceptance and rejection, to make comments, and to engage socially with others. Her pronunciation was generally good; however, she could sometimes be difficult to understand due to a low volume, varied intonation, and word substitutions and omissions. She could identify common items and pictures in her environment receptively and expressively. She received a score of 6 years 8 months on the PPVT-4B. Brad was a 19-year-old man who had been diagnosed with a pervasive developmental disorder. He communicated vocally to request, label, and comment on people, actions, and objects in short phrases. He could identify many common objects and pictures receptively and expressively. He received a score of 5 years 7 months on the PPVT-4B. All participants had received a diagnosis before entry into this study by a provider not affiliated with the researchers or the facility that provided clinical services.

Setting and Materials

We conducted all sessions in the participant's academic workspace or in a therapy room at the participant's school. Participation in the study took place for 20 to 30 min per day, three to five times per week. Materials included one table and two chairs, a video camera, and specific materials needed for each condition.

Materials included the Peabody Picture Vocabulary Test (Form 4B; PPVT) assessment booklet (Dunn & Dunn, 2007), previously identified highly preferred reinforcers, and various materials to assess OL across five tasks (see Table 1). Within the five tasks, we had materials for three different variations of those tasks. For example, in the hidden items task, items were hidden in either cloth green boxes, cloth blue boxes, or plastic cups. In the computer games task, the variation included the position of the button that produced reinforcement in the form of a preferred video. A forced choice preference assessment (Fisher et al., 1992) for edible items, toys, and video clips was conducted before the start of the study to ensure adequate reinforcers and preferred activities necessary for various aspects of the study. Data are available from the first author.

Description of Observational Learning Tasks

We deemed it important that the assessment of OL occur across a variety of leisure and academic tasks because previous research assessed OL across only one task with no novel variations of that task or different tasks (e.g., Pereira Delgado & Greer, 2009). Several tasks were selected from seminal articles related to imitation, video modeling, and observational learning and were modified for this study. We did not assess preference for task materials before the study.

Hidden item task. This task was similar to one described in Miller and Dollard (1941) involving a preferred object hidden in one of three possible places. Hidden items were preferred as determined via a preference assessment. The neutral item was selected as something in the participant's environment that was not identified as preferred during assessments. We placed a preferred item in one box or cup, a neutral item in another box or cup, and placed nothing in the final box or cup. The experimenter said, "Watch me as I choose from the different places." We hid items in three cups or cloth boxes. The experimenter opened the box or cup on the left, pulled out the preferred item, and said "Yum, this is really good." Next, the experimenter opened the box or cup in the middle, pulled out the neutral edible item, and said, "Yuck, this is gross," and grimaced. Finally, the experimenter opened up the box or cup on the right and said, "There is nothing inside, oh well." After each observation, the experimenter asked the participant to choose one box. During subsequent trials, we placed each item in a different box or cup so the participant attended to the actions of the experimenter as well as to the location of the preferred item. One session consisted of six trials so that each box was systematically in each position twice. Task variations included finding the preferred item in green boxes, blue boxes, and plastic cups.

Computer task. This task was similar to procedures described by MacDonald, Dickson, Martineau, and Ahearn (2015), which assessed learning following observation of an experimenter playing a computer game. The participant and the experimenter sat next to each other

Table 1
Task Variations for Observational Learning Tasks

OL tasks	Possible variations
Hidden items	Green box, blue box, plastic cups
Computer game	Right, left, or top correct position
Academic	Various nonsense symbols or words
Construction toy	Dump truck, front loader, back hoe
Building toy	Elmo, Cookie Monster, Lego, Thomas

in front of the computer. The experimenter said, "Watch me as I play this computer game," pushed the button on the right, and said "This button is broken. I am going to try the other button." The experimenter pushed the button on the left and a preferred movie clip appeared on the screen. The experimenter exclaimed "Cool! Let's watch." When the movie clip finished, the experimenter asked the participant to push a button. Correct button locations were counterbalanced across four trials (one session) so that each button could be on either side twice. After each observation trial, the experimenter gave the participant the opportunity to choose one button to push. Task variations included the correct button on the right and the incorrect button on the left, the correct button on the left and the incorrect button on the right, and the correct button on the top and the incorrect button on the bottom.

Academic task. This task was similar to procedures described by Schoen and Ogden (1995) and Orelove (1982), which assessed OL in the context of an academic task. Both the confederate and participant sat at the desk with the experimenter on the opposite side of the table. The experimenter told the participant to watch the confederate. Over six trials, the experimenter asked the confederate to pick the corresponding stimulus out of an array of three stimuli when presented with a sample. For instance, when told, "pick zing," the confederate pointed to the correct symbol and the experimenter delivered a preferred edible item. If the confederate pointed to an incorrect symbol, the experimenter said "no, that's wrong." We counterbalanced the confederate's correct and incorrect answers across trials, giving the participant exposure to correct and incorrect responding. After completion of six trials, the experimenter said "[Participant], it is your turn, pick zing." Correct responses were followed by reinforcement, whereas incorrect responses were followed by the verbal feedback, "no, that's wrong." We completed six trials (one session) following observation of the confederate engaging in responding so that each symbol was in each spot twice. Variations included different nonsense pictures when given nonsense words.

Construction toy. This task assessed OL in the context of a play response similar to those included in Varni, Lovaas, Koegel, and Everett (1979). Two similar toys were present, but one toy was broken. We counterbalanced the toys across four trials so that each toy could be in each spot twice. The experimenter started each session by saying "Watch me play with these toys." The participant observed the experimenter pull back each toy, push a button, and observe the effects (i.e., boulder falls out, lights flash) or no effects (i.e., boulder does not fall out, lights do not flash). While playing with each toy, the experimenter commented about the toys working or being broken. Then the experimenter asked the participant to choose a toy. For example, the experimenter and participant sat at a table with two identical dump trucks. Both dump trucks had a boulder in the back, but one boulder was glued into the vehicle. The experimenter said "Watch me play with these toys." The experimenter backed up the dump truck across the table, and pulled up on the back of the dump truck. Depending on the contents, the experimenter commented on the action with an excited tone if the contents fell out, or with a frustrated tone if the contents did not fall out. The experimenter played with the other truck in an identical manner. Next, the experimenter asked the participant to pick one truck to dump. After each observation of the experimenter engaging with both toys, the experimenter gave the participant the opportunity to engage with either toy. Task variations included a dump truck, a front loader, and a backhoe.

Building toys. This task was similar to procedures described by Werts, Caldwell, and Wolery (1996), which assessed OL in the context of building a character toy when a toy with all of the pieces and a toy with missing pieces were available. Each toy consisted of a four-step building sequence, and all toys were assessed to be of similar difficulty. The experimenter placed identical toy pieces (with the exception of one missing piece) in baggies. We counterbalanced the position of the baggies across four trials (one session) so that each baggie was placed in each spot twice. For instance, at the start of each session, the experimenter said, "Watch me as I build a toy!" The experimenter built both toys and commented on contents within each bag. After completion of each toy construction, the experimenter said, "You pick one toy to build." After each observation of engagement with both toys, the experimenter gave the participant the opportunity to engage with either baggie with toy pieces and to build a toy. Task variations included an Elmo toy, Cookie Monster toy, Thomas the Tank toy, and Lego toys, depending on the age of the student.

Design

We used a multiple-probe design across participants to demonstrate functional control of training over responding. We used a multiple probe within participants (across OL task variations) to gauge functional control of training over responding and to assess generalization. We assessed and trained only one task per day. The experimenter did not conduct probes for acquisition of OL on the same day as training sessions; we conducted probes on the following day. Probes occurred both before training started and when all skills targeted for increase during training met the mastery criterion (one session with 100% accurate and independent responding).

Dependent and Independent Variables

The dependent variable in the preassessment and postassessment phases was the percentage of correct, independent performances for each task variation across OL tasks. The preassessment phase served as a baseline measure to demonstrate how each participant performed before training, and the postassessment phase served as a control measure to demonstrate how participants performed after training and, in addition, to assess for generalization across OL task variations and OL tasks.

The dependent variable in the training phase was the percentage of correct, independent responding across a nine-trial block (one session). The independent variable was the use of the least-to-most physical prompting to teach each skill. After mastery of one skill, the next skill was taught (i.e., when imitation was considered mastered, delayed imitation was taught).

Data Collection and Interobserver Agreement

During the pre- and postassessments, we collected data on the percentage of correct, independent performance on each task variation following observation. For a task variation to be considered mastered, the participant chose the correct task materials necessary for the delivery of the reinforcing consequence and independently and accurately completed the task (i.e., chose the baggie with all of the pieces and built the toy). We assessed only one OL task (across three task variations) per day. During the training phase, we collected data on the percentage of correct, independent responding on each skill. A skill was deemed mastered after 100% accurate and independent responding across nine consecutive trials in one session. A second observer watched each session on videotape. Exact agreement was calculated (i.e., chose correct toy and imitated action) by the first observer and the second observer. 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. Agreement was calculated during 99% of assessment sessions across all participants and was at or above 98% (range, 98% to 100%).

An independent observer reviewed the data sheet and a checklist that contained all of the procedures described below to assess procedural integrity (e.g., reinforcer delivery, environmental arrangement, and task delivery). The independent observer scored correct and incorrect responses of the experimenter on the above skills across all phases. The percentage of skills correctly implemented was calculated on a trial-by-trial basis by dividing the number of skills correctly implemented by the total number of skills and converting the result to a percentage. Procedural integrity was calculated during 99% of all sessions across all participants and was 100%.

General Procedure

We chose tasks novel to the participant to assess OL. The experiment consisted of three phases; preassessment, training, and postassessment (see Figure 1). The preassessment phase provided the opportunity for the participant to observe the completion of each task under both reinforcing consequences and neutral consequences. After preassessment, we arbitrarily selected one OL task to be taught. Training included teaching attending to the model, imitation of the task, delayed imitation, and consequence discrimination using the materials from the specific OL task variation. After training, we implemented a postassessment phase to assess skill acquisition and generalization.

Preassessment phase. Before the start of a session, the experimenter said "watch me" and engaged with items required for that session. During the preassessment phase, the experimenter gave the participant the opportunity to observe the experimenter engage in responses followed by either reinforcing or neutral consequences for each OL task and task variation. We did not deliver programmed reinforcement or prompts. The experimenter demonstrated each task four to six times. The number of demonstrations depended on the number of stimuli in each task. For instance, in the academic task, six trials were demonstrated to allow the participant to see each stimulus followed by a positive consequence and a negative or neutral consequence. The participant observed the experimenter engage in the task and, after the 2-min interval, the experimenter said, "It's your turn," and provided the opportunity to engage in the response. After observation, the experimenter waited 2 min before allowing the participant the opportunity to engage in the task. The rationale behind the 2-min interval was to make sure there was a delay between the opportunity to observe and the opportunity to engage with the materials. We assessed participant performance on all OL tasks and task variations during this phase. After the assessment phase, we selected one OL task variation for initial exposure to training.

Training phase. We taught skills in a nine-trial data block. Multiple blocks were conducted in the same day. We used least-to-most physical prompting to facilitate correct responding. The experimenter used prompting to facilitate independent prompting by providing the least restrictive prompt before a more restrictive prompt. Verbal praise paired with a preferred edible item or activity (based on a forced-choice preference assessment) followed correct independent or prompted responding. We ignored incorrect responding and presented another trial. A skill was deemed mastered after 100% accurate and independent responding across nine consecutive trials in one session. After mastery of each skill, training began with the next skill.

The first skill taught was attending. Attending was considered independent if, after the experimenter's statement ("watch me"), the participant oriented his or her face toward the experimenter and provided brief eye contact. Materials were present on the table, but the participant was not required to interact with them. When prompting attending using a full physical guidance prompt,

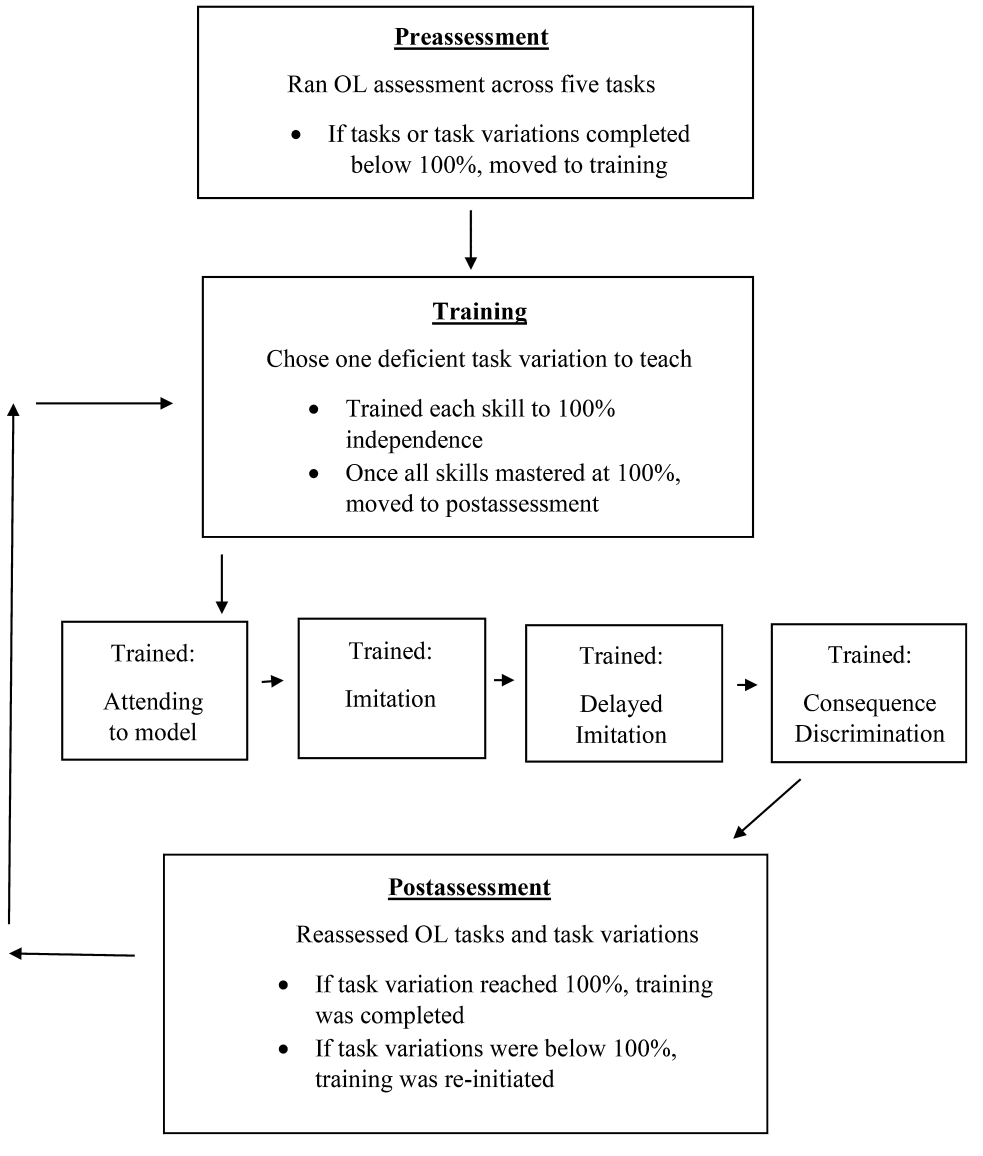

Figure 1. Flow chart of general procedures.

the experimenter provided a hand shield to block the participant's gaze in directions other than toward the experimenter. To increase the probability of eye contact, the experimenter placed a preferred edible item next to her eyes for all prompting except the gesture prompt.

The next skill taught was imitation. Imitation consisted of repeating the action of the experimenter following the statement, "watch me," paired with an action. Imitation training ensured that participants engaged with the item correctly. Only the materials necessary for the delivery of the reinforcing consequence were present. During imitation training, the experimenter said, "watch me," and completed the specific OL task variation. Then the experimenter said,"It's your turn," and put the task materials in front of the participant.

The third skill taught was delayed imitation. Delayed imitation consisted of repeating the action of the experimenter following a 5-s delay. This ensured that the participant learned to imitate the actions of the experimenter after a delay. Only the materials necessary for the delivery of the reinforcing consequence were present. During delayed imitation training, the experimenter said, "watch me," and completed the specific OL task variation. The experimenter waited 5 s, then said, "It's your turn," and put the task materials in front of the participant.

The final skill taught was consequence discrimination. Training was similar to the assessment phase with the addition of prompting and reinforcement for choosing the materials necessary for the delivery of the reinforcing consequence and refraining from modeled behavior that had been followed by no reinforcement. Materials necessary for the delivery of the reinforcing consequence and the neutral consequence were present. During consequence discrimination training, the experimenter said, "watch me," and completed the specific OL task variation with all task materials (similar to the assessment phase). Then the experimenter said, "It's your turn." Consequence discrimination was considered mastered if the participant chose the item that had previously been followed by positive consequences in addition to completing the action independently and refraining from observed behavior followed by no reinforcement. After nine trials of correct independent responding, the postassessment phase occurred.

Postassessment phase. We completed the postassessment phase on the day after completion of the training phase. Procedures were identical to the preassessment phase. If performance with all task variations in the specifically taught OL task was 100% accurate, we probed all OL tasks and task variations to assess generalization. If performance on any OL task variations did not increase to 100% correct responding, we introduced the training phase for those tasks. We taught OL tasks and task variations until the participant acquired all tasks and task variations after training or generalization across tasks (see Figure 1).

RESULTS

Preassessment

None of the participants performed all OL tasks independently (Figure 2). After observing the experimenter engaging in the tasks during the assessment phase, George correctly performed the hidden items task and the computer task. Fred and Teddy did not perform any of the tasks correctly after observation. Emma and Sally correctly performed the academic task, and Brad correctly performed the computer task.

Training

The amount of training needed for each participant to acquire all OL skills in the assessment ranged from training only one variation of one of the tasks (i.e., Fred) to training OL with task variations for three tasks (i.e., George and Teddy). Some generalization to untrained tasks occurred during posttests with all participants. Data from each training session are available from the first author. Table 2 indicates the number of training sessions needed for each participant across each skill. More training was needed for mastery of imitation for Teddy and Fred, whereas George required more training for delayed imitation. Emma, Sally, and Brad required the most training on consequence discrimination. Fred performed all task variations correctly after training on just one task variation; however, all other participants required training on at least two tasks before they performed all tasks correctly.

Postassessment

For each participant, each leg in Figures 3 through 5 depicts a different task. It is important to note that skills performed correctly during the assessment phase were not included. Figure 3

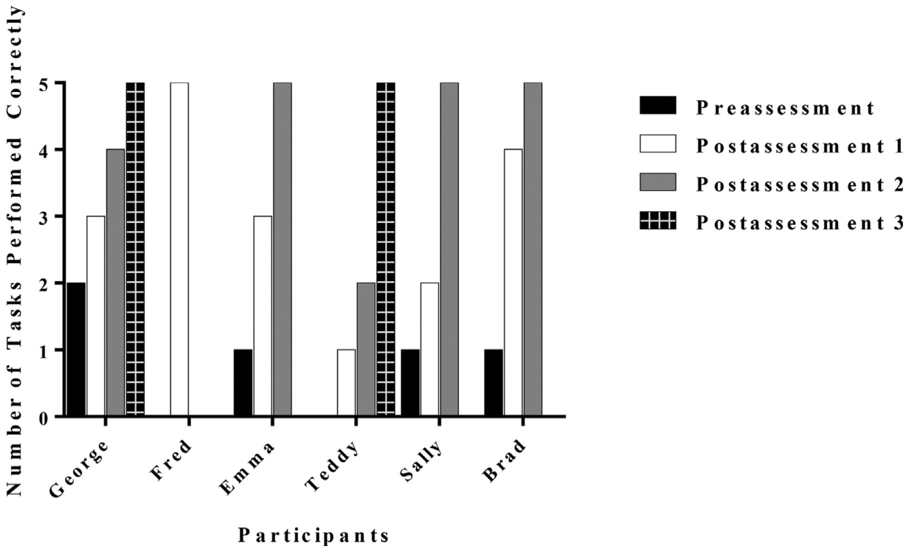

Figure 2. Results of preassessment and postassessment of OL tasks across participants. An OL task was not considered to be performed correctly until all variations of that task were performed correctly. Postassessments 2 and 3 occurred after all variations of the taught OL task were performed correctly.

(left) illustrates the results across all tasks and task variations for George. Performance was low across all task variations during the preassessment. After training for each task, George demonstrated skill acquisition in the absence of training for the two untrained task variations for each of the three tasks. However, generalization was not demonstrated across tasks. For Fred (Figure 3, right), generalization occurred across all other tasks and variations following training with only one task variation of one task. Figure 4 (left) depicts the results across all tasks for Emma. During the preassessment phase, performance was low across all tasks. After training for two construction task variations and one computer game task variation, Emma demonstrated generalization for all remaining tasks and task variations.

Figure 4 (right) illustrates the results across all tasks for Teddy. During preassessment, performance was low across all task variations. After training on each task, Teddy demonstrated skill acquisition for all variations following training for two variations of the building task, three variations for the construction task, and one variation in the academic task. Generalization was demonstrated across one task variation in the building task, two task variations in the academic task, and across all variations in the computer game task and the hidden items task.

Figure 5 depicts the results for Sally (left) and Brad (right). Performance was low across all task variations during the preassessment for Sally. After training for one variation of the hidden items task, generalization occurred across the remaining two variations. After training for one variation for the building task, generalization occurred across the remaining two variations. In addition, after training of the previous variations, generalization of performance was demonstrated across all variations of the construction task and the computer game task. Performance was low across all task variations for Brad in the preassessment phase. After training for one variation of the construction task, generalization occurred across the remaining two variations. After training for one variation for the academic task, generalization was demonstrated across the

Table 2
Total Training Sessions Required for Mastery of Each Skill

Participant	Attending	Imitation	Delayed imitation	Consequence discrimination
George	3	9	14	7
Fred	2	7	1	3
Emma	3	9	3	11
Teddy	8	40	13	10
Sally	2	5	2	8
Brad	2	4	2	5

remaining two variations. In addition, after training of the previous variations, generalization of performance was demonstrated across all variations of the building and hidden items task.

Figure 2 depicts the results of the preassessment and postassessment for each participant. It is important to note that a task was not considered to be performed correctly until each variation of that task was performed correctly. George acquired all tasks after training on one task variation for each task but did not demonstrate generalization across tasks. Fred acquired correct OL performance on all tasks after training on one task variation for one task; thus, performance generalized across all task types. Emma acquired all tasks after training on two task variations and demonstrated generalization across two tasks. Teddy acquired all tasks following training on six task variations and demonstrated generalization across two tasks. Sally and Brad acquired all tasks after training on two task variations and demonstrated generalization across two tasks.

DISCUSSION

This study demonstrated that OL can be taught to children with ASD across a variety of leisure and academic tasks. Specifically, performance of OL on untrained tasks emerged for five of the six participants after training of a specific task (George did not display generalization across tasks). This study is the first to assess a variety of OL tasks, in that previous studies primarily focused on a specific task. Our initial assessment was effective in identifying whether participants engaged in OL before training. Four of the six participants demonstrated OL in at least one task during the preassessment. This may have been due to a previous history of reinforcement with similar tasks. For instance, Emma demonstrated OL during the academic tasks but not during leisure activities. It is possible that she had previously learned to observe others during discrete-trial instruction in a small-group format but did not demonstrate generalized OL skills. These results suggest the utility of arranging multiple exemplars when OL skills are targeted. Specific skills can be taught, while others are reserved for assessment of the generalization of OL. Given that the goal of training OL is to produce a generalized repertoire, it is important to use this form of design to assess generalization. It was imperative that during the preassessment phase, each OL task variation was conducted to ensure that responding did not occur after observation across each task variation.

After the preassessment, a specific task variation was used to train skills presumed to be necessary to engage in OL. Teaching skills related to the task might also have produced a repertoire of behavior that facilitated task engagement. Palmer (2012) suggested that, to respond after observation, some fine-grained unit of behavior must already be present. It is likely that participants did not initially engage in OL skills because they lacked some or all of the skills necessary to learn observationally within the confines of the OL tasks of this study. For instance, it may have been likely that participants did not build a toy after observation of the experimenter because engaging in the response after the model's actions was not present in their behavioral repertoire (see Table 3). However, after training, all participants acquired the trained skills and correctly responded with some tasks that received no training. Based on previous recommendations (i.e., Taylor &

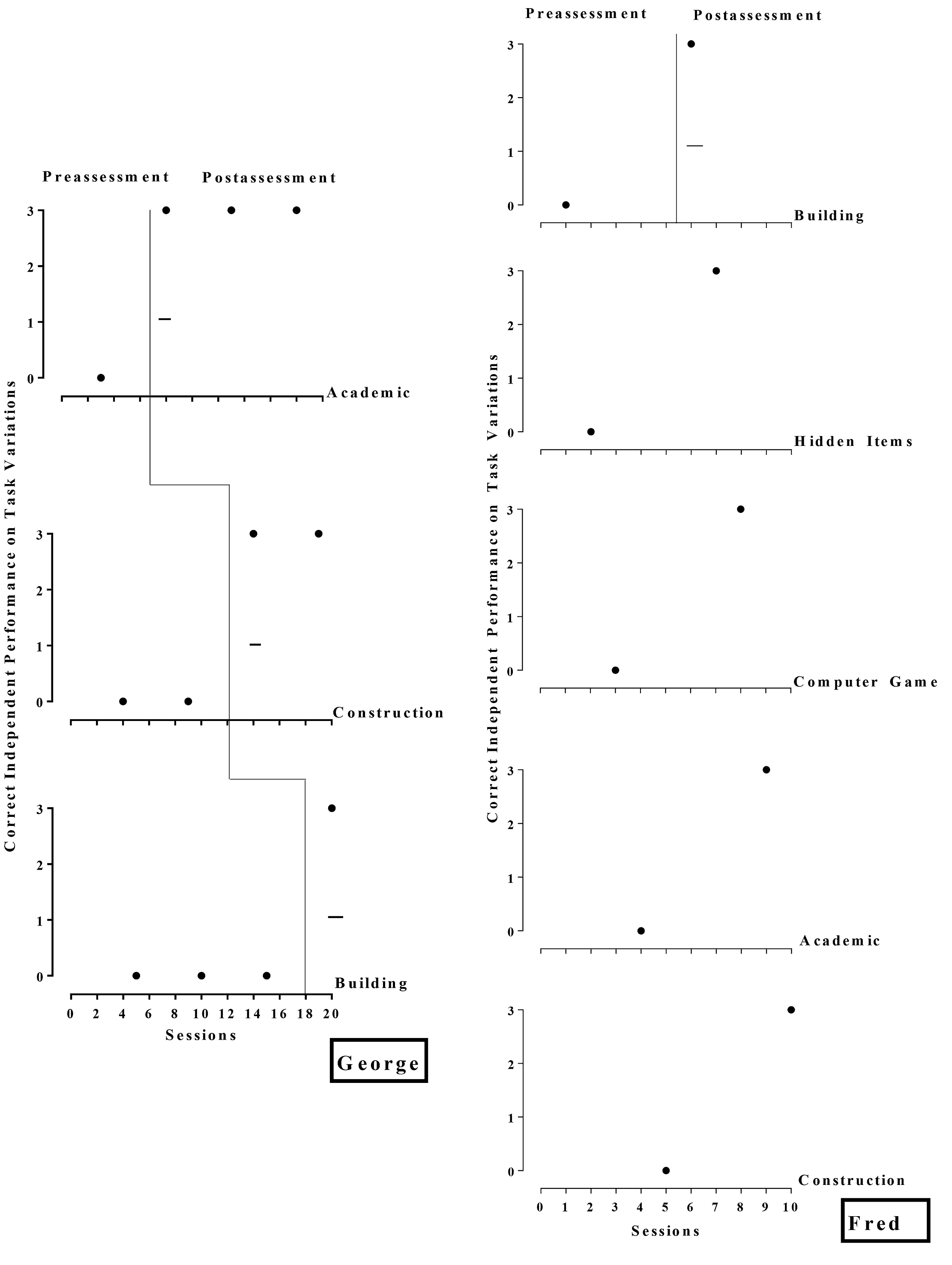

Figure 3. Number of OL task variations performed correctly during preassessment and postassessment across OL tasks for George (left) and Fred (right).

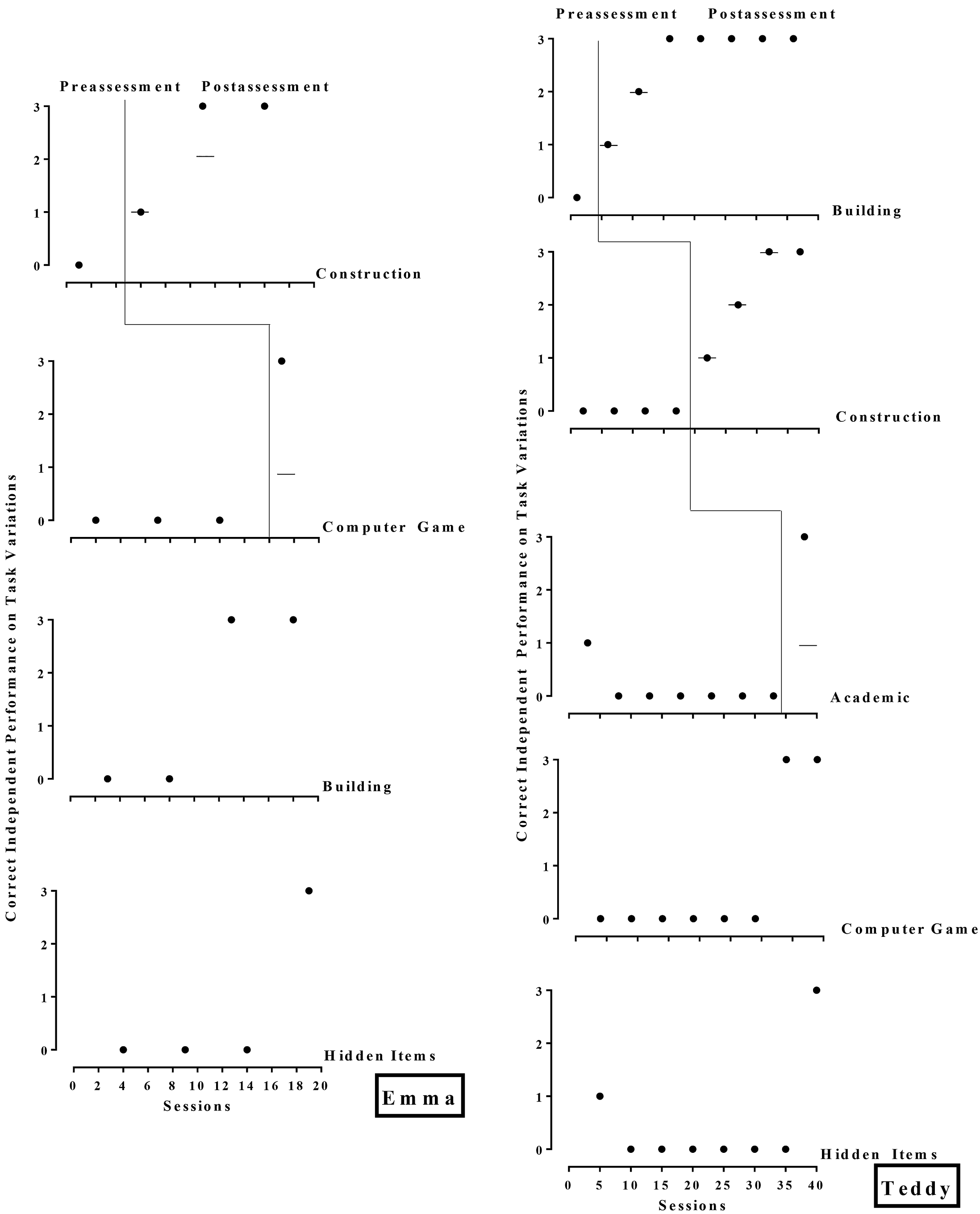

Figure 4. Number of OL task variations performed correctly during preassessment and postassessment across OL tasks for Emma (left) and Teddy (right).

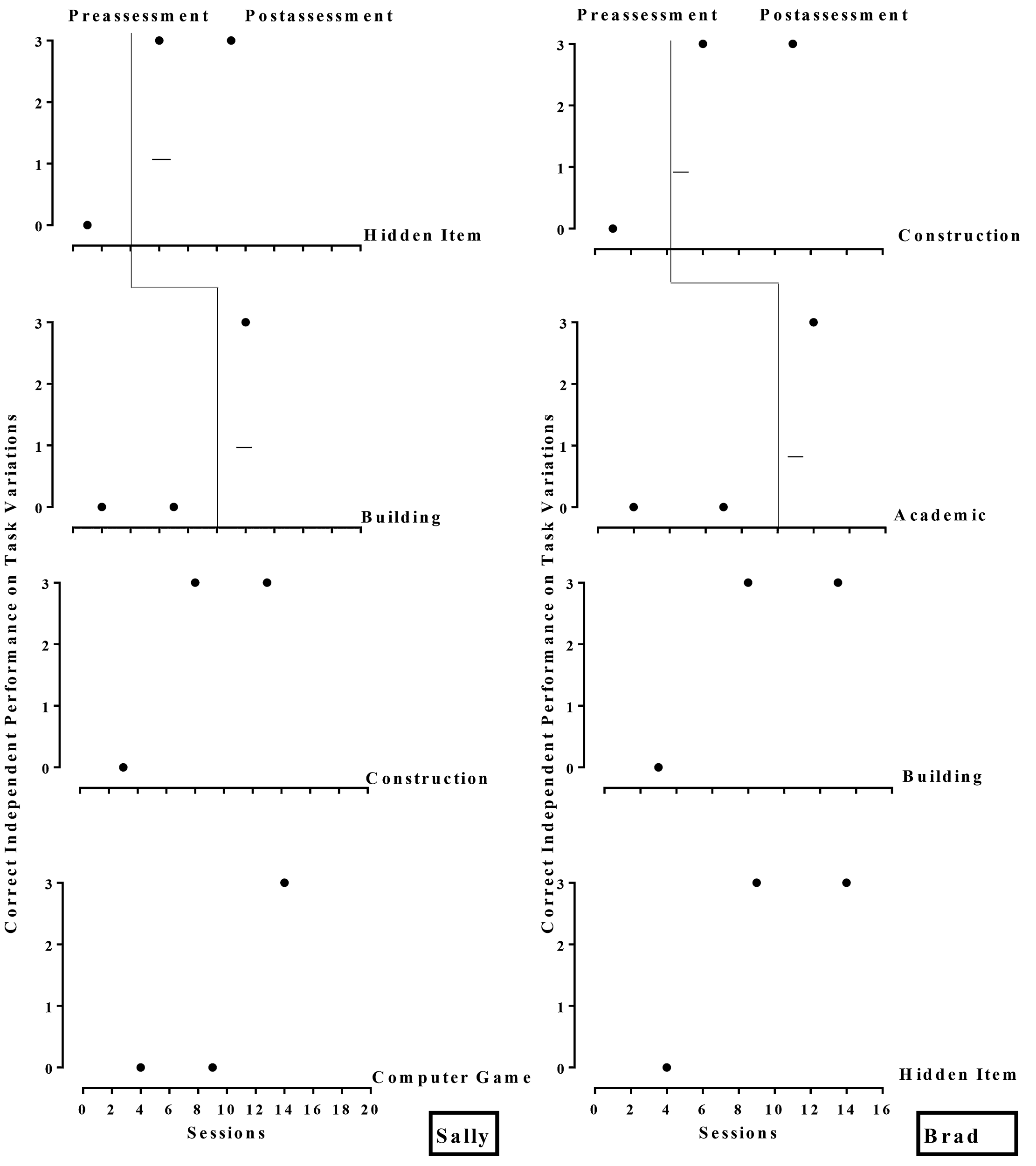

Figure 5. Number of OL task variations performed correctly during preassessment and postassessment across OL tasks for Sally (left) and Brad (right).

DeQuinzio, 2012), this study empirically demonstrated that teaching skills that may be deficient in the individual's repertoire (e.g., imitation, consequence discrimination) might facilitate the performance of OL.

Although generalization of OL across tasks was observed for five of the six participants, it is important to note that all participants were exposed multiple times to the OL tasks. OL tasks without previous exposure were not assessed.

However, one benefit of using five task types and three task variations was that this afforded multiple opportunities to assess the generalization of OL performance. Some degree of stimulus generalization was seen across all participants after training of a specific task variation. Stimulus generalization was demonstrated across all novel task variations for four participants and one of two task variations for two participants. In addition, after training, OL was demonstrated without explicit training across four tasks for Fred, three of four tasks for Sally and Brad, one of three tasks for Emma, and two of five tasks for Teddy. Generalization of OL did not occur across tasks for George. Based on these results, training skills suggested in the previous literature (e.g., Taylor & DeQuinzio, 2012) had a positive effect on the acquisition of OL skills. Broadly, teaching skills and assessing across similar but novel task variations increased the probability that generalization and OL might be detected across different contexts. Future research is warranted to investigate the necessary components to increase the probability of generalization across various tasks. Because we did not evaluate preference for tasks, a possible limitation of the current study is that completing the task might be automatically reinforcing for the student, increasing the probability of OL. Ultimately, for OL to be persistent, engaging in the activity to completion must in itself be reinforcing in order for the behavior to be more likely in the future. Further research should investigate the effects of automatic reinforcement and its place in teaching an OL repertoire. It might also be beneficial to investigate whether OL skills will generalize across different tasks of varying difficulty because of automatic reinforcement for completing the skill.

Although generalized OL was demonstrated for five of the six participants, asserting functional control over responding might be limited. For instance, Fred only required training on one task variation before he engaged in all tasks and task variations. These results do not show an increase in responding when and only when we implemented training across OL tasks. In addition, the timing of increased performance for some tasks did not always correspond to the introduction of training, thus limiting the strength of the assertion of functional control. The design does not permit ruling out any extraneous variables that might have played a part in increasing performance across OL tasks. Future research should attempt to strengthen functional control by replicating this study using an extended concurrent multiple baseline design across subjects in which preteaching measures are collected across all tasks to show the direct and indirect effects of teaching on observational learning.

Another limitation of the present study was that it is unclear exactly whether all skills trained or a subset of the skills (i.e., imitation and consequence discrimination) would produce the same results. Future research should conduct a component analysis of the trained skills to determine exactly what skills are required for OL. This could be completed by specifically training one skill in the component package and then assessing the demonstration of OL with multiple similar, but novel, tasks. In this way, each skill could increase the demonstration of OL without repeated practice effects by using the same stimuli.

A final limitation of the present study was that consequences were not assessed for their reinforcing qualities before the start of the experiment. To address this limitation, future research should use a reinforcer assessment before the study to make sure that the consequences are in fact reinforcers and the neutral consequences are in fact not acting as reinforcers. To that end, to further extend our results, future research should also demonstrate the effects of the establishing operation for the consequences before starting OL training to ensure that the student will be motivated to engage in the action.

Finally, future research should examine the effects of the treatment package across different types of actions. In the present study, only motor responses were required to complete the observed actions. However, mixed results in the literature suggest generalization may or may not occur across varying responses (i.e., motor and vocal responses). Some research has revealed generalization across varying response types (i.e., Baer & Sherman, 1964; Poulson & Kymissis, 1988), but generalization has not been evident across various responses types in all studies (i.e., Young, Krantz, McClannahan, & Poulson, 1994). More research is warranted to identify conditions under which generalization will and will not occur across conditions.

Teaching attending, imitation, delayed imitation, and consequence discrimination was useful for teaching OL, and it might also be useful to teach other related skills such as video modeling or higher order social skills. For instance, Tereshko, MacDonald, and Ahearn (2010) and MacDonald, Dickson, Martineau, and Ahearn (2015) suggested that imitation and delayed imitation were highly correlated with performing well using video modeling. Given that these skills are taught, video modeling might be more successful at producing the targeted behavior after this type of training. Future research should assess the applicability of the treatment package on increasing acquisition via video modeling.

REFERENCES

Baer, D. M., & Sherman, J. A. (1964). Reinforcement control of generalized imitation in young children. *Journal of Experimental Child Psychology*, *1*, 37–49. doi: 10.1016/0022-0965(64) 90005-0

Birnbrauer, J. S. (1976). Mental retardation. In H. Leitenberg (Ed.), *Handbook of behavior modification and behavior therapy* (pp. 361–404). Englewood Cliffs, NJ: Prentice Hall.

Browder, D. M., Schoen, S. F., & Lentz, F. E. (1986). Learning to learn through observation. *The Journal of Special Education*, *20*, 447–461. doi: 10.1177/002246698602000406

Dunn, L. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Circle Pines, MN: American Guidance Service.

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I . (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis*, *25*, 491–498. doi: 10.1901/jaba.1992.25-491

Garcia, E. E. (1976). The development and generalization of delayed imitation. *Journal of Applied Behavior Analysis*, *9*, 499. doi: 10.1901/jaba.1976.9-499

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

MacDonald, R. P. F., Dickson, C. A., Martineau, M. E., & Ahearn, W. H. (2015). Prerequisite skills that support learning through video modeling. *Education and Treatment of Children*, *38*, 33–47. doi:10.1353/etc.2015.0004

Miller, N. E., & Dollard, J. (1941). *Social learning and imitation*. New Haven, CT: Yale University Press.

Orelove, F. P. (1982). Acquisition of incidental learning in moderately and severely handicapped adults. *Education and Training of the Mentally Retarded*, *17*, 131–136.

Palmer, D. C. (2012). The role of atomic repertoires in complex behavior. *The Behavior Analyst*, *35*, 59–73.

Patten, E., & Watson, L. R. (2011). Interventions targeting attention in young children with autism. *American Journal of Speech-Language Pathology*, *20*, 60–69. doi: 10.1044/1058-0360(2010/09-0081)

Pereira Delgado , J. A., & 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.

Poulson, C. L., & Kymissis, E. (1988). Generalization imitation in infants. *Journal of Experimental Child Psychology*, *46*, 324–336. doi: 10.1016/0022-0965(88) 90064-1

Rothstein, M. B., & Gautreaux, G. G. (2007). The effects of a peer-yoked contingency on observational learning and the collateral emergence of naming. *Journal of Early and Intensive Behavior Intervention*, *4*, 453–470. doi: 10.1037/h0100384

Schoen, S. F., & Ogden, S. (1995). Impact of time delay, observational learning, and attentional cuing upon word recognition during integrated small-group instruction. *Journal of Autism and Developmental Disorders*, *25*, 503–519.

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

Tereshko, L., MacDonald, R., & Ahearn, W. H. (2010). Strategies for teaching children with autism to imitate response chains using video modeling. *Research in Autism Spectrum Disorders*, *4*, 479–489. doi: 10.1016/j.rasd.2009.11.005

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

Weiss, M. J., & Harris, S. L. (2001). *Reaching out, joining in: Teaching social skills to young children with autism*. Bethesda, MD: Woodbine House.

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.

Young, J. M., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1994). Generalized imitation and response-class formation in children with autism. *Journal of Applied Behavior Analysis*, *27*, 685–697. doi: 10.1901/jaba.1994.27-685

Received October 3, 2014
Final acceptance May 7, 2015
Action Editor, David Wilder