Comparison of the Effects of Video Modeling with Narration vs. Video Modeling on the Functional Skill Acquisition of Adolescents with Autism

Molly Smith and Kevin Ayres  
University of Georgia

Linda Mechling  
University of North Carolina Wilmington

Katie Smith  
University of Georgia

Abstract: The purpose of this study was to compare the effects of two forms of video modeling: video modeling that includes narration (VMN) and video models without narration (VM) on skill acquisition of four adolescent boys with a primary diagnosis of autism enrolled in an Extended School Year (ESY) summer program. An adapted alternating treatment design was used to compare the effectiveness and efficiency of both video modeling interventions with a baseline, comparison, replication, best treatment (when applicable), and maintenance condition (when applicable). Results indicated VMN was more efficient for two out of four participants. Narration was not a critical component of video modeling for the other two participants, although both indicated a preference for VMN over VM.

Researchers have suggested that many individuals with autism are strong visual learners and subsequently may benefit more from visually based rather than language based interventions (Quill, 1995; West, 2008). Video is one means of presenting instructional stimuli visually and researchers report that this is frequently an effective and efficient intervention to teach children and adolescents with autism spectrum disorders a range of skills (Ayres & Langone, 2007; Bellini & Akullian, 2007; Coy & Hermansen, 2007). One popular form of video based instruction is video modeling which involves a learner watching the performance of a target behavior and subsequently being instructed to imitate or perform the viewed behavior (Bellini & Akullian, 2007). Video modeling has been used to teach cooking (Mechling, Gast, & Seid, 2009; Mechling & Gustafon, 2008; Murzynski & Bourett, 2007; Shipley-Benamou, Lutzker, & Taubman, 2002; Sigafoos et al., 2005), vocational skills (Allen, Wallace, Renes, Bowen, & Burke, 2010; Cihak & Schrader, 2008; Van Laarhoven, Johnson, Van Laarhoven-Myers, Grider, & Grider, 2009), community skills (Alcantara, 1994; Ayres & Langone, 2007), social and play skills (Charlop, Dennis, Carpenter, & Greenberg, 2010; Tetreault & Lerman, 2010; Nikopoulos & Keenan, 2007); and leisure skills (Blum-Dimaya, Reeve, Reeve, & Hoch, 2010; Hammond, Whatley, Ayres, & Gast, 2010).

Researchers have used several means to present video models to learners including computer programs (Ayres & Langone, 2007), television (Shipley-Benamou et al., 2002), video games (Blum-Dimaya et al., 2010), and portable DVD players (Mechling, Gast, & Fields, 2008). Though Davies, Stock, & Wehmeyer began using Palm Pilots in 2002 to present still images and Mechling et al. (2009) used them to present video, the emergence of popular and pervasive mobile technologies like the iPod and iPhone increase the range of...
possibilities for delivering instruction. Further, because of the popularity of these devices for entertainment and communication among many adolescents and adults, they can be less stigmatizing than other prompting tools such as a laptop or television and they easily blend into the environment. Researchers are just beginning to facilitate instruction using these portable devices to teach functional skills to individuals with autism. Video models presented on an iPod have been used to teach four students with autism to transition appropriately (Cihak, Fahrenkrog, Ayres, & Smith, 2010) and as a prompting device in an employment setting for an adult with developmental disabilities (Van Laarhoven et al., 2009).

Although video modeling is an effective intervention, researchers have only begun exploring the various components of video modeling and what contributes to its effectiveness. Through the isolation of different variables, researchers will have a better idea of what contributes to the effectiveness of video modeling (Ayres & Langone, 2007). For example, studies have compared self-modeling (using footage of the student) and an adult or peer model (Cihak & Schrader, 2008) revealing that the procedures were equally effective and efficient in the acquisition of vocational skills. Ayres and Langone (2007) evaluated whether point of view modeling was more powerful than third person video modeling. They found no differences between the procedures; students learned efficiently with both procedures. Mechling and Bishop (2011) examined differences in the size of presentation of video models (screen sizes) and found individual variation on the number of approaches made towards each device with two of three participants approaching a larger screen more often than a smaller screen. Additionally, Mechling and Ayres (2012) found a higher percentage of correct responses with the larger screen than smaller screen, although both screen sizes resulted in increased performance for all participants. Through the isolation of various components of video modeling, researchers will pinpoint those that contribute to the success of the intervention, which will guide instruction in classroom settings.

The vast majority of studies using video models or video prompts have included narration or verbal cues overlaying the video model. These cues serve as verbal directions of each step of the chained task analysis. Rayner, Denholm, and Sigafoos (2009) raised the question of whether narration overlaying video models might enhance or possibly hinder the learning for students with autism. Only one study has isolated the verbal cue variable from video modeling in a comparative study (Mechling & Collins, 2012). The results of this study show that video models with narration were more effective than video models without narration for three out of four students with moderate intellectual disabilities. For one of the participants, narration was not an essential part of skill acquisition.

Because individuals with autism are strong visual learners, this topic should be further explored with this population. This study was a systematic replication of a study conducted by Mechling and Collins (2012) comparing the effectiveness and efficiency of narration on video models. Though the current study focused on adolescents with autism spectrum disorders instead of moderate intellectual disability, the purpose of the study was the same: to evaluate functional skill acquisition using video models with narration (VMN) compared to functional skill acquisition using video models (VM).

Method

Participants

Four students aged 13–16 years old with a diagnosis of autism were chosen to participate in the current study (see Table 1). All students attended a public middle school for Extended School Year (ESY) program. ESY was provided to students who either had a history of regressing on skill acquisition over the summers or were at critical places in skill acquisition when the summer break arrived. Participants had generalized gross and fine motor imitation and could physically complete all steps of the chosen functional skills. Each student had previous experience using computers and iPhones.

Grady transitioned to high school following the completion of the ESY program. Grady spoke in complete sentences, but often responded to questions with one-word answers.
His pragmatic language disorder impacted his social relationships in the classroom and across all settings. Grady had an extensive repertoire of functional and vocational sign sight word vocabulary, recited and wrote personal information, and engaged in many functional skills independently. Due to his constant scripting and distractibility, Grady needed frequent reminders to remain on task. Additionally, he was often unable to respond to group instructions and required individual instruction to complete a task. ESY was recommended for Grady to continue progressing on functional and academic skills throughout the summer.

Cole communicated using complete sentences and was often concerned with the emotions, thoughts, and feelings of others. He struggled with fluent speech during non-structured conversations, requiring frequent verbal cues to "slow down" or "breathe." When frustrated, he often said, "I’m sorry" in response to task directions. Cole initiated conversations with peers and adults, but was unable to carry on a conversation about topics other than his preferred interests. Cole responded well to structure and routine and instructional materials presented in a visual manner. Academically, he could decode grade level texts, but had trouble comprehending what was read. He was unable to follow simple written directions because of the lack of comprehension. Cole often engaged in noncompliance in response to changes in his daily schedule. He required individual teacher prompts to remain on task during group instruction. ESY was recommended for Cole to maintain skills previously mastered during the school year, as well as to target current goals and objectives.

Rawlins communicated using complete sentences and was exceptional in asking questions related to the current activity, although he often asked the same question more than once. Verbal praise from his peers or teachers was motivating for him. His speech was often unintelligible due to an articulation disorder...
which negatively impacted his oral communication in all settings. He often needed prompts to slow down or repeat what he had just said. Rawlins required supervision to complete functional and vocational tasks such as stacking serving trays and cleaning tables in the cafeteria and required frequent prompts to remain on task. Rawlins needed continued instruction in reading and comprehending vocational and functional sight words. He often experienced loss of skills over extended breaks from instruction, therefore ESY was recommended to maintain performance over the summer.

John transitioned from middle to high school following ESY. He communicated using complete sentences and responded well to verbal praise. John enjoyed writing and drawing as well as creating to-do lists. John provided his basic personal information (name, address, phone number, city, state, zip, mother’s name) in writing on a variety of job applications and forms. He independently completed simple cooking activities, such as using the microwave and making a grilled cheese sandwich. John read and comprehended texts two grade levels below his grade level. He engaged in self-stimulatory scripting behavior up to 40 times per school day. He struggled to maintain appropriate communication exchanges, often reverting to scripts to gain attention of his peers and adults. Due to the impending transition to high school, ESY was recommended to continue social and functional skill instruction.

Setting

All sessions of the study were conducted in the family consumer science classroom during the ESY program. The classroom measured approximately 30 ft by 40 ft and included a kitchen and living area. The living area consisted of desks and tables with cabinet space and a bulletin board. Sessions took place in the kitchen of the consumer science room which was furnished with a table, an island with a sink, two refrigerators, two sinks, a stove, cabinets, and ample counter space. The kitchen was stocked with food and cooking supplies.

Training Materials and Video Models

Materials were placed in the same area on the counter and in the refrigerator during all phases of the study. Distracters (i.e. different color spoons, different sized bowls, etc.) were mixed in with target materials on the counter to evaluate discrimination of the items used in the video models. Each of the three behavior sets (described in response definitions) included a specific set of materials. Set 1 consisted of salsa, streamers, a happy birthday sign, flowers, and a vase. Set 2 included a box of crackers, a 2-liter bottle of a soft drink, a glass, the board game Life, an envelope, and labels. Set 3 contained fruit punch, a pitcher, spoon, plates, cups, confetti, and the board game Twister.

Video models were presented using an iPhone 4 running iOS 4. Videos of an adult model performing the tasks were filmed in the classroom. An adult was chosen as the model since results of studies comparing self-modeling and using others as models have shown no difference in performance (Sherer et al., 2001). Narration was overlaid on each video for the VMN sessions. The volume was muted on videos presented in the VM condition. Videos included numbers indicating the start of the next step against a black background. For example, the number “1” would appear on the screen paired with the verbal cue. The video clip of each step followed immediately after the completion of the verbal cue. The narration did not overlap with the video depiction in order to focus attention on the critical components of both the verbal cue and the video model. Videos lasted from 24 seconds to one minute and twelve seconds.

Response Definitions and Data Collection

Target behaviors. Participants were screened on 18 behavior chains aimed at helping the students entertain friends or guests in their homes including preparing appetizers and drinks, decorating the room, and setting up an age appropriate game. This domain was selected as a target because all students had goals related to social interaction and it is age appropriate for adolescents to have parties and entertain friends. Targeting behaviors that prepare participants for age-appropriate
interactions with typical peers increases the social validity of the study. Each behavior chain consisted of a single task including between 3 and 8 steps. From the 18 behavior chains screened, 12 were selected to compare the effects of video modeling with and without narration (See Table 2). Four behavior chains were eliminated because all participants could complete at least 33% of steps. Two additional behavior chains were eliminated to make three equal behavior sets.

Following procedures outlined in Mechling and Collins (2012), the individual behaviors chains were divided into three behavioral sets that each received differing interventions (VMN, VM, and one served as a control set). The behavior sets were functionally independent to ensure learning one behavior did not result in an increase in accuracy in performance on another behavior. The behavior sets were also functionally equivalent, assessing similar skills such as preparing food items, putting out decorations, and setting up board games. Great effort was taken to ensure the behaviors sets were of equal difficulty. All behaviors were close to the same number of steps and assessed similar fine motor skills (i.e. twisting, pouring, and opening). Behaviors varied from three to eight steps each with the total number of steps for each behavioral set equaling 20 to 21 steps.

**Scoring responses.** The dependent variable was the percentage of independent correct responses exhibited during each set of behaviors. Correct and incorrect responses were recorded separately across all trials of each behavioral set. A correct response was defined as the participant initiating the task within five seconds of viewing the video, completing each step within 10 seconds of initiation, and com-

| TABLE 2 | Task Analyses for Behavior Sets |
|-----------------|-----------------|-----------------|-----------------|
| **Set 1** | **Set 2** | **Control Set** |
| **Put Salsa in bowl** | **Put crackers on platter** | **Make fruit punch** |
| 1. Select salsa | 1. Select crackers | 1. Select fruit punch |
| 2. Twist off cap | 2. Open box | 2. Open fruit punch |
| **Put streamer on wall** | **Pour drink in glass** | **Set out dinnerware** |
| 1. Select streamer | 1. Select drink | 1. Select stack of plates |
| 2. Tear off streamer | 2. Untwist cap | 2. Place on table |
| 3. Tear off tape | 3. Pour in glass | 3. Select stack of cups |
| 4. Tape one side to wall | 4. Stop | 4. Place on table |
| 5. Tear off tape | **Set up Life** | 5. Stop |
| 6. Tape other side to wall | 1. Select Life | **Put confetti on table** |
| 7. Stop | 2. Open box | 1. Select confetti |
| **Hang up party sign** | 3. Take out board | 2. Pour confetti on table |
| 1. Select party sign | 4. Open board | 3. Stop |
| 2. Tie one side up on wall | 5. Place spinner on board | **Set up Twister** |
| 3. Tie other side on wall | 6. Place car on board | 1. Select Twister |
| 4. Stop | 7. Stop | 2. Open box |
| **Put flowers in vase** | **Address Envelope** | 3. Lay mat on floor |
| 1. Select vase | 1. Select envelope | **Set up Twister** |
| 2. Put water in vase | 2. Select labels | 1. Select Twister |
| 3. Select flowers | 3. Peel off label | 2. Open box |
| 4. Put flowers in vase | 4. Place label on envelope | 3. Lay mat on floor |
| 5. Stop | 5. Stop | 4. Place spinner on floor |
| **Make fruit punch** | **Set out dinnerware** | **Put confetti on table** |
| 1. Select fruit punch | 1. Select stack of plates | 1. Select confetti |
| 2. Open fruit punch | 2. Place on table | 2. Pour confetti on table |
| 4. Put mix in pitcher | 4. Place on table | **Set up Twister** |
| 5. Put water in pitcher | 5. Stop | 1. Select Twister |
| 6. Select spoon | **Set up Twister** | 2. Open box |
| 7. Stir | 3. Lay mat on floor | 3. Lay mat on floor |
| 8. Stop | **Set up Twister** | 4. Place spinner on floor |

168 / Education and Training in Autism and Developmental Disabilities-June 2013
Completing the task within one minute from initiation of the first step. Incorrect responses were recorded if the student failed to initiate the response within five seconds, failed to complete an individual step within 10 seconds, or failed to complete the behavior within one minute. To receive credit for a response, participants had to discriminate between the correct appliance (refrigerator) and distractors as seen in the video model. Replicating Mechling and Collin’s (2012) procedures, single opportunity probes were used to assess performance for each behavior chain (Cooper, Heron, & Heward, 2007). An incorrect response was interrupted and the next instructor direction and video was provided. Efficiency of VM versus VMN was assessed through evaluation of the total trials to criterion and the total number of errors to criterion.

Experimental Design and General Procedure

An adapted alternating treatment design (AATD) (Sindelar, Rosenberg, & Wilson, 1985; Gast, 2010) with a baseline phase, comparison phase, replication phase, best treatment phase (when applicable), and maintenance phase (when applicable) was used to compare the effectiveness of VMN versus VM interventions on three sets of party preparation behavior sets. All participants were presented with the same control set of behaviors, which received neither intervention (until the replication phase), which intended to help detect history, maturation, and carryover threats. Interventions of VMN and VM were implemented to assess the performance of the remaining two behavior sets. Interventions were counterbalanced across participants. For example, two participants received VMN on Set 1 and VM on Set 2 and vice versa. Interventions were counterbalanced across sets to control for possible sequencing effects.

Best treatment was identified by calculating number of trials and errors to criterion. If the number of trials to criterion for VMN and VM was equal for a participant, the number of errors that occurred was evaluated to identify the best treatment. If the number of trials and errors were equal, the student was asked which instructional strategy he preferred (i.e. “Do you like the video to have sound?”). Criteria for the comparison phase were 90% correct across 2 sessions for one behavior set. Following completion of the comparison phase, the superior treatment was implemented with the control behavior set for each participant during the replication phase. After completion of the replication phase, the superior treatment was applied to any behavior set that did not reach criterion during the comparison phase in order to bring that behavior to criterion level. Including a Best Treatment condition evaluated for multitreatment effects to internal validity by assessing the alternate set of behavior with the superior treatment. On the final day of ESY (at least two weeks after completion of the comparison phase), a maintenance probe was conducted on the two sets of behaviors evaluated during the comparison phase.

Baseline procedure. A baseline phase was implemented to determine each participant’s performance without any intervention on each behavior set. Baseline sessions began with the presentation of task direction (e.g. “Put salsa in bowl”). The instructor provided non-descriptive verbal feedback for attention to task and correct responses on a CRF schedule. Once the task direction was provided, the student was given five seconds to initiate and 10 seconds to complete each step in the task analysis. Participants were given one minute to complete the entire task. Similar to Mechling and Collins (2012), an incorrect response was interrupted (i.e. “Stop”); the session for that behavioral chain was ended; and the next task direction was presented (single opportunity probe). Each behavioral set was assessed one time during each baseline trial, with the order of behavior sets alternating across all sessions to control for a sequencing threat to internal validity. All materials present during the comparison phase were present in order to make the baseline condition as naturalistic as possible. This ensured that all participants had to make conditional discriminations on all materials each time to make it as naturalistic as possible and promote discrimination. The baseline condition was concluded following a stable or decreasing trend across three sessions of each behavior set. Procedures used during the maintenance condition were identical to those used in the baseline condition.
Comparison of VMN versus VM. During VMN trials, the student was provided an iPhone with a video of the required behavior including verbal directions of each step. For example, the student was provided the task direction, “Put salsa in bowl.” Immediately following the task direction, the participant viewed a video of the behavior chain with the directions, “Select salsa, twist off top, pour salsa in bowl, stop.” The student was provided five seconds to initiate the response, 10 seconds to complete each step, and one minute to complete the behavior chain. Failing to initiate the behavior within five seconds, engaging in an incorrect behavior, failing to complete an individual step in 10 seconds, or failing to complete the behavior chain in one minute resulted in the end of a trial and presentation of the task direction and video of the next behavior chain in the set. As in the baseline condition, correct responses were reinforced on a CRF schedule until the participants performed 90% of the behavior independently. Then, reinforcement was faded to a VR-4 schedule of reinforcement (i.e. one time per behavior in each set). The same instructional procedures were used during the VM intervention. The iPhone was muted so the participants viewed the video without audio. For the control set of behaviors, the instructor provided a verbal direction of each behavior, as seen in baseline procedures. Behavioral sets of each intervention were varied across sessions and counterbalanced across participants.

Replication and best treatment phases. Procedures used during the replication phase and best treatment phase were identical to those implemented in the comparison phase as described above. Following the comparison phase, a replication phase was implemented in which the best treatment was applied to the control set for each participant. The replication phase involved the implementation of the intervention on the behavior set that remained in the no-treatment condition while the other two behavior sets received intervention. In order to determine the treatment for the replication phase, the researcher first looked at the effectiveness and efficiency data of VMN and VM. If one intervention was more effective and efficient than the other, this intervention was selected to be applied to the control behavior set. If the two interventions were equally effective and efficient, the participants were asked which intervention they preferred. The best treatment phase was only implemented for behavior sets that did not reach criteria in the comparison phase. The same treatment used for the replication phase was then applied to the behavior set for three days in an effort to boost performance to criteria levels. The replication condition for each participant was three sessions and the best treatment condition was three sessions due to time constraints. No specific criteria were used in these phases.

Reliability
Interobserver agreement was taken between 25%–33% of baseline sessions and 33% of comparison, replication, and best treatment phases for each student. Interobserver agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements of each step of all three sets of behaviors. Interobserver agreement was 100% across all participants, as behaviors were operationally defined and easily observable.

Procedural fidelity was taken at the same time as the reliability checks. The reliability data collector took data on the following researcher behaviors: a) all materials were presented; b) instructor provided task direction; c) instructor showed correct video model; d) instructor reinforced on correct schedule; e) instructor engaged in error correction procedures; f) instructor provided no additional prompts; and g) instructor provided the next video in the behavior set. Procedural fidelity was calculated by dividing the number of correct instructor behaviors by the total number of possible behaviors and multiplying by 100. Mean procedural fidelity was 98.5% (range = 84%–100%). Errors were from additional teacher prompts for one student who required repeating the task direction following the video for two trials.

Social Validity
Following the completion of the study, three teachers were asked to complete a survey taken from Mechling and Collins (2012). Teachers were asked whether video models
with narration or video models were easier to create, which type they would rather use, and if they viewed video models as an effective intervention for classroom instruction. The teachers were asked one additional question: what types of skills would they target with video modeling? In addition to the teacher questionnaire, the researcher asked the four participants if they liked using the video models, if they wanted to learn other skills using video modeling and whether they preferred the video models with or without narration.

Results

Comparison Phase

Three out of four participants reached 100% accuracy using VMN and VM the same day, with one participant reaching 100% accuracy with VMN and over 80% with VM (see Figures 1, 2, 3, and 4). VMN resulted in an immediate increase in performance for John, who reached 100% accuracy on the first session. The comparable drastic increase in correct responses in both VMN and VM makes a superior treatment indiscernible. When VMN was introduced in the comparison phase, Rawlins’ performance when provided VMN increased at a more rapid pace while VM demonstrated a gradual change. Rawlins reached 100% accuracy after the same number of sessions for both sets of behaviors. However, he made five more errors when presented with VM than VMN, indicating the latter intervention as more efficient.

Initially when presented with both types of video models under the comparison phase, Grady’s performance did not change abruptly. He continued to complete the first step of each behavior in each behavior chain as reflected in his baseline performance. Video correction involves presenting the video clip again once an error has occurred during video modeling or video prompting procedures and has been used to teach daily living skills to adults with disabilities (Cannella-Malone et al., 2006; Goodson, Sigafoos, O’Reilly, Cannella, & Lancioni, 2007). Following a zero celerating trend across three stable data points, a video correction trial was introduced in which Grady was presented the step of the video where an error occurred and the prompt, “No, Grady. Do what you see in the video.” After one video correction trial, Grady’s performance jumped to 100% accuracy across both sets of behaviors. This indicated

![Figure 1. Percentage of correct steps performed on three behavior sets for Cole in baseline, comparison of VMN and VM, replication, and best treatment phases.](image-url)
that Grady did not understand that he was supposed to actually complete each step. Once he was corrected one time, he completed each step of both sets accurately. Cole’s performance increased abruptly when provided VMN intervention. He required an extra prompt during the first session of both videos presented using VMN and VM to complete the behavior (i.e. “It’s your turn.”). The first presentation of VM intervention resulted in near-baseline performance as Cole was perseverating on the numbers in each video.

Figure 2. Percentage of correct steps performed on three behavior sets for Grady in baseline, comparison of VMN and VM, replication, and maintenance phases.

Figure 3. Percentage of correct steps performed on three behavior sets for John in baseline, comparison of VMN and VM, replication and maintenance phases.
However, during the second session his accuracy jumped to 50%. It took Cole five sessions to reach 100% criterion with the narration and he responded above 80% criterion after five trials of VM training. Cole also made more errors with VM than with VMN. Fewer errors made and fewer trials to reach criterion deemed VMN intervention as the superior treatment for Cole.

Replication Phase

Cole responded with 100% accuracy with narration and 81% accuracy without narration. VMN intervention was selected as the best treatment for Cole. The remaining three participants reached 100% accuracy on both sets of behaviors on the same day; therefore, number of errors was compared to determine best treatment. Rawlins made five more errors when presented VM intervention, therefore VMN was chosen as the superior treatment. Grady and John made the same number of errors across both interventions, so they were asked which intervention they preferred. Both participants selected video modeling with narration (see Table 3).

All participants reached over 76% accuracy

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Determining Best Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>Trials to Criterion</td>
</tr>
<tr>
<td>Cole</td>
<td>VM-Required Best Tx</td>
</tr>
<tr>
<td></td>
<td>VMN-5</td>
</tr>
<tr>
<td>Grady</td>
<td>VM-6</td>
</tr>
<tr>
<td></td>
<td>VMN-6</td>
</tr>
<tr>
<td>John</td>
<td>VM-1</td>
</tr>
<tr>
<td></td>
<td>VMN-1</td>
</tr>
<tr>
<td>Rawlins</td>
<td>VM-4</td>
</tr>
<tr>
<td></td>
<td>VMN-4</td>
</tr>
</tbody>
</table>

Figure 4. Percentage of correct steps performed on three behavior sets for Rawlins in baseline, comparison of VMN and VM, replication, and maintenance phases.
on their control set within three trials when VMN was applied. John and Grady reached 100% accuracy when presented the narration. Grady’s performance decreased to 86% during the last trial as a result of a sequential error. Fifty five percent of the errors that occurred during the replication phase were sequential errors.

**Best Treatment Phase**

Because three of the four participants reached criteria for behavior sets taught with VM and VMN in the same sessions, a best treatment condition was not necessary. VMN was chosen as the best treatment for Cole and was applied to the behavior set originally instructed with VM. Procedures identical to those used in the comparison phase were implemented on Behavior Set 2 during the Best Treatment Phase. Cole responded with 86% accuracy during the comparison phase using VM on Set 2. Once narration was added, Cole immediately responded with 100% accuracy. This immediate increase strengthens the internal validity of video modeling with narration as the superior intervention for Cole. His performance dropped to 62% on the second trial due to visitors arriving in the room and subsequent noncompliance during the session. His performance jumped back up to 100% accuracy again during the final session.

**Maintenance Phase**

On the final day of ESY and 15–20 days after treatment ended, a maintenance probe was conducted to assess performance of both sets over time for John, Grady, and Rawlins. John and Grady responded with 100% accuracy for Set 1 and Set 2 when assessed for maintenance. Rawlins also responded with 100% on Set 1 which received VMN intervention. However, Rawlins’s performance on Set 2 (VM) decreased drastically from 100% during the comparison phase to 48% during the maintenance probe. Cole reached criterion for the best treatment phase one day prior to the completion of ESY; therefore a maintenance probe was not conducted.

**Social Validity**

All four participants indicated they enjoyed learning new skills using the video models and they would like to learn more skills using video modeling presented on an iPhone. When verbally asked, each participant stated a preference for the video models with narration over video models. The three teachers who participated in the survey thought it would be easier to develop video models over video models with narration, but they would rather use VMN over VM. All three teachers felt video modeling would be an effective tool to use for instruction in their classrooms. The teachers indicated they would use video models to teach other complex chained tasks such as daily living, self-care, community skills and social skills.

**Discussion**

The purpose of this study was to compare the effectiveness and efficiency of video models with narration and video models without narration on skill acquisition of four adolescent boys with autism. Individuals with autism are generally noted to be strong visual learners, and visually cued instruction for individuals with autism has become widely used (Quill, 1995). Although individuals with autism may learn better from visual stimuli, this study attempted to assess if audio narration is a critical component to video modeling or if video modeling alone, because of the visual learning strength for individuals with autism, would be equally effective.

All participants responded with over 80% accuracy following the comparison phase with both types of video models, increasing the external validity of video modeling as an intervention to teach complex chained tasks. The effectiveness and efficiency of each type was determined by analyzing the percentage of correct responses in each behavior set, the number of trials to reach criterion, and the number of errors that occurred. John and Grady reached criterion for both sets after the same number of trials and made the same amount of errors, making a superior treatment indiscernible based on data. John responded with 100% accuracy on the first trial of both VMN and VM during the comparison
phase, indicating the skills may have been too easy for him or that he may have already been able to complete the steps of the task but did not understand the task direction. After one video correction trial, Grady’s performance jumped to 100% accuracy across both behavior sets. This indicated that Grady did not understand that he was supposed to actually complete each step. Video modeling training prior to implementation of the study could have prevented Grady’s errors and provided a better representation of the more effective type of video modeling.

Although Rawlins reached criterion for both types of video models after the same number of trials, he made more errors during the VM intervention than in the VMN. Because of the number of errors, VMN was the more efficient intervention in acquiring functional skills for Rawlins. Also, Rawlins maintained over 50% more of the skills taught using the VMN over VM when assessed for maintenance. Cole did not reach 100% accuracy on the behavior set taught with VM during the comparison phase and only reached 100% accuracy during the Best Treatment Phase, indicating VMN as a more effective intervention for Cole.

One interesting development of the study was an increasing trend of sequential errors once criterion had been met. During the replication phase, 55% of the errors that occurred were sequential errors. Once the students reached a level of fluency with the sets, they often found an easier and more efficient way to complete the steps, resulting in more sequential errors. For example, during the first few trials of the replication phase, Grady correctly completed all five steps of “setting out dinnerware.” This included selecting the stack of plates, placing it on the table, selecting the stack of cups, placing it on the table, and stopping. After two trials of 100% accuracy, Grady engaged in a sequential error since he picked up both the cups and plates at the same time. Grady knew the final outcome of the behavior and found a more efficient way to achieve the outcome. An increase in sequential errors could have been prevented had the instructor chosen to use multiple opportunity probes, indicating the use of single opportunity probes as a limitation of the study. Stopping the student once an error had occurred did not give the most accurate depiction of student performance. Using multiple opportunity probes would be a more accurate measurement of performance and would have controlled more for these sequential errors.

Limitations

When evaluating the effectiveness of two interventions using an adapted alternating treatment design, behaviors assessed must be of equal difficulty (Gast, 2010). Identifying 18 party preparation skills that were equally difficult was a struggle. Participants were probed on the behaviors identified and through close analysis of their performance as a group, 12 behavior chains were chosen to evaluate VMN versus VM interventions. Each behavior set had a similar number of skills, similar sets of behavior chains, and the complexity and similarity of discriminations required were consistent. Three out of four participants completed three or less steps correctly across all sets, indicating low levels of baseline performance for all participants on all behavior chains. On the other hand, John performed one behavior (addressing the envelope) independently prior to receiving any intervention. This was a limitation of the study that affected the validity of the findings, as John was already performing 29% of one behavior set prior to beginning the comparison phase. In the future, more strategic selection of behaviors, such as a larger pool of behaviors to select from, may result in behaviors being closer to equal difficulty.

An additional limitation of the study included the time constraints due to data being collected during ESY. Because of the time constraints, behavior sets originally in the control condition were only provided intervention for three days. For three out of four participants, this behavior set did not reach criteria. The researcher was also unable to collect maintenance data on Cole.

Implications for Practice

The VMN intervention proved more efficient for two out of four participants while narration was not a critical component of video modeling for the remaining participants. All
four participants indicated enjoying learning new skills using the video models on the iPhone 4, while all three teachers who participated in the social validity survey indicated a preference for VMN over VM. Teachers can implement video technology within their classrooms to teach a variety of functional skills such as pre-vocational skills, cooking skills, and leisure skills. Because all teachers indicated an interest in using VMN and participants enjoyed learning skills with video modeling, more research should be conducted to determine the critical components of video models. Regardless of how video is presented to students, it is likely worth the teacher or interventionist’s efforts to incorporate concise verbal directions with the video. Ultimately, if stimulus control can be transferred to audio cues, a student could work through a task wearing earphones and with a device like a smartphone in their pocket providing them the necessary information. This would allow a hands free approach to task completion. Further, if stimulus control were transferred to audio stimuli (instead of the video model), this would allow teachers and other support staff greater flexibility in being able to provide verbal cues as needed. Coupled with this, students may begin to acquire some of the associated vocabulary. For students who find the audio distracting or distasteful, they can always be taught to mute the audio narration.

Implications for Research

Future researchers should identify behaviors that are independent, functionally equivalent, and of equal difficulty to evaluate the effectiveness of VM and VMN. Due to the small sample size of the current study, more research should be conducted evaluating the effectiveness of the interventions with a greater number of participants with autism to increase the external validity of the findings. An additional research area includes the importance of narration for different types of tasks. Because each of the behavior chains in the current study is between 4–8 steps, it is important to evaluate these procedures with longer, more complex tasks. Additionally, researchers should look into the importance of narration for other age groups of individuals with autism including both younger children and adults with autism. Also, for individuals with autism who display scripting behaviors, narration could be beneficial for remembering tasks, or could be hindering if the individual perseverates on the narration. Future research needs to look at individual variation within individuals with autism and the effectiveness of VMN and VM.

Lastly, because VMN and VM were equally effective for three out of four participants, it is important that all future research that seeks to isolate important components of video modeling include social validity measures. All participants in the current study stated a preference for VMN.

References


---

Received: 5 January 2012
Initial Acceptance: 11 March 2012
Final Acceptance: 9 July 2012