Teaching Social Skills to Children with Autism using Point-of-View Video Modeling

Allison Serra Tetreault and Dorothea C. Lerman

University of Houston – Clear Lake

Abstract

Video-modeling (VM) is a widely used instructional technique that has been applied to the education of children with developmental disabilities. One form of VM that lacks in-depth analysis is point-of-view video modeling (POVM). The current study investigated the use of POVM to teach three children diagnosed with autism to initiate and maintain a conversation with a conversant. Using a multiple baseline across scripts design, the participants were taught to engage in both eye contact and vocal behavior without the presentation of a vocal discriminative stimulus from the conversant. The treatment package included both the presentation of the target video and reinforcement of target behavior. Although this combination proved successful for increasing the social behavior of two participants, prompts were necessary to achieve acquisition for a third. These data suggest that while POVM may be a successful technique for teaching some social skills, limitations exist that should be further investigated.

The effective use of video modeling to help remediate the behavioral excesses and deficits of children with autism is well documented (Bellini & Akullian, 2007). This strategy has been shown to help establish a variety of skills, including those related to joint attention (e.g., LeBlanc et al., 2003), play (e.g., D'Ateno, Mangiapanello, & Taylor, 2003), self help (e.g., Shipley-Benamou, Lutzker & Taubman, 2002), academic instruction (e.g., Kinney, Vedora, & Stromer, 2003), communication (Wert & Neisworth, 2003), and community survival (e.g., Haring, Kennedy, Adams & Pitts-Conway, 1987). Additionally, video modeling is potentially more effective than teaching through in vivo modeling (Charlop-Christy, Le & Freeman, 2000), and can improve the effectiveness of instructional prompts (Murzynski & Bourret, 2007).

The use of videos to teach social skills has been examined in a recently expanding body of literature. The majority of studies investigating social skills instruction via video models, however, focused on relatively simple behaviors. For example, Bidwell and Rehfeldt (2004)

used video models and contingent praise to teach adults with severe disabilities to initiate an interaction by bringing a cup of coffee to an adult peer. Nikopoulos and Keenan (2004) demonstrated that video models alone were sufficient for teaching three children with autism to initiate an interaction by gesturing or vocally requesting an adult to join the child in play.

A few studies investigated video-based training for more complex social skills. Using video models alone, Maione and Mirenda (2006) obtained increases in the frequency of social initiations and responses of a young boy with autism during two different play contexts. The participant watched videos of two adults engaging in appropriate verbalizations and playing with the target activities. With the implementation of video modeling, the frequency of the participant's use of both scripted and unscripted verbalizations (including initiations and responses) increased during these play sessions. However, reinforcement, video feedback, and prompting were needed to increase behavior in a third play context. The authors reported that some of the modeled statements were novel, while others already existed in the child's repertoire. Charlop and Milstein (1989) showed that video models and reinforcement increased conversational responding for three children with autism. Prior to the intervention, each child exhibited utterances of three to four words in length. The children were taught scripted exchanges consisting of statements with up to eight words per utterance. Each scripted exchange involved an appropriate response to the conversant's question, followed by a reciprocal question to the conversant. While this target represents the most complex set of social behavior taught through video models to date, all of the exchanges were initiated by the conversant. Thus, further investigation of the utility of video models for teaching complex social skills, including those involving initiation of conversation in the appropriate context, is warranted.

Several authors suggest that video modeling is effective because it reduces the amount of irrelevant stimuli in the learning environment, increasing the likelihood that the participant will focus on the most relevant cues (Charlop-Christy et al, 2000; Krantz, MacDuff, Wadstrom, & McClannahan, 1991). If so, video formats that further reduce irrelevant stimuli may help promote learning. One format that may serve to reduce additional irrelevant stimuli in the learning environment is point-of-view modeling (POVM). In this type of modeling, the camera angle is presented at the participant's eye level and shows only what the participant might see within the context of the targeted activity, skill, or context (i.e., from his or her own viewpoint). Depending on the target skill, the participant might view a specific setting or a pair of hands completing a task.

One potential advantage of POVM over the typical, or scene view, video model is that it further restricts the stimuli to those that are directly related to the target behavior, eliminating the necessity of identifying optimal characteristics of the model (Hine & Wolery, 2006). The extent to which POVM has been utilized is unclear, however, because many prior studies did not include detailed descriptions of the video modeling procedures. To date, only four studies have explicitly evaluated the POVM technique (Alberto, Cihak, & Gama, 2005; Hine & Wolery, 2006; Schreibman, Whalen, & Stahmer, 2000; Shipley-Benamou et al., 2002) and these studies included a few specific procedural features in common while other features varied. For example, participants received prompts and praise for watching the video model in each case and some degree of generalization and maintenance of behavioral gains was observed in each study. However, many aspects of POVM differed across studies. For example, only Alberto et al. (2005) prompted rehearsal during video viewing and incorporated least-to-most prompts during post-viewing practice sessions. The delay between video viewing and practice opportunities was not specified in this study and in Schreibman et al. (2000), whereas practice occurred immediately after the presentation of the video in Shipley-Benamou et al. (2002) and Hine and Wolery (2006). Finally, participants received reinforcement for correct responding during practice sessions in some studies (e.g., Shipley–Benamou et al., 2002; Hine & Wolery, 2006) with no explicit mention in others.

The above investigations demonstrated the effectiveness of POVM for teaching self-help skills, play skills, and compliance with transitions. However, additional research is needed to determine the ease with which social and communication skills – two primary core deficit areas for children diagnosed with autism -- may be acquired through this teaching approach. The purpose of the current study was to investigate the efficacy of POVM for teaching children with autism to initiate and maintain social interactions with others. The extent to which these skills generalized across materials and maintained over time also was evaluated.

Method

Participants

Participants were three children diagnosed with mild-moderate to severe autism by an independent psychologist. During the course of the study, all participants received behavior analytic services at a private center for 6 hours each day, 5 days per week. They lived at home with their parents and received various other therapies outside of the private center (e.g., occupational therapy, auditory integration

training, dietary supplementation). Each child's language abilities and autism severity were assessed prior to the study using the *Preschool Language Scale, Fourth Edition* (*PLS-4*; Zimmerman, Steiner, & Pond, 2002) and the *Childhood Autism Rating Scale* (*CARS*; Schopler, Reichler, & Renner, 1988), respectively. Zhane was 5 years, 5 months at the time of the study and had attended the center for 2.5 years. His receptive language abilities were assessed to be the age-equivalent of 2 years, 3 months, and his expressive language abilities were 2 years, 9 months. Zhane's autism severity score was 39, which falls in the severe range of symptomatology. Randall was 8 years, 2 months and had attended the center for 2 years, 3 months. His receptive language abilities were assessed to be the age-equivalent of 3 years, 4 months, and his expressive language abilities were 3 years, 1 month. Randall's autism severity score was 35.5, which falls in the mild-moderate range of symptomatology. Janet was 4 years, 4 months and had attended the center for 10 months. Her receptive and expressive language abilities both were assessed to be the age-equivalent of 3 years, 10 months. Janet's autism severity score was 32.5, which falls in the mild-moderate range of symptomatology. The children were selected for the study because they did not engage in spontaneous social initiations but could imitate three- to four-word sentences. None of the participants had exposure to video models as an instructional strategy prior to the study.

Setting and Stimulus Materials

All sessions were conducted in a small (2.4 m by 4.6 m) room at the day treatment center. The room contained a child-sized table and chairs, a filing cabinet, a bookcase, and a tripod-mounted video camera, as well as any materials relevant to the session (described below). During training sessions, a portable DVD player played the video of the designated script.

Three scripted sequences of social initiations were prepared (see sample script in Table 1). Each script focused on a different situation that would set the occasion for a social initiation by the participant. These activities were selected in consultation with the day treatment center supervisor and incorporated free-play items available during breaks from instructional time. Each script modeled on the video was associated with specific materials. The "Get Attention" script involved getting a conversant's attention for the purpose of displaying a creation made with a marker and a dry erase board. The "Request Assistance" script was designed to teach a request for a conversant's assistance in attaining and opening a clasped plastic box that contained a bottle of bubble solution. The "Share a Toy" script involved offering a Mr. Potatohead® doll to a conversant and then requesting it back

again. In addition, two sets of generalization materials were selected for each script. The scripted statements were designed to allow for different materials to be presented in the situation. For example, the conversation created for "Share a Toy" could be used to offer and request any toy, not just the Mr. Potatohead® doll used in the video and practice sessions (see Table 1). The materials shown in each video clip and the materials used to evaluate generalization are listed in Table 2. Relevant materials were present in the baseline and post-viewing practice sessions for each target script.

Each video model began with a brief visual introduction (separate slides which showed "1", "2", "3", "GO!") followed by three repetitions of the target script. Subsequent to the first presentation, a brief visual transition ("READY!", "GO!") preceded the two consecutive presentations. The total durations for the "Get Attention", "Request Assistance", and "Share a Toy" videos were 2:21, 2:33, and 2:31, respectively. All camera angles on the video were shot from the first person perspective (i.e., POVMs) as illustrated in the Figure 1 screen shots that correspond with the script in Table 1. During filming, the camera was swiveled on the tripod to mimic natural head movements and brief (e.g., 2 s to 3 s) eye contact with the conversant, who was an unfamiliar graduate student. A female adult who was not in view (the first author) spoke the target verbalizations. Because of this person's proximity to the camera, the participant's lines were spoken more loudly than the conversant's lines in the final videos. All videos were recorded in a location unfamiliar to the children.

Response Measurement and Reliability

All post-viewing practice sessions were videotaped for data collection purposes. During these sessions, the tripod and camera were placed in an unobtrusive position over the left shoulder of the conversant (i.e., the adult with whom the participant practiced the target skills) to adequately capture eye contact. Data were collected on the behavior in the target script, as well as on all of the children's novel vocal behavior. All scripts were composed of five specific exchanges (see example in Table 1). For the purpose of this study, an exchange was defined as eye contact and vocal behavior from the child that occurred prior to the vocal behavior of the conversant. Each script began with a social initiation from the child in the form of a greeting ("Get Attention": "Hi there!"; "Request Assistance": "I'm glad to see you!"; "Share a Toy": "Hey there!"). Correct and incorrect vocal behavior and eye contact were scored for each exchange. A correct vocal behavior was scored if the child said the exact sentence from the video or a sentence that differed by no more than two words (added or deleted)

Table 1

An Example Script: "Share a Toy"

	Conversant		Participant	
Exchange	Actions	Statement	Actions	Statement
1	enters the room, looks at the participant		looks up from the toy, looks at the conversant	"Hey there!"
2	maintains eye contact, sits at the table	"Hi!"	looks at the toy, looks at the conversant	"I'm playing with Mr. Potatohead®.
3	looks at the toy, looks at the participant	"That looks like fun!"	looks at the toy, looks at the conversant	"Would you like to play?"
4	looks at the toy, looks at the participant	"Yes! Thank you!"	looks at the toy while the conversant plays, looks at the conversant	"May I play some more?"
5	looks at the toy, looks at the participant	"Sure. Here you go."	looks at the toy, plays with the toy, looks at conversant	"Thank you!"

from the target script (e.g., "Circle" instead of "It's a circle."). For the initial social exchange, any appropriate greeting (e.g., "Hi", "Hello", and the script examples above) said by the child was scored as correct regardless of the modeled greeting for that script. Correct eye contact was scored if the child looked at the conversant for any amount of time immediately prior to, during, or following (within 2 s) the target vocal behavior. These data were collected using pen-and-paper data sheets that listed the target vocal behavior for each exchange. Each sheet also included space to transcribe novel vocal behavior; however, no increases in appropriate novel language occurred for any of the children during treatment and therefore no data are presented for this measure. The number of exchanges consisting of both correct eye contact and vocal behavior was totaled for each post-viewing practice session. Data were collected during sessions by the experimenter and were verified by videotape at the end of each day.

A secondary observer collected data from video independently during 38%, 42%, and 43% of post-viewing practice sessions for
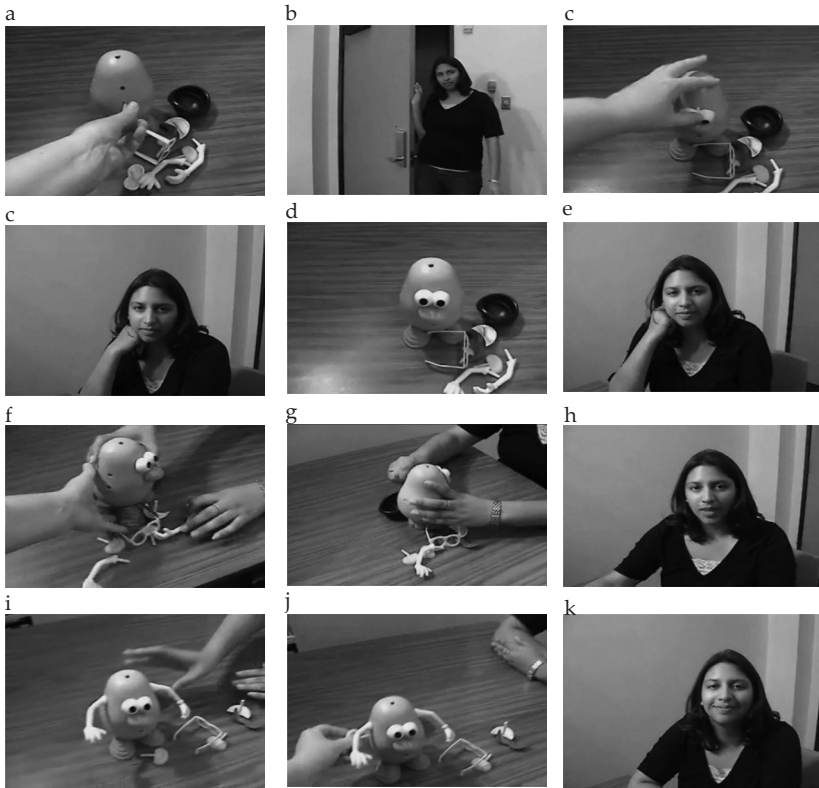
Figure 1. Screen shots from the "Share a Toy" script.

Zhane, Randall, and Janet, respectively. These data were compared for each instance of eye contact and vocal behavior during a session. An agreement was scored if both the primary and secondary observer mutually recorded the occurrence or nonoccurrence of a specific vocal or play target behavior. Interobserver agreement was calculated by dividing agreements by agreements plus disagreements and multiplying by 100. Across all subjects, scripts, and conditions, interobserver agreement averaged 93% (range, 70% to 100%). Observers also collected data on the conversant's presentation of scripted statements to ensure integrity. Across all subjects, scripts, and conditions, accuracy of scripted conversant behavior averaged 99% (range, 80% to 100%). Data were not collected on other forms of conversant or experimenter behavior (e.g., reinforcer and prompt delivery).

Table 2

Materials for Video-Model and Generalization Scripts

	"Get Attention"	"Request Assistance"	"Share a Toy"
Video Model	Dry erase board Dry erase marker	Clasped plastic box Bottle of blowing bubbles Screw-top plastic container	Mr. Potatohead®
Generalization Set A	Playdoh®	Key-locked shape sorter Bottle of blowing bubbles	Plastic toy bus
Generalization Set B	Interlocking building blocks	Bottle of blowing bubbles	Plastic toy dragon

Note. For "Share a Toy", generalization to a third and fourth toy were tested for Randall. Generalization Set C included a plastic toy dinosaur and Set D included a twirling toy.

Experimental Design and Procedure

The study employed a multiple baseline across behaviors (scripts) design. Initial script assignment was counterbalanced across participants. Each participant began treatment on one of the three scripts while baseline data were collected for the remaining two scripts. Probes for generalization to novel sets of materials were conducted throughout all baseline and treatment phases. Once a participant attained mastery on the first intervened-upon script (see below), treatment began on a subsequent script. Therapists from the treatment center acted as conversants and were randomly rotated throughout all conditions and across all participants. The adult presented in the video model was not employed by the day treatment center and was never present.

Baseline. During baseline sessions, the child sat at a table with the relevant stimulus materials for the target script. One adult conversant was present. The child was instructed that the conversant

would return shortly and that he or she should play nicely at the table until then. No video was presented. Within 20 s of exiting the treatment room, the conversant re-entered the room by knocking lightly on the door, stepping into the room, and closing the door behind her. The conversant performed each action and stated each assigned line within 10 s of the beginning of an exchange in the script (see Table 1), regardless of participant behavior. For example, if the participant did not respond within 10 s of the conversant entering the room, the conversant stated the scripted line of Exchange 1 and waited another 10 s for the participant to initiate the next exchange. This procedure ensured that each participant behavior could occur with equal opportunity in every session. No contingencies were programmed for eye contact or vocal behavior.

Video plus food. This treatment phase was conducted by one adult acting as the conversant and a second acting as the trainer (the first author) during video viewing and practice sessions. The trainer was responsible for setting up session materials, delivering reinforcers, and cueing the conversant (as described below). The trainer was constant across all treatment conditions for all participants.

The participant was seated at the table with the DVD player and the video model for the target script. The trainer sat behind the participant for the duration of the video viewing and the post-viewing practice session. Prior to beginning the video, the trainer stated, "Let's watch a movie!" During viewing, reinforcement was periodically provided contingent on attention to the video model (i.e., every 10 to 15 s for eye contact with the television screen) using food items identified via a multiple-stimulus-without-replacement preference assessment (DeLeon & Iwata, 1996) conducted immediately before each training session. Immediately following the video viewing, the trainer initiated a post-training practice session by placing the relevant materials on the table and stating, "Let's practice." These sessions were identical to baseline with the following exceptions. Food items were delivered by the trainer contingent on scripted exchanges with the conversant (see Table 1). Immediately following the child's scripted behavior, the conversant engaged in the scripted vocal response. If the child did not engage in any part of an exchange, the trainer cued the conversant when 10 s had elapsed by holding up the next written statement on an index card behind the participant and out of the participant's view. A participant attained mastery if any 8 (out of 10) scripted eye contact and vocal behaviors occurred per session across three consecutive sessions. Under this criterion, a given trial might have included the vocal behavior without the eye contact, for example, and have been counted as one correct behavior (i.e., mastery was not based on the occurrence

of both correct behaviors on every exchange).

Video only. During this condition (Janet only), the video model was shown as described above and the adult roles were the same as described above, but no additional components of the treatment package were in effect. That is, no food reinforcers were delivered during video viewing or during the post-viewing practice session. During the video-plus-food condition, Janet began to orient towards the trainer instead of the conversant each time the conversant spoke suggesting that conversant statements became discriminative for the delivery of food reinforcers although they were only delivered contingent on correct participant behavior. Previous research using video models indicated that it was possible for some participants to acquire skills through video modeling without the inclusion of programmed reinforcer delivery (e.g., Charlop-Christy et al., 2000, see Krantz et al, 1991, for a review). Therefore, this condition was implemented to test for acquisition in the absence of arbitrary reinforcers, as well as to eliminate the strengthening of inappropriate stimulus control (i.e., looking for a food item each time the conversant spoke). Sessions were otherwise identical to those in the video-plus-food condition.

Least-to-most prompts. This phase was introduced to facilitation acquisition of the target behaviors for Janet after she did not meet the mastery criteria during prior conditions. These sessions were identical to video-plus-food sessions with the following alteration. If a correct vocal response did not occur within 10 s of an opportunity during post-viewing practice sessions, the trainer instituted a three-step least-to-most prompting procedure (Horner & Keilitz, 1975). First, the trainer provided a gesture prompt by pointing to the conversant. If the child did not engage in the vocal response within 2 s to 3 s, the trainer continued to provide a gesture prompt along with a partial verbal model of the beginning of the child's scripted line (e.g., "I'm glad …"). If the child did not engage in the target behavior within 2 s to 3 s, the trainer combined the gesture prompt with a full vocal model of the child's scripted line (e.g., "I'm glad to see you"). In this phase, food items were delivered if behavior occurred independently or with only a partial model; edibles were not delivered if a full model was used. In a later repetition of this phase, the trainer was eliminated from the post-viewing practice session and the conversant delivered prompts and food items. This alteration was made because, as mentioned above, Janet began to attend to and engage with the trainer instead of the conversant. The mastery criteria were identical to those in the video-plus-food condition.

Generalization.   Probes with the generalization materials were conducted throughout all conditions of the study using the procedures described in the baseline condition.

Maintenance. The procedures were identical to those in the baseline condition with a modification introduced for Randall and Janet after a decrement in responding was observed. Contingent food items were reintroduced to determine if this decrease in behavior was due to extinction effects (labeled "Food Only" on graphs). Sessions were identical to those in the video-plus-food condition except that the children did not watch the video prior to the practice session, and the trainer was not present during these sessions. Instead, the conversant delivered food items contingent on correct behavior.

Follow-up. Procedures were identical to those in baseline. Follow-up data were collected 10 days after the last maintenance session for Zhane only, as his rapid performance during treatment allowed for follow-up assessment during the study timeline.

Results

For each participant, two figures are presented with the first illustrating the number of correct (i.e., both eye contact and vocal behavior) exchanges in each session while the second depicts data on the occurrence of eye contact and scripted vocal behavior separately. Because either behavior could occur in the absence of the other on each exchange, the information presented in the second figure provides a more sensitive analysis of behavior over the course of intervention. Furthermore, the mastery criterion was based on the independent occurrence of eye contact and vocal behavior, regardless of whether they occurred together during the same exchange (i.e., any 8 of 10 behaviors across three consecutive sessions). However, data in the first figure also are important to evaluate because reinforcement was delivered contingent upon a fully correct exchange (i.e., both aspects correct for a given exchange).

Zhane's performance is depicted in Figures 2 and 3. During the "Request Assistance" script (top graph of each figure), Zhane did not engage in any correct exchanges (Figure 2) during baseline with a gradual increase during the video-plus-food condition and mastery in 14 sessions. A return to the baseline condition during maintenance produced a brief decrease in exchanges with subsequent maintenance at or above the mastery criterion and skill maintenance at follow-up. No correct exchanges were observed during the generalization probes until the maintenance condition, and generalization sessions for both sets of stimuli met mastery levels during the follow-up condition. Figure 3 shows that eye contact occurred more frequently than vocal behavior in baseline. With the implementation of the video-plus-food condition, eye contact increased more rapidly than vocal behavior, and eye contact continued to occur at a higher frequency than vocal
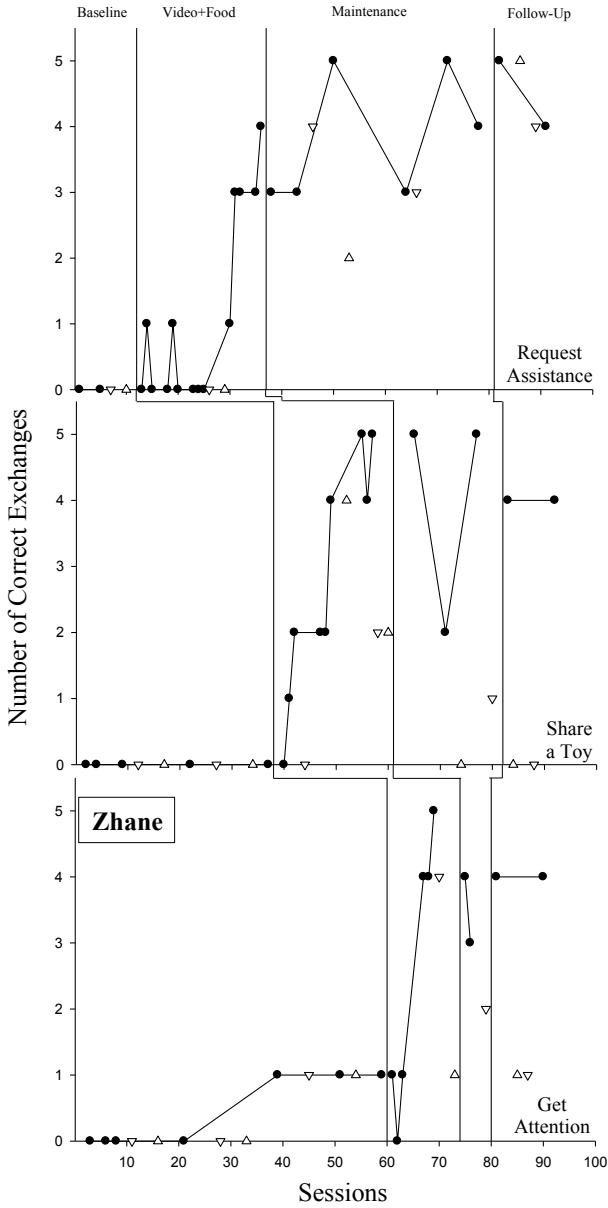
Figure 2. Number of correct exchanges during baseline and treatment phases for Zhane. Open triangles represent generalization probes.
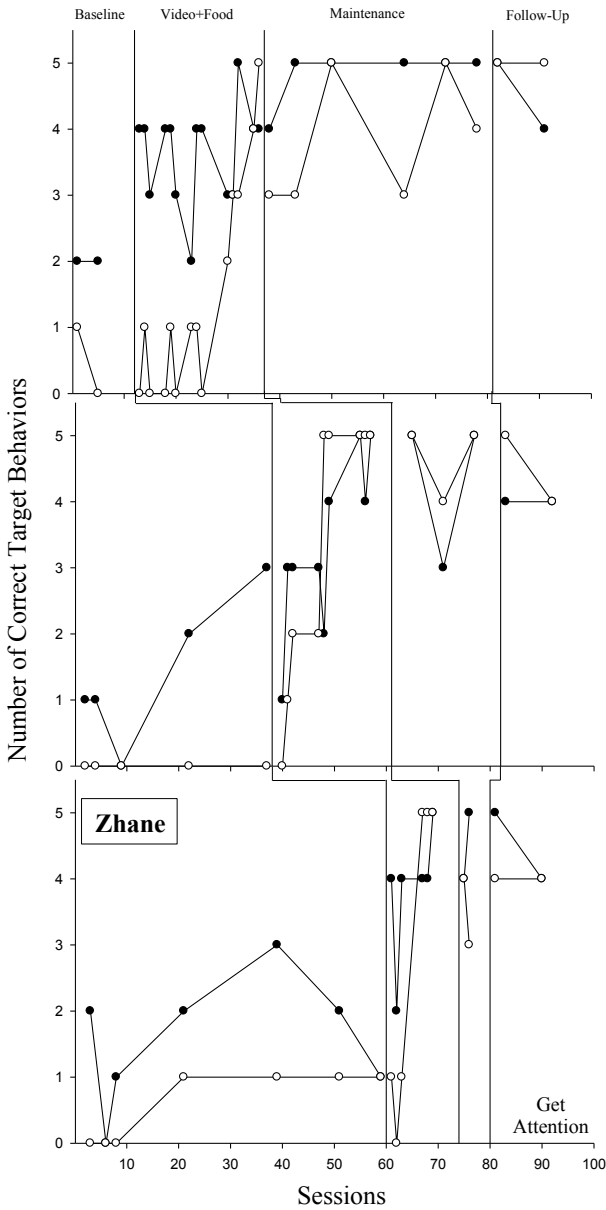
Figure 3. Number of correct eye contacts (closed circles) and scripted responses (open circles) during baseline and treatment phases for Zhane. (Generalization probes are not shown.)

responses during maintenance. For the "Share a Toy" script (middle graphs), no correct exchanges (Figure 2) occurred during baseline but increases occurred once treatment was implemented with mastery in 9 sessions. Correct exchanges continued during the maintenance and follow-up sessions. In the generalization probes, no correct responding occurred until treatment and the increases were not maintained during the maintenance and follow-up phases. Figure 3 illustrates that no correct vocal responses occurred during baseline, though eye contact did increase during baseline. The video-plus-food condition resulted in increases in both eye contact and vocal responses that maintained at approximately the same frequency during maintenance and follow-up. Similar results were obtained for the third script, "Get Attention" (bottom graphs), with the mastery criterion for exchanges (Figure 2) obtained in 6 sessions of intervention and maintained during maintenance and in follow-up. Eye contact and vocal behavior (Figure 3) increased simultaneously during intervention; however, little generalization to the novel materials was observed.

Randall's performance is depicted in Figures 4 and 5. During the "Share a Toy" script (top graphs), Randall did not engage in any correct exchanges (Figure 4) during baseline, and there was no increase during the video-plus-food condition. Figure 5 shows that both eye contact and vocal responses occurred at baseline levels. However, Randall's therapists at the day treatment center reported that he was using the scripted vocal behavior appropriately during his extra-experimental teaching sessions. It was hypothesized that the presentation of the video immediately before practice sessions may have created an abolishing operation for responding. Therefore, a baseline probe was conducted, during which Randall responded with all 10 of the modeled behaviors. Subsequent baseline sessions were conducted (labeled "Maintenance" on the figures), but there was a drastic decrease in correct exchanges after 4 sessions. Figure 5 shows that a decrease occurred for both eye contact and vocal behavior. Because a change from treatment to baseline conditions included the removal of both the video viewing component and the delivery of response-contingent food items, a food-only condition was initiated and correct exchanges met mastery in 4 sessions and responding continued in maintenance (only 2 sessions were conducted due to study timeline). Across all phases, Randall's exchanges during generalization probes did not increase above baseline levels.

With the implementation of treatment for the "Request Assistance" script (middle graphs), correct exchanges gradually increased to mastery in 15 sessions and continued in maintenance, though generalization was limited (Figure 4). Eye contact increased during

baseline for this script when intervention began with the first script (Figure 5) and maintained with the introduction of video-plus-food condition, although both eye contact and vocal behavior continued at approximately the same frequency during maintenance. During baseline for the "Get Attention" script, correct exchanges never exceeded one (the initial greeting; Figure 4, bottom graph). With initiation of the video-plus-food condition, Randall's responding reached mastery in 14 sessions and continued in maintenance, though little generalization occurred. Although eye contact increased during baseline with this script as well, the behavior decreased prior to the intervention. Similar increases in eye contact and vocal behavior occurred during the video-plus-food condition (Figure 5).

Data for Janet are displayed in Figure 6 and Figure 7. Janet's baseline responding during the "Get Attention" script (top graphs) was at zero levels with little increase after 12 sessions in the video-plus-food condition. Both eye contact and scripted vocal behavior remained infrequent (Figure 7). Anecdotally, it was noted that Janet was attending more to the experimenter seated behind her (who provided the food reinforcers) than to the target conversant, although no eye contact or other forms of attention were delivered by the experimenter. To control for this behavior, the video-only phase was initiated after a return to baseline; however, correct exchanges did not increase (Figure 6) and both eye contact and vocal behavior decreased (Figure 7). The least-to-most-prompts condition was then implemented correct exchanges quickly increased with mastery in 10 sessions. Figure 7 shows that there was a more rapid increase in vocal responses than eye contact. However, an immediate decrease in exchanges occurred during a return to baseline. Because the video, prompts, and response-contingent food items had been removed, a food-only condition was introduced based on the assumption that the removal of reinforcement had extinguished correct responding. Nonetheless, correct exchanges did not increase under this condition. It was hypothesized that the decrease in exchanges during maintenance could instead have been due to the absence of the experimenter (who had previously delivered prompts). To establish stimulus control in the presence of the conversant alone, the conversant began to provide prompts in the next phase (labeled "Conversant Prompts" on the graph) and correct exchanges met mastery in 2 sessions with maintenance across 3 additional sessions. Figure 7 shows that both eye contact and scripted vocal responses increased concurrently. In the next phase, the prompts were removed while the delivery of response-contingent food items continued and Janet's performance maintained. Across all phases, generalization to new stimuli did not occur.
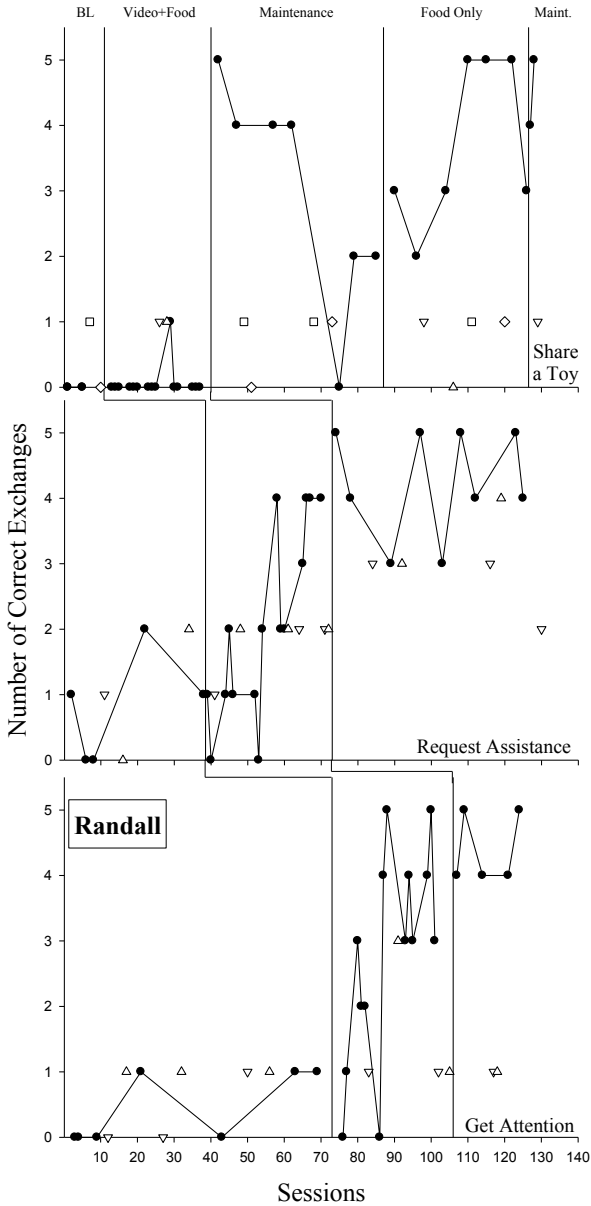
Figure 4. Number of correct exchanges during baseline and treatment phases for Randall. Open triangles represent generalization probes (Sets A and B). Generalization Sets C and D of "Share a Toy" are represented by open diamonds and open squares respectively.
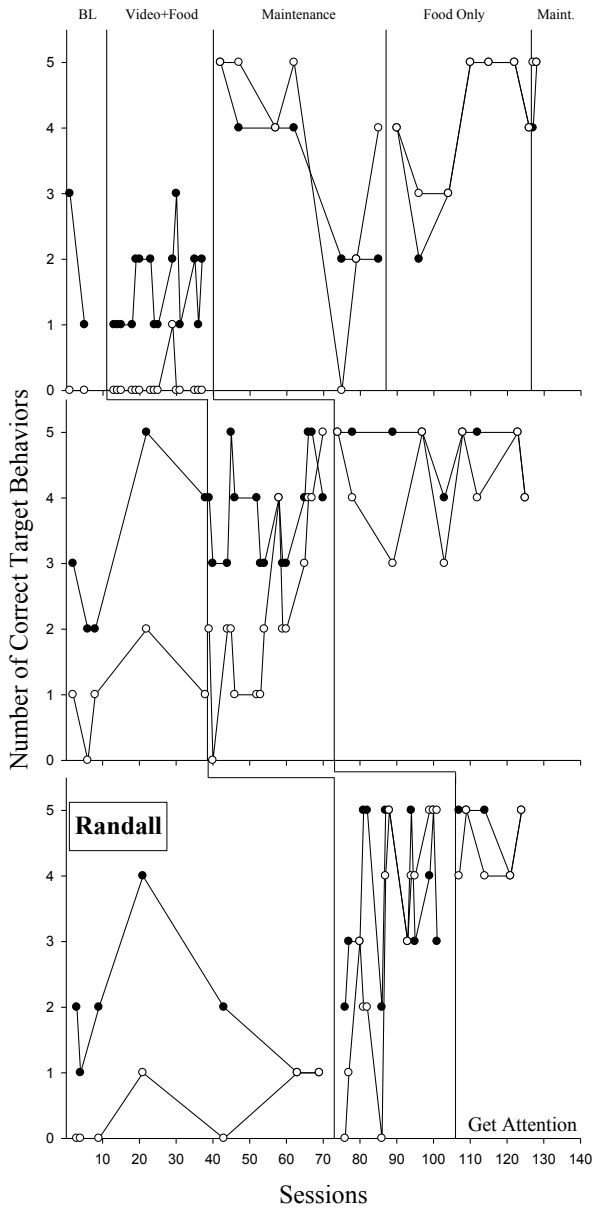
Figure 5. Number of correct eye contacts (closed circles) and scripted responses (open circles) during baseline and treatment phases for Randall. (Generalization probes are not shown.)
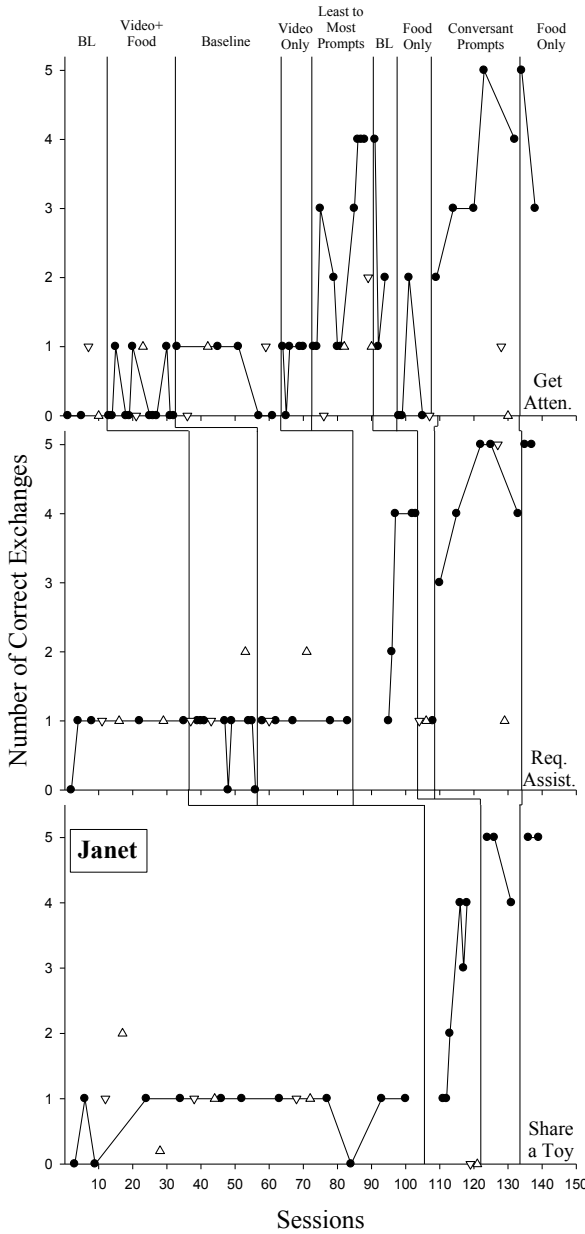
Figure 6. Number of correct exchanges during baseline and treatment phases for Janet. Open triangles represent generalization probes.
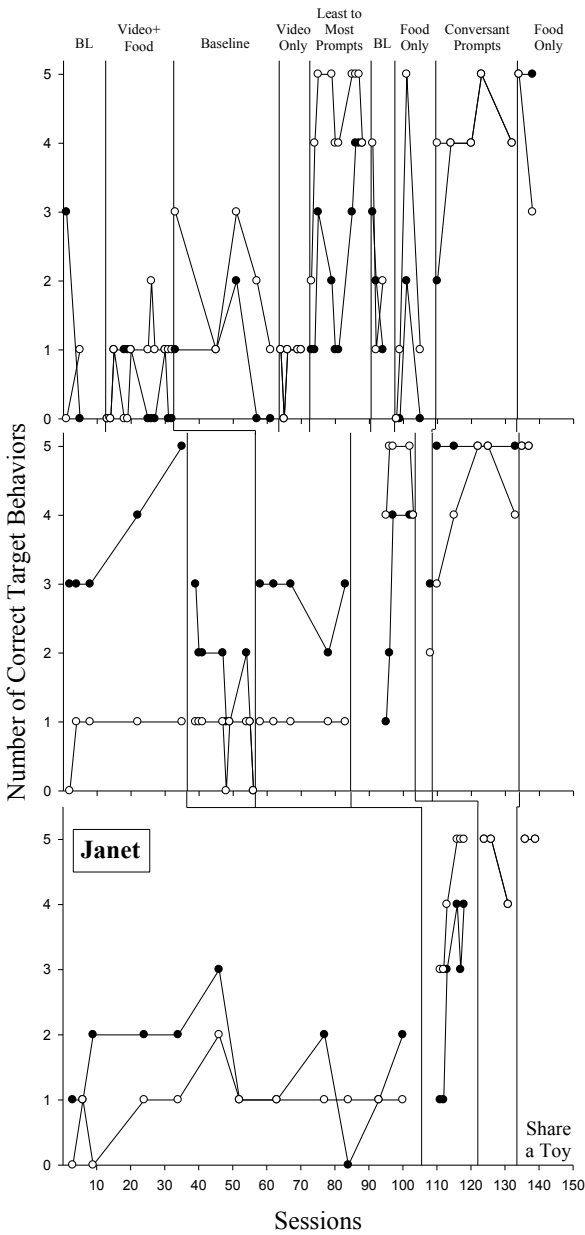
Figure 7. Number of correct eye contacts (closed circles) and scripted responses (open circles) during baseline and treatment phases for Janet. (Generalization probes are not included.)

Similar results were obtained for the "Request Assistance" script (middle graphs). Treatment began with the least-to-most-prompts condition, and correct exchanges reached mastery in 5 sessions. As with the first script, a food-only condition was introduced briefly, but behavior decreased to only one correct exchange immediately. Correct exchanges returned to mastery in 2 sessions and maintained across 3 additional sessions during the conversant-prompts condition, and maintained when prompts were removed in the following phase. Figure 7 shows that frequency of eye contact increased during baseline but decreased during the video-plus-food condition. As with the first script, vocal behavior increased more rapidly than eye contact when least-to-most prompts were introduced, but both behaviors occurred at approximately the same level during the conversant-prompts and food-only conditions. There was limited generalization across all conditions. With the implementation of the least-to-most-prompts condition for the "Share a Toy" script (bottom graphs), mastery was met in 6 sessions and correct exchanges continued in the conversant-prompts and food-only conditions, though no generalization occurred. It can be seen in Figure 7 that both eye contact and vocal behavior increased simultaneously with the implementation of treatment for this script.

Discussion

Results of this study are inconclusive regarding the overall effectiveness of POVM to teach social exchanges to children with autism. Responding on all three scripts came under the control of the video and reinforcement contingencies for 1 of the 3 participants (Zhane). For a second participant (Randall), two scripts were readily taught using the video modeling package intervention whereas an additional script required modification. For a third participant (Janet), response prompts were necessary to increase the frequency of eye contact and social initiations.

For all participants, eye contact appeared to generalize across baselines to some extent and was acquired and maintained somewhat more often than scripted vocal behavior (see Figures 3, 5 and 7). There are two likely explanations for these findings. First, the eye contact modeled in each video involved the same topography of shifting gaze from the materials to the person in all three scripts, while target vocal behavior was different in each case. Second, the action of eye contact (e.g., the motion of the camera) was clearly visible in the video model, whereas the scripted vocal responses were stated by a person not seen on the video which is a potential drawback of the POVM compared to a scene model. However, these findings were not robust, so further analysis of POVM for teaching various forms of social behavior should be investigated.

Although Zhane's frequency of correct exchanges clearly increased with the introduction of treatment across all scripts, his behavior did not generalize to the materials used during probes for two of the three scripts. Randall's response to treatment was perplexing. Anecdotal reports indicated that he had acquired the behavior shown in the video (i.e., saying lines from the script while engaging in eye contact), but it is not clear why he did not engage in these behaviors during post-viewing practice sessions for the first script. His mother reported that he frequently engaged in delayed echoing of lines from his favorite movies at home. However, it appeared that he did not generalize from the video model to the in-vivo practice session. The intervention was nonetheless effective with the other two scripts. For Janet, the video model and reinforcement alone were insufficient to increase correct exchanges and response prompts were necessary to increase her eye contact and vocal behavior. The obtained results suggested that Janet's responding was at least partly controlled by the behavior of the experimenter, who during treatment delivered prompts and reinforcement for Janet's exchanges with the conversant. However, the treatment components responsible for the increase in correct exchanges are unclear. Additional analyses comparing the efficacy of video modeling alone to the prompting procedure alone would provide more information about Janet's acquisition of social behavior.

Although the general treatment effects were replicated across three scripts for each participant, few correct social exchanges occurred in the presence of materials that did not appear in the videos. For example, Zhane could talk about and share Mr. Potatohead®, but he did not do so with a toy bus or with a toy dragon. These results suggest that generalization is unlikely to occur if training is restricted to a single set of materials. However, the scripts were designed to be maximally different while including components of social referencing and verbal initiations. The "Get Attention" script involved showing off an item that the child had created (a drawing, a model built from Playdoh®, or a structure built from blocks), the "Request Assistance" script focused on requesting an out-of-reach item and assistance to open it (a clasped, screw-top, or locked container with bubbles visible inside), and the "Share a Toy" script was about sharing a toy and then requesting it back again (Mr. Potatohead®, a bus, or a dragon).

It is unclear why the current study failed to replicate the results of previous research on POVM. One possibility is our use of an off-screen modeled response (e.g., the scripted vocal statements), as mentioned above. When the intended model is not clearly visible on camera, as in the case of hands manipulating materials, the stimuli that should signal behavior may be more ambiguous. Another possibility

is the complexity of the social exchanges examined in this study. The current analysis selected target behaviors that have not been previously studied using POVM and that have rarely been studied with traditional video models. Participants not only were required to make brief eye contact with each social exchange, but they also were required to make a statement that was not dependent on the previous statement of another person (i.e., initiation of the interchange). It is possible that these skills would not have been acquired through traditional scene video modeling either. Further analysis of the usefulness of the POVM technique to teach social skills to children with autism is needed. An intermediary step between simple social skills (e.g., greetings) and more complex skills like those assessed here is warranted.

It would also be beneficial to determine which components of the present procedure were necessary to produce the desired results. For example, while a model and reinforcement were sufficient to change the behavior of one participant, these components were not entirely sufficient for another, and prompts were necessary for behavior change in a third. Also, it may have been unnecessary to include both a trainer and a conversant. From Janet's results, one person may have sufficed to implement the intervention. A component analysis may help to identify the necessary ingredients for an effective video modeling treatment package. It is unclear to what extent the addition of an arbitrary reinforcer aided in the acquisition of target behavior. It is possible that the inclusion of highly preferred activities, assessed for each participant, may have made the social interaction itself more reinforcing and reduced the need for arbitrary reinforcers. Although not included in many other studies of video modeling, reinforcement was included here because of the unlikelihood that parity of behavior alone (i.e., similarity to a model) would have acted as a reinforcer (e.g., Horne & Erjavec, 2007). This assumption should be tested further with video modeling techniques.

Anecdotally, the intervention was associated with collateral decreases in self-stimulatory vocalizations in the post-viewing practice sessions for all children. During baseline sessions prior to treatment, both Zhane and Randall mumbled statements to themselves that were difficult for others to hear. Once treatment began, these responses did not occur in practice sessions, although they were noted in probes for generalization. Future studies should further analyze the potential relationship between video-viewing and verbal behavior that appears to be maintained by automatic reinforcement.

More research needs to be conducted to determine if POVM is limited in its capacity to teach these or other behaviors (e.g., academic, other expressive skills, other social behavior). Future investigations

should determine the characteristics of children who would be considered best responders to the point-of-view procedure or skills best suited to point-of-view perspective in models. Additionally, a comparative analysis should test the relative efficacy of traditional video modeling and POVM.

Note

References

Alberto, P. A., Cihak, D. F., & Gama, R. I. (2005). Use of static picture prompts versus video modeling during simulation. *Research in Developmental Disabilities, 26*, 327–339.

Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional Children*, *73*, 264–287.

Bidwell, M. A., & Rehfeldt, R. A. (2004). Using video modeling to teach a domestic skill with an embedded social skill to adults with severe mental retardation. *Behavioral Interventions*, *19*, 263–274.

Charlop, M. H. & Milstein, J. P. (1989). Teaching autistic children conversational speech using video modeling. *Journal of Applied Behavior Analysis, 22*, 275–285.

Charlop-Christy, M. H., Le, L. & Freeman, K. A. (2000). A comparison of video modeling with in vivo modeling for teaching children with autism. *Journal of Autism and Developmental Disorders*, *30*, 537–552.

D'Ateno, P., Mangiapanello, K. & Taylor, B. A. (2003). Using video modeling to teach complex play sequences to a preschooler with autism. *Journal of Positive Behavior Interventions, 5*, 5–11.

DeLeon, I. G. & Iwata, B.A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*, 519–532.

Haring, T. G., Kennedy, C. H., Adams, M. J. & Pitts-Conway, V. (1987). Teaching generalization of purchasing skills across community settings to autistic youth using videotape modeling. *Journal of Applied Behavior Analysis, 20*, 89–96.

Hine, J. F. & Wolery, M. (2006). Using point-of-view modeling to teach play to preschoolers with autism. *Topics in Early Childhood Special Education, 26*, 83–93.

Horne, P. J. & Erjavec, M. (2007). Do infants show generalized imitation of gestures? *Journal of the Experimental Analysis of Behavior, 87*, 63–87.

Horner, R.D. & Keilitz, I. (1975). Training mentally retarded adolescents to brush their teeth. *Journal of Applied Behavior Analysis, 8*, 301–309.

Kinney, E. M., Vedora, J. & Stromer, R. (2003). Computer-generated video models to teach generative spelling to a child with an autism spectrum disorder. *Journal of Positive Behavior Interventions, 5*, 22–29.

Krantz, P. J., MacDuff, G. S., Wadstrom, O., & McClannahan, L. E. (1991). Using video with developmentally disabled learners. In P.W. Dowrick (Ed.) *Practical guide to using video in the behavioral science*s (pp. 256–266). Oxford, England: John Wiley & Sons.

LeBlanc, L. A., Coates, A. M., Daneshvar, S., Charlop-Christy, M. H., Morris, C. & Lancaster, B. M. (2003). Using video modeling and reinforcement to teach perspective-taking skills to children with autism. *Journal of Applied Behavior Analysis, 36*, 253–257.

Maione, L. & Mirenda, P. (2006). Effects of video modeling and video feedback on peer-directed social language skills of a child with autism. *Journal of Positive Behavior Interventions, 8*, 106–118.

Murzynski, N. T., & Bourret, J. C. (2007). Combining video modeling and least-to-most prompting for establishing response chains. *Behavioral Interventions, 22*, 147–152.

Nikopoulos, C. K. & Keenan, M. (2004). Effects of video modeling on social initiations by children with autism. *Journal of Applied Behavior Analysis, 37*, 93–96.

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The Childhood Autism Rating Scale (CARS)*. Los Angeles: Western Psychological Services.

Schreibman, L., Whalen, C., & Stahmer, A. C. (2000). The use of video priming to reduce disruptive transition behavior in children

with autism. *Journal of Positive Behavior Interventions, 2*, 3–11.

Shipley-Benamou, R., Lutzker, J. R., & Taubman, M. (2002). Teaching daily living skills to children with autism through instructional video modeling. *Journal of Positive Behavior Interventions, 4*, 165–175.

Wert, B. Y. & Neisworth, J. T. (2003). Effects of video self-modeling on spontaneous requesting in children with autism. *Journal of Positive Behavior Interventions, 5*, 30–34.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). , Fourth Edition (PLS-4). San Antonio, TX: Harcourt Assessment, Inc.