5-6-2021

Effects of first- and third-person point of view on the acquisition of behaviors using video modeling

Robert Harper III

James Madison University

Effects of First- and Third-Person Point of View on the Acquisition of Behaviors Using Video Modeling

Robert Harper III

A thesis submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Education

Department of Education

May 2021

FACULTY COMMITTEE:

Committee Chair: Dr. Sara Snyder

Committee Members/ Readers:

Dr. Laura Desportes

Dr. Benjamin Riden

Table of Contents

List of Figures

Abstract

Video modeling is an evidence-based practice for teaching behaviors and chains of behaviors to children with Autism Spectrum Disorder (ASD). The purpose of this study was to determine whether the viewing perspective of these models played a role in influencing a learner's acquisition of the target behavior or behaviors. An adapted alternating treatments design was used to examine the effects of these different perspectives affected the learning of two similar behavioral chains in a learner with ASD. Video models from both viewing perspectives were provided to the learner with no additional prompting other than brief verbal acknowledgement of a step's completion.

Keywords: in vivo/live modeling, first-person perspective/POV modeling, model, third-person perspective/scene modeling, video modeling

Introduction

Imitation skills are a foundation block of learning and allow for the quick acquisition of skills and behaviors. Imitative skill, however, is an area of deficit for some children with autism and developmental disabilities (Cooper et al., 2020). Imitation is when a target skill or action is modeled, then recreated with formal similarity by an observer (Cooper et al*.,* 2020). Targeted imitation training and modeling have been used to teach a variety of skills to children with autism including functional tasks, conversation skills, expressing emotions, and play skills amongst others (Charlop-Christy et al., 2000; Charlop et al., 1983; Delano, 2007; Kleeberger & Mirenda, 2010). The ability of a child with autism to effectively follow modeled behavior is important to their ability to be successful in a mainstream setting or natural environment (Charlop-Christy et al., 2000).

Initially in research, live or "in vivo" modeling was the most common method of task demonstration used with learners (Kleeberger & Mirenda, 2010). As this method of teaching has evolved, the use of video has been incorporated into its use. Video modeling still consists of the demonstration of a target skill or behavior through demonstration to the learner, but the demonstration is delivered via a pre-recorded video rather than live and in the moment (Charlop-Christy et al., 2003; Delano, 2007). Video modeling does offer several potential advantages, such as easy repeatability and consistency of the model across trials (Goldsmith & LeBlanc, 2004). Video modeling is often combined with some method of in vivo instruction, particularly prompting or discrete trial training (DTT) techniques (Kleeberger & Mirenda, 2010).

Video modeling as a method of instruction has been utilized going back as early as the 1980s (Delano, 2007). A study conducted by Charlop-Christy et al. (2000) showed that when comparing the effectiveness of video and in vivo modeling with five subjects, video modeling was generally a more effective and efficient teaching tool than live models. Of the five subjects, four of them acquired the target skill in less time with a video model, and the fifth acquired the skill at the same rate with both methodologies. The same study also noted that related to the difference in time, video modeling was a more cost-effective method of instruction for the learners in this study.

In vivo modeling provides the opportunity to utilize either peer or adult models (Charlop et al., 1983). Video modeling can also make use of an adult or peer model but adds the additional option of changing the perspective of the task presentation (Spriggs et al., 2016; Cotter, 2010). By the nature of observing the action performed, in vivo modeling is viewed by the learner from a 3rd person or “scene” perspective (Cotter, 2010). When using a video to present the model, however, it is possible for the recorded model to mirror a first-person perspective, where the presentation of the model is more similar visually to what the learner will see when performing the task (Cotter, 2010). In a study teaching functional skills to preschool aged children with autism, the researchers utilized a first-person (in this case “over the shoulder”) video model that was not only effective in teaching each of the target skills to the subjects, but that the skills maintained at mastery criteria in a follow up a month post instruction (Shipley-Benamou et al., 2002).

Statement of the Problem

The ability to imitate actions and sequences of behaviors is a foundational skill that influences a person’s ability to acquire new skills throughout their life. Research shows that imitation training and modeling are effective abilities to teach children with autism and developmental disabilities (Charlop et al., 1983; Quill, 1997; Spriggs et al., 2016). Additionally, there is a large amount of research indicating that the delivery of these models through a video is effective in teaching the target skill (Charlop-Christy et al., 2000; Delano, 2007; Goldsmith & LeBlanc, 2004; Kleeberger & Miranda, 2010; Spriggs et al., 2016). Some of this research even indicates that video modeling is a more effective method of instruction than in vivo presentations (Charlop-Christy et al., 2000). Video modeling interventions are more accessible than ever with the proliferation of high-quality video technology readily available, making it a realistic avenue of instruction in most cases. Although in recent years there has been some research comparing the effectiveness of first- and third-person video models, there is a lack of replication present to lend credence to these results.

Purpose of the Study

The purpose of this study is to examine the effect of differing viewpoints when teaching skills using a video modeling presentation. Specifically, the study will examine these questions:

1. Will the use of first- versus third-person video modeling (i.e., point of view vs. scene) influence the speed at which a learner acquires a skill?
2. In a follow up probe coming after the discontinuation of instruction, will the first- or third-person model have an effect on the durability of the skill learned?

This study adds to the existing literature on the use of video modeling by further examining an area with a relative paucity of research. If one method of instruction is indicated to be more effective, this is an area that future treatment could utilize when implementing video models.

Literature Review

This literature review begins with a description of the key terms that will recur throughout the study. Following this, the theoretical framework of the research will be examined. Then, the review will conclude with a description of the different video modeling methodologies that are commonly used in existent studies. The sources utilized in this review range from textbooks, to dissertation, and peer-reviewed articles.

Definition of Terms

1. *In Vivo/Live Modeling* - A visual demonstration of the target action or sequence of actions presented in the immediate physical vicinity and view of the learner
2. *First-Person Perspective/POV Modeling* – A model in which the video field of view corresponds with that of the learner
3. *Model* – "[A presentation] that shows, demonstrates, or conveys exactly the behavior the learner is expected to perform." (Cooper et al., pg. 533, 2020)
4. *Third-Person Perspective/Scene Modeling* – A model in which the learner is viewing another person complete the action. Can be live modeled or video modeled.
5. *Video Modeling* – "A technique that involves demonstration of desired behaviors though video representation of the behavior." (Bellini & Akullian, 2007)

Theoretical Framework

In describing the use of modeling, Cooper et al. (2020) note that modeling can consist of live actions conducted in the moment, or "symbolic" in nature. Live modeling would be the way that imitation is generally taught: a learner is cued to complete an action or task following someone completing it immediately before them. Symbolic models can be anything ranging from pictures, to videos to a mixed media presentation. Modeling, both with in vivo and video presentations are evidence-based practices for the instruction of new skills to children with autism, who may otherwise struggle to learn observationally (Spriggs et al., 2016). Charlop-Christy et al. (2000) demonstrated that the use of video models may in fact be more effective in teaching children with autism than in vivo models.

The specific use of video modeling has been used as early as the 1960s in the “Bobo Doll” experiments headed by Albert Bandura (Bandura, 1965; Bandura et al., 1963). In these experiments, children showed approximations and replications of the modeled behaviors when placed in similar scenarios (Bandura, 1965; Bandura et al., 1963). Specifically, one experiment found that a visible reward or punishing contingency was effective in altering the learner’s behavior (Bandura, 1965).

Throughout the Bobo Doll experiments, Bandura was able to demonstrate the effectiveness of modeling, specifically through a video presentation. However, each of these studies used typically developing children and not children with autism. Charlop et al. (1983) were able to demonstrate the effectiveness of modeling in teaching skills to children with autism using peer modeling, displaying its utility in this area. Others would go on to display video modeling using peers as an effective means to teach children with autism, as well subjects with severe intellectual disabilities (Cannella-Malone et al., 2011; Delano, 2007). Others were successful in teaching children with autism using video models containing adult models, models filmed in a first-person point of view, or even self-modeling (Delano, 2007; Goldsmith & LeBlanc, 2004).

It is important to note that not all experiments have been effective in solely using video models to train behavior (Bandura, 1965; Bandura et al., 1963; Delano, 2007). In the Bandura experiments, most subjects were not successful in recreating the entirety of the modeled behavior or were not able to recreate the model with accuracy. Delano (2007) specifically notes a study that was not successful in utilizing video models until the delivery of outside reinforcement and a self-management procedure were implemented. Their results indicated that without the additional supports and procedures, that video modeling alone would not have been entirely effective in this case.

Video Modeling as a Method of Instruction

As presented above, the use of video models builds upon the existing research showing modeling as an effective method of instruction. In traditional live modeling, the learner experiences the task performed in a third-person perspective, simply by virtue of being an observer to the task or action. By making use of video modeling, there are multiple new avenues that can be explored both in making use of a variety of models, but also in the presentation perspectives. Both peer and adult modeling as described below are from scene modeling, or third-person perspective. These are by far the most popular methodologies used in existing research, when combined accounting for almost six times as many studies as those conducted from a first-person point of view (Cotter, 2010).

Peer Modeling

The use of peers to model behavior dates back to before video modeling was a popular methodology of instruction. Mentioned previously, Charlop et al. (1983) used peer models to effectively teach children with autism to receptively label a target. In this study, they compared the effectiveness of a modeled condition and a trial-and-error procedure to train the subjects to label the presented stimuli. Using this method, the researchers found that in these learners, both trial-and-error and peer modeling could be used to eventually train to criteria. However, the modeled condition led to better generalization in a follow up session, and better maintenance in the trained targets that those taught via trial-and-error.

Peer modeling provides the benefit of the model likely looking more similar to the subject than when adult models are used. Per Cooper et al. (2020) "the similarity between the individual providing the model and the learner can influence the likelihood that imitative behavior will occur" (pg. 533). Bellini & Akullian (2007) echoed this, noting that there is an increased likelihood that learners will attend to a peer model. On the contrary, Charlop-Christy et al. (2000) argued that previous studies did not show an appreciable difference in both in vivo and video modeled presentations when comparing adult to peer models. In the studies analyzed by Delano (2007), five made use of adult models, while seven made use of peers, all of which were effective to varying extents in the acquisition of the target skill.

Although peer modeling does appear to be effective in teaching various skills to subjects with autism, this methodology does present a number of barriers to its application. While using a peer does offer the benefit of formal similarity as referenced by Cooper et al. (2020), finding a suitable peer that is similar in appearance may prove to be challenging. In addition, this will likely require additional parental consent to be obtained, as well as the time necessary to train the peer to correctly complete the task (Shipley-Benamou, et al., 2002). Although for simple tasks that may already be existent in the peer's repertoire and could be easy to capture on video, it may be necessary to train them to complete the task before video may be taken. This stands in comparison to the relative ease of using adult models, who would require less training and permissions to be obtained (Bellini & Akullian, 2007).

Related to peer modeling in the third-person perspective viewpoint is video self-modeling (VSM). VSM is a specific presentation of a third-person video model in which the learner observes a video of themselves completing the target task correctly (Cooper et al., 2020). Obviously, this does satisfy the potential benefit from the model being similar in appearance to the learner but presents additional difficulties. The first and foremost challenge to this approach is that it requires the learner to be able to complete the task with some independence for the video to be made. There is evidence to support this intervention in maintaining learned skills over time but can be challenging for learners that lack the skill to complete the action even once for the model to be made.

Adult Modeling

Adults have also shown to be effective models when using a video presentation (Bellini & Akullian, 2007; Charlop-Christy, et al., 2000; Delano, 2007). Stretching back to the Bobo Doll experiments conducted by Bandura and his team across their research, the use of an adult model on the screen did lead to the learner's behavior acquisition (Bandura, 1965; Bandura, et al., 1963). In her dissertation, Cotter (2010) was able to teach a variety of skills to the study subjects using adult models both in first- and third- person perspectives. In their study comparing the efficacy of in vivo and video modeling, Charlop-Christy et al. (2000) made use of adult models in both conditions. Again, in both conditions no instruction nor reinforcement were provided to the learners outside of cues and praise for attending to the model. A study teaching learners how to access video games also made use of an adult model. This study also used observation of peers as an instruction condition but found the video models to lead to quicker mastery in most subjects (Spriggs, et al., 2016).

Adult models were also utilized by Kleeberger & Mirenda (2010) when teaching imitated actions in play or songs. Where the previously mentioned study was effective in teaching the target skills with just a video model, Kleeberger & Mirenda required greater prompts to train the imitative skill. Here, additional conditions such as highlighting (calling attention to the action) and highlighting with positive reinforcement were necessary to train the response. After these steps, the skills did generally maintain at that level in an additional video model only condition. In this study it should be noted that the target skills in the video model included an adult roleplaying the teacher as they did in the teaching trials, but with two additional adults roleplaying the student and acting out the target response. It is possible that the presence of so many actors in the model could be distracting to the learner, even if all were completing the action.

For the previously mentioned studies, the target action was either a brief, imitated action or a short exchange of words (Charlop-Christy, et al., 2000; Kleeberger & Mirenda, 2010). In another study using an adult model, the steps necessary to turn on, start, play and conclude 4 different video games were taught using video modeling (Spriggs, et al., 2016). In this study, both a video model of the chained target behavior was used in instruction, as well as observation of a peer in another condition. This study found the video model condition to be effective in teaching the skills to the learner across multiple learners and target chains. No additional reinforcers other than access to the game were provided to the learners in this study, but additional prompts in the form of narrating the task analysis were present in the video (Spriggs et al., 2016).

First-Person POV Modeling

Perhaps the most interesting method of presenting a video model is from a first-person perspective. In these models, the camera is positioned in such a way that the visible field is very similar to how a person would perceive the world naturally. Here, the benefit to the learner is that the field of the view in the video is more or less the same as what they will experience when learning the task. In most instances, only the model’s hands will be visible in the video.

Utilizing a multiple probe design, one study used first-person video models to teach activities of daily living (ADLs) to two different participants (Aldi et al., 2016). Both participants acquired the targeted ADLs. Within the video models was narration providing additional verbal prompting to the visuals of the video. Participants maintained at least 50% of the steps, often more, in post-intervention maintenance probes. It is also notable that the models in these interventions were family members of the subjects, indicating that similar interventions could be implemented without the practitioner being physically present, provided they could coach the family in the creation of these videos.

Shipley-Benamou et al. (2002) also utilized a multiple probe design to teach functional skills to multiple subjects. Unlike the research done by Aldi et al. (2016) however, no additional prompts were included in the video for each step of the task. At the start of each model a cue was given to “watch the friend” on the screen but following this no additional verbal prompts were given. All subjects demonstrated growth in each of the target skills in the intervention phase. Additionally, the learned skills showed strong maintenance in probes conducted following the conclusion of the video modeling condition. The authors of the paper note that the existing skillsets of the learners were “significant features” (p. 174) in the results. While none of the subjects showed great performance in baseline, all three had existing imitative repertoires as well as the ability to follow one and two step instructions.

Where the other discussed studies have all been specific to learners diagnosed with ASD, Cannella-Malone et al.’s (2011) study also utilized a first-person perspective in modeling, but in teaching to seven subjects diagnosed with severe intellectual disabilities. In addition, one of the subjects was deaf, and unable to respond to any verbal instructions or cues delivered in the intervention. This study used a combination of a multiple probe and alternating treatments design. Rather than just using a video model, the researchers here alternated the delivery by using both video models and “video prompts.” In this study, video models refer to videos containing the entirety of the behavior being taught. Video prompts, however, are the same videos as those of the models, but broken down into individual clips corresponding to the task analysis. In doing this, the researchers were able to deliver more targeted modeling to the individual steps of the chain, rather than a total task presentation.

Although there are a multitude of studies that make use of a first-person perspective in presenting the video, only one published piece directly examined the effects of first-person modeling with third-person modeling (Cotter, 2010). The researcher measured the rate of acquisition across multiple subjects and targets in these areas, but also in attending to the actual model. Using an adapted alternating treatments design, this study utilized point of view and scene models combined with some variations to assess the acquisition and attending. Notably, the researcher did not find a strong correlation between increased learning speed and video perspective, though they did note slightly more positive results with first-person perspective in some subjects. In fact, in this study there was difficulty in even completing the tasks at all, regardless of perspective. Only one participant was successful in acquiring the skills using video models alone, and the remaining subjects required additional instructional techniques such as narration in video, error correction, and in vivo modeling.

Research Gap

Many of the mentioned studies show that video modeling is an effective means of teaching a wide scope of target skills to learners, especially when paired with additional methods of instruction and reinforcement. The existing literature provides multiple examples of this but has a relative paucity of examining the most effective means of implementing a video model. Only one source found in doing research for this study examined the difference that first-and third-person perspectives may cause in the acquisition of video modeled skills. This study will add to that area of inquiry in the existing video modeling literature.

Method

This section discusses the method and design implemented to examine participants' skill acquisition when presented with first- and third-person perspective video models. The following section includes a description of the sample, data collection instruments and procedures, and the design of the study. In discussing the design of the study, the topics of both internal and external validity, reliability, generalizability, and limitations will be covered as well.

Participants, Selection Procedures, and Setting

The selection procedures initially laid out for this study were as follows: individual diagnosed with Autism Spectrum Disorder (ASD), able to follow at least some multiple step sequences of actions without reliance on prompts, and strong enough motor skills to navigate their environment. Selection and exclusion procedures for this study were left intentionally broad for a number of reasons. First, the research base reviewed for this study contained subjects ranging in age from early childhood to teenagers. This wide range of subjects in the existing literature indicates that success in following video models is not reliant on a specific age. Secondly, as this was a single subject study, the meaningful comparisons drawn will be against their own performance in the other phase of the intervention, not against an outside population. This left the only area of need being pinpointing two functional skills of similar length and complexity to target in the different phases of the intervention.

The selected subject was a child receiving Applied Behavior Analysis (ABA) services through a home and clinic-based agency. It should be noted at this point that the primary researcher worked at the agency providing services to the client and was temporarily added to their treatment team for the purpose of this intervention. Prior to their receiving services with the agency, the child was diagnosed with ASD. At the time of intake, parent interviews, the Autism Treatment Evaluation Checklist (ATEC), and the Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP) were completed by members of their treatment team. Through these assessments, the researcher developed an inventory of potential skill chains to target in the intervention. Additionally, the research sought input from the subject's family regarding skills the client needed for success in the home setting.

The selected participant was a 4-year-old girl of Southeast Asian descent, diagnosed with ASD (Jess). Jess lives at home with two parents and a younger brother. Jess's most recent Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP) score was a 74.0, completed in March of 2021. It showed primarily that while Jess does well particularly in areas pertaining to play, requesting, and receptive language, she still shows deficits in other areas of expressive language. Jess was also shown to have appropriate imitation skills for a child her age. In addition, Jess has a history of struggling to remain seated in a work area for more than seconds at a time but had recently met a goal of working in an assigned area on an adult directed task for 5 minutes. Despite meeting this objective prior to the selection process, Jess did show regression in the days immediately prior to the intervention.

All sessions were completed in the agency's clinic. All clinic sessions were carried out in the same room, either sitting or standing at the table. The room was an observation room with a one-way mirror on one wall and was about 5 feet by 9 feet inside. In the clinic, all sessions were recorded on a small camera that was positioned either from an overhead view or head-on at an upper angle. All videos of trials completed by the learner were saved on the secure server used by the agency providing services, however videos were labeled and stored under a pseudonym as a precaution to protect the subject's identity. All video modeled prompts were presented on the same tablet regardless of target or video modeling perspective. This tablet was not the client's and was not available to them for any activities other than the prompting videos to reduce the potential of any other distractions occurring.

Materials

The materials and reinforcers used were all obtained from the subject's natural environment or selected to specifically mimic stimuli present in the subject's natural environment. The video models were presented on a tablet device, which was not available to the subject other than during the presentation of a model. The videos were presented on a 3rd generation iPad Air, which has a 10.5-inch diagonal screen. The behaviors modeled in each of the seven total videos were all completed by the same person, wearing the same clothes. This was done to maintain formal similarity across all models and reduce any potentially confounding affects. For each instance of a model being presented, the tablet was operated by an adult present in the setting, then removed as the subject began to initiate the behavior chain. Reinforcers were made available to the subject at the conclusion of each attempt.

Description of Data Collection Instruments

Task analyses for the behavioral chains were created prior to the filming of the video model. Data were collected on the subject based on their adherence to the steps modeled in the video. Responses were recorded as either a '+' for a correct step completion or a '-' for failing to complete the step as modeled. Failing to complete the step as modeled could entail either an incorrect behavior or attempt, or even a failure to initiate any attempt. From this, percentage data were calculated by dividing the total number of correct steps by the total steps in the task analysis. If any major problem behavior events or potential barriers to future success were observed in the sessions, these were notated using duration recording and a narrative description of the behavior observed. Additionally, due to an existing program with this client, any appropriate mands for a break or to conclude the activity were met with an immediate stop to the trial, and the same short break that followed other attempts was given. If the learner had begun to initiate the action, data collection was stopped at the point the attempt ended. If the mand came prior to initiating the action, this was noted, and the trial was omitted entirely from the final percentage data. This was done in the event that any part of the intervention required alteration due to danger to the client or others.

Description of the Research Design

An adapted alternating treatments design was utilized for its utility in comparing the acquisition of two separate skills concurrently, while ruling out the potential confounding effects of maturation (Ledford and Gast, 2018). Baseline and intervention data were collected within the same instructional session, with one hour in between to provide space between trials. Additionally, the order of instructional trials was randomized using random.org before the start of each session to avoid any potential effects from maintaining the same sequence. A post intervention maintenance probe was planned for each target behavior chain two weeks following the discontinuation of the intervention. This was to be done to gauge if one skill maintained with greater accuracy following a break in instruction (Ledford & Gast, 2018). Ultimately, this was not able to be completed due to scheduling issues that delayed the start of initial data collection.

Procedures

Baseline

The subject entered the contrived instructional area selected for the specific skill and was provided the necessary materials to complete the task. At that time, they were given a relevant task instruction (e.g. "Do ____________________") and were allowed to attempt the behavior. In baseline, more potential target tasks were observed than were later targeted in the intervention. Following the collection of stable baseline data, the researcher graphed the data, and targets for the yoked intervention were selected based on similar level and length. In the baseline setting, no additional instruction or reinforcement were provided to the learner.

Because of this, baseline data collection was kept brief to avoid potential problem behavior on the part of the subject. Each potential target's baseline was completed separately, with time in between given to the subject to engage in preferred activities.

Video Modeling

The participant entered the instructional setting necessary for the target and was presented a tablet with a video preloaded by an adult. After being given an instruction functionally equivalent to "watch this," the model video was started. At the conclusion of the video clip, two verbal cues were given. The first a brief "your turn!" followed by a presentation of the task direction, "do ___________________." If the learner successfully completed the task modeled in the video, verbal praise and a break were given, where reinforcers were available to access. A list of potential reinforcers were selected prior to the start of the intervention using a multiple stimulus without replacement (MSWO) preference assessment. Items that Jess showed high preference for were kept on hand for the intervention, though her mands for those or other stimuli were honored for other reinforcers. After determining a general hierarchy via the preference assessment, the subject was given a choice of reinforcers prior to the start of the session. For each session, the reinforcer was kept consistent across conditions to protect against any potentially confounding affects. Throughout the attempt, correct steps were given mild, specific praise in the form of "good doing ____________" and errors resulted in "no" being said. Regardless of the learner's response in that moment, no additional corrective feedback was given to the learner during the trial. If an error occurred during the sequence of events, the trial was discontinued and brief praise for the attempt was given before a brief break. During the break following an error, access to the highly preferred item was not given, but the researcher did engage in some verbal exchanges and general rapport building.

Description of Data Analysis

Visual analysis was used to analyze the data. The visual analysis conducted followed the procedures laid out by Ledford and Gast (2018), where they detailed the systematic process for conducting visual analysis. This was done with graphs used to represent the data collected in each session. Using the method described previously (percentage of trials correct), the subject's performance was charted across the baseline and intervention phases. Each individual treatment was charted on its own data path within the same graph, with a solid a dotted line differentiating the two visibly.

As mentioned previously, selection of the target treatments was done following the graphing of baseline data. This was also in line with the Ledford and Gast (2018) procedure, which called for consistent graphing of data throughout a single case design study. This was also assessed using visual analysis. Each of the treatments was graphed using the described procedure, at which point the selection was made based on stability of responding at similar levels of performance in both final targets. As there was only one "phase" of the intervention, visual analysis was utilized throughout the intervention process to gauge level of performance session to session. This was done by looking at the trend, level, and variability of the data throughout this process. In addition, the magnitude of the change was looked at, as it was important for insight into the efficacy of each video modeling procedure.

Visual analysis was also used to assess the final findings of the study. A comparison of baseline to intervention data, and the resultant change in behavior lent credence to the intervention being responsible for the change in behavior. Due to the selection of the adapted alternating treatment design, the individual targets were also examined against one another. Here, the researcher evaluated the following primary factors: the immediacy of the change observed and the stability of the change.

Internal and External Validity, Reliability, and Generalizability

When working with young subjects, maturation, or development and growth due to aging, is a consistent threat to the validity of a study. However, the use of the adapted alternating treatments design mitigates some of these concerns (Ledford & Gast, 2018). By collecting both baseline and intervention data in yoked sessions, there is no danger of maturation affecting the data collected within sessions. This study also took place over less than two weeks, mitigating this risk as well. On a session-to-session basis, the entirety of the intervention lasted from around 15 minutes to a half hour of time from start to conclusion. Additionally, by randomizing the presentation of the learning targets on a session-to-session basis, any potential effects from time of day or fatigue are eliminated when comparing the treatments to one another.

Interobserver agreement (IOA) was collected across at least 25% of sessions. The second data collector watched the pre-recorded sessions and recorded data on the participant's correctly completed steps on the respective task analyses (TAs). IOA was calculated by dividing the number of agreements on correctly completed steps by the total number steps in the task analysis (Cooper et al., 2020). Data collectors were other Registered Behavior Technicians (RBT)and Board Certified Behavior Analysts (BCBAs) also working in the clinical setting. The secondary data recorders were trained using practice videos that were recorded specifically for that purpose. The two individuals collected practice data until 95% agreement was met in the data. IOA data and session data collected by the researcher are in Appendix A at the end of this document.

Procedural fidelity data were also collected to ensure accuracy and consistency on the part of the researcher. The RBTs and BCBAs trained in data collection for IOA data were also tasked with collecting procedural fidelity data in other sessions. Similar to the IOA data collected from observing the learner, procedural fidelity data were recorded using a task analysis of steps the researcher was expected to complete. These data were also taken in 25% of trials. A procedural integrity checklist and task analysis is included in Appendix B.

IOA and procedural integrity data were taken for a total of five of the 13 sessions. IOA data averaged 98% agreement across the sessions, and no single session was lower than 94% agreement. There were two issues in the comparison of IOA data. In one session, a camera issue resulted in the recording ending before the session was completed. For this session, the data taken by the IOA recorder were compared through that point with the researcher's. Additionally, problem behaviors that occurred prior to the start of an attempt were notated differently by the researcher and IOA recorder. After noting these discrepancies, the two reviewed the video and agreed on how this should be notated. Procedural integrity was at a similarly high level. Across all sessions, it averaged 99% fidelity of implementation, with no sessions having more than one error observed.

Results

The research questions for this study were: 1) Will the use of first- versus third-person video modeling (i.e., point of view vs. scene) influence the speed at which a learner acquires a skill? And 2) In a follow up probe coming after the discontinuation of instruction, will the first- or third-person model have an effect on the durability of the skill learned? The following section will discuss the results of the intervention, and how they pertain to these central questions.

Baseline

As discussed in the methods section, a variety of targets were evaluated in baseline with subject. Per discussion with the subject's family as well as input from their BCBA and RBT who routinely work with them, the following behaviors were selected: labeling 4 different household items (spoon, cup, fork, plate), putting on a shirt, putting on shoes, and writing their name. For all of these, a task analysis was developed by the researcher to orient data collection prior to baseline, though some minor adjustments were made to the TA's for the two behaviors that were eventually selected for the intervention. Rather than the planned 3 baseline sessions, Jess was observed for a total of four baseline sessions due to ease with scheduling and implementation.

Following the four baseline sessions, all of the labeling targets were withdrawn from consideration after Jess was successful in labeling the target for three of the four items. Additionally, labeling was included because it is an area of current parent concern, but the relatively short task made it difficult to compare against the other, longer skills.

Across the four sessions, Jess was able to complete 17% of the steps for putting on a shirt without a video model and performance was consistent across session. Every time the task directive was given she did pick up the shirt and hold it up, but made no further moves to put it on. Performance in the two remaining behaviors was slightly more variable. For putting on shoes, Jess's performance was 0%, 50%, 33%, and 33% steps completed correctly across the sessions. For writing her name, she completed 29%, 0%, 43%, and 14% steps correctly.

Based on the relatively low level and the stability of the data, "put on shirt" was selected as the first behavior to target in the video modeling intervention. Between the two remaining behaviors, "put on shoes" and "write name", the selection was more challenging due to the relatively similar level and variability of the behaviors. Ultimately write name was selected due to a slightly lower level of performance, and due to how dissimilar it was to the other selected behavior allowing for less risk of sequence effects or confounding discriminative stimuli (S^Ds). While dissimilar in terms of the action, the number of steps and length of time required for the learner to complete the action was similar, and appropriate to draw comparisons between. These task analyses can be seen in Appendix A.

Video Modeling

Jess's intervention sessions took place in nine sessions across five weekdays. Data collection sessions were done bookending her typically scheduled ABA sessions in the clinic; one immediately upon arrival and one immediately preceding leaving with about an hour in between. Jess showed some improvement with both behaviors but did not reach mastery for either within the time available for the intervention. For both behaviors, 5 trials or attempts at the behavior were given in each session, though in some trials it was not possible to complete each of these, which will be discussed later.

Jess showed the greatest improvement in the behavior taught using a first-person perspective model. Over the course of the intervention Jess went from completing an average of 21% of the task analysis in baseline to completing as much as 74% of the steps across the five attempts correctly in the 11th session. Jess did not maintain this behavior at that high a level for the remaining two sessions, but did not show severe regression (60% and 63% of steps completed accurately in those respective sessions). All of these data can be seen in Figure 1 located below.

For the behavior presented in the third-person perspective, there was less change overall in level throughout the intervention, though there was a slight upward trend. Additionally, there was very little variability in these data, all points falling within ten percentage points of one another. Following a baseline average of 17% of the task analysis being completed, Jess's peak session for this behavior had her independently completing 42% of the behaviors in sequence. This was, consequently, Jess's last data collection session.

Due to scheduling issues, the post instruction probe was not able to be completed prior to this write up of the study. Initially, a post instruction probe had been planned for two weeks following instruction to assess the durability of behavior change. This did not occur due to the compounding factors of an appropriate subject not being available until later than planned, followed by this subject not being available for a week further due to a potential COVID-19 exposure.

Though the initial results would show an indication that the use of the first-person model was more effective in this study, there were confounding factors that impact the strength of any conclusions that can be drawn. These confounds will be discussed in greater depth in the discussion section. One overall take away is that the results of this study do show that video modeling can be used to teach these behaviors. Both behaviors did show some increase over the course of the nine sessions with video models as the only teaching strategy used.

Discussion

The purpose of this study was to evaluate the effectiveness of two different methods of video modeling and observe if one may result in more durable results compared to the other. Due to scheduling conflicts, the latter of those two goals was not met. The start of data collection with the subject was delayed by a COVID-19 related scare within their home, which required time for isolation and negative tests before in person and specifically in clinic sessions could resume. Because this was delayed, there was not an opportunity to complete post instruction probes with Jess prior to the completion of this write up. As it stood, it was only possible to reach the minimum required sessions (13) in the intervention phase itself. While there were overall indications as to the success of both video models in the short time of the intervention, additional confounds make it difficult to draw meaningful conclusions about the efficacy of one perspective over the other.

The root of these issues was the specific task analysis used to create the video model for putting on a shirt. Per this TA, the subject was supposed to put in their arms prior to pulling the shirt over their head. While in baseline, the subject did not perform enough of this behavior for the discrepancy to become apparent to the researcher prior to the intervention phase. As the protocol is written, any error or out of step piece of the behavior is to result in the immediate discontinuation of the attempt. This led to multiple trials in which Jess appeared to be completing the action of putting on a shirt with no errors but performing the steps out of sequence with the model. In multiple trials where the researcher was either unsuccessful in interrupting the behavior or Jess persisted through the attempt to end the trial, she was able to complete the task in its entirety, just not in the same sequence as the video. After observing this, Jess's mother confirmed that when dressing her that is the sequence they complete the steps in.

Data from this part of the intervention are contained in Figure 1. While Jess's performance overall appears low in this area, these observations make it appear likely that had the task analysis and video model been slightly different, Jess would have a much higher percentage of correct trials overall. While this does indicate a shortcoming in assessing the social validity of this method of putting on a shirt for Jess, it may also indicate poor attending to the video model. Jess is only four years old, and at times the length of the video models did appear to be a barrier when it regarded maintaining her attention. While neither model was longer than 15 seconds, she did not consistently attend to the video.

Another factor that influenced this study was the prevalence of problem behaviors. While Jess had mastered a behavioral objective recently for working on adult-directed tasks at a table for increasing durations, her RBT noted there had been regression in this area. During data collection for this study, there were problem behaviors, primarily elopement attempts that occurred. Following discussion with Jess's BCBA, attempts to elope were blocked, however, if Jess were to vocally say "all done" or a functionally similar equivalent, the attempt was abandoned, and she was allowed to leave the area.

Lastly, the amount of access to reinforcement may have been too thin of a schedule for this particular learner. Due to prevalent errors in the beginning of the intervention, the only reinforcement contacted was the behavior specific praise for the completion of steps. In later sessions where Jess was able to complete the entire action and contact the identified preferred item, she was likely to either do the same or at a similarly high level within that session. These behaviors did not always consistently maintain across sessions however, which presented similar issues in those sessions.

Limitations

The researcher identified 4 limitations for this study. First, the sparsity of reinforcement that was discussed above. Second, Jess's learning history with the task "putting on shirt" did have an impact on the study. A third limitation is the size of the population for the study, in this case, one. While it was appropriate based on the design of this intervention to work with only one subject, any results obtained and conclusions drawn from the study are only from this one learner. Although the findings lend further credence to the evidence-based practice of video modeling, based on the sample size and other limitations it is difficult to come to other concrete conclusions.

Lastly, the relatively few numbers of sessions were a limitation of this study. Four baseline sessions and nine intervention sessions are enough to begin to see and assess trends, level, and variability in the data, but could have benefited from more intervention sessions. For both behaviors, there were trends towards increased independent success that could have shown meaningful findings.

Contribution to the Current Research

This study does contribute to the existing literature of video modeling as a tool for teaching skills to people with autism. While the barriers encountered over the course of the study make comparisons of the two different perspectives challenging, the upward trend in both behaviors could indicate that both are feasible methods of demonstrating behaviors to learners. This is important to have in the existing scholarship as some behaviors simply cannot effectively be filmed from varying perspectives.

Areas for Future Research

There is an existing knowledge base for the use of video modeling in teaching skills to people with autism. With that said, there is a dearth of information concerning the effect viewing perspective has on skill acquisition. Further replications or similar studies to this one could increase our understanding of the efficacy of these models.

Additionally, combining this intervention with other evidence-based practices may lead to more success or meaningful results dependent on the learners. In the case of this specific learner, it is possible that the inclusion of narration prompts in the video could have led to increased success or attending. Similarly, the video prompting discussed in the literature review where the video is broken down into pieces reflective of the task analysis may be more feasible for younger learners that do not attend well for longer durations. Attending in general is an area that future research may benefit from taking data on in tandem with the other intervention data.

Recommendations for Practice

Taking into account the general reinforcement schedule for the learner is important. While Jess did show some improvement with the behaviors over the study, increased opportunities for reinforcement could have assisted in reducing problem behaviors and better skill acquisition. Additionally, in actual practice the inclusion of error correction procedures would expedite this learning process. These procedures were not included in this study in order to get a cleaner depiction of solely video modeling, but in the applied setting these would likely be beneficial to the learner and lead to less frustration.

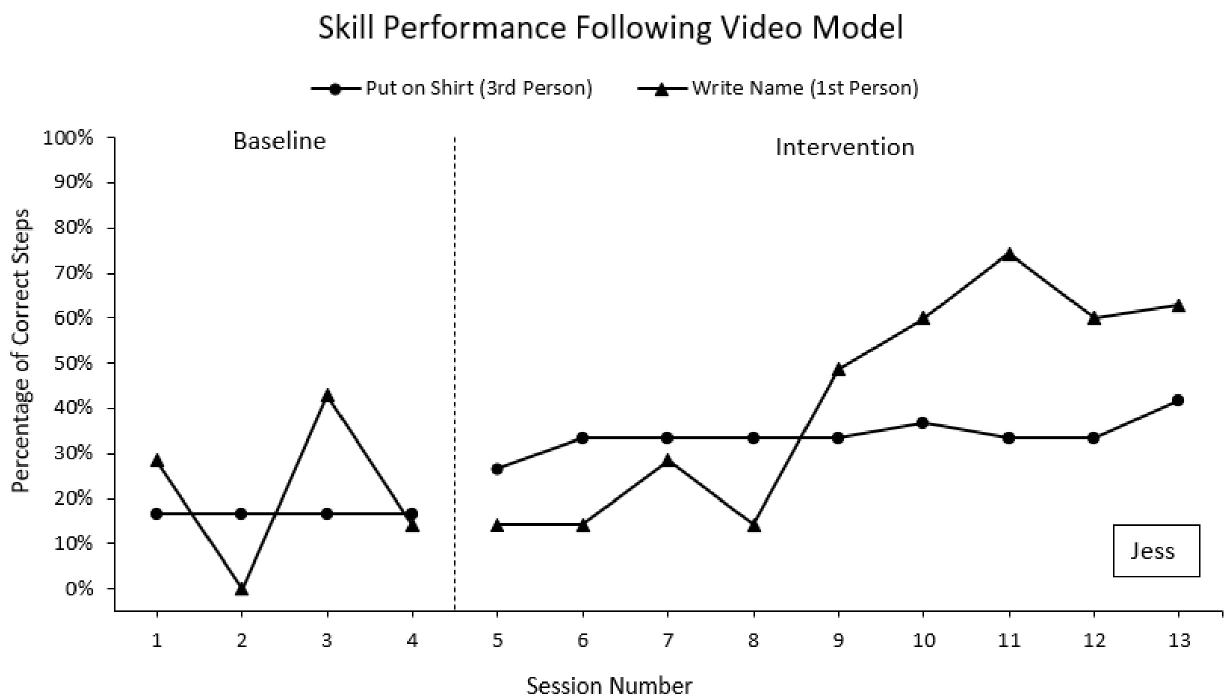
Another area that led to learner frustration in this study was the rigidity of the protocol in regard to the sequence of the task analysis. As described previously, the subject here showed signs of independently completing one of the tasks, but out of sequence from the created task analysis. Due to the construction of this protocol, incorrect or out of order steps led to discontinuation of the attempts, and not access to the highly preferred item. The first step to addressing this could be to observe this skill as it occurs in the home now, especially regarding the aid and prompts given by family members. This will be useful in developing a better task analysis of the skill. This rigidity was maintained in this target to ensure the procedure was kept consistent for the purpose of comparison to the other target, and its method of implementation. When working in an applied setting, this rigidity can be greatly lessened when appropriate for the target skill.

The two targets of this study are suitable examples of when this would and would not be appropriate. In putting on a shirt, whether the arms or head is put through first is overall unimportant, so long as the learner can do it consistently and independently, and the shirt ends in the correct orientation. Here, flexibility to the learner may be appropriate. In the other target of writing their name, maintaining the correct sequence of letters is important to being able to write neatly and efficiently. For this target, it would be more important to maintain the sequence of steps. Whether or not the target skill is one that has a specific sequence will be a determination that is made on a case-by-case basis. Overall, however, a similar intervention in an applied setting does not need to be stringently structured and maintained when the learner is showing signs of functional, appropriate success in the skill even if it does not match the precise sequence of the task analysis.

The last area to take into account during real world application is the occurrence of problem behaviors. This can be more or less of a concern based on that particular learner's history, but the intervention as described here may be particularly challenging for younger learners. The density and type of reinforcement provided should be altered dependent on that learner's history. Any alterations should be done with respect to protocols already in place with this learner, particularly if they have particular behaviors they often engage in.

Figure 1

Baseline and Intervention Performance for "Put on Shirt" and "Write Name"



Note: Data were collected in typically in sessions of five attempts per target, with some sessions being less due to problem behaviors and appropriate requests to end the task from the participant.

APPENDIX A

Client Pseudonym Inits.: _________________ Data Collector: _________________

Date:__________

First-Person POV Task Analysis and Datasheet

Task:	**Trial:**									
Task Analysis:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1.										
2.										
3.										
4.										
5.										
6.										
7.										
8.										
9.										
10.										
11.										
12.										
13.										
14.										
Totals (only + counts as "correct"):	__	__	__	__	__	__	__	__	__	__

Key:	+: Successful Completion	-: Unsuccessful	P: Problem behavior

Client Pseudonym Inits.: ________________ Data Collector: ________________
Date:_________

Third-Person POV and Datasheet

Task:	**Trial:**									
Task Analysis:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1.										
2.										
3.										
4.										
5.										
6.										
7.										
8.										
9.										
10.										
11.										
12.										
13.										
14.										
Totals (only + counts as "correct"):	__	__	__	__	__	__	__	__	__	__

Key:	+: Successful Completion	-: Unsuccessful	P: Problem behavior

Put on Shirt Task Analysis and Datasheet (3rd Person)

Task:	**Trial:**									
Task Analysis:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1. Pick up shirt										
2. Hold shirt upside down, so the front is along torso										
3. Put in right arm										
4. Put in left arm										
5. Pull up shirt with arms, and put head through hole										
6. Pull bottom of shirt down to cover body										
Totals (only + counts as "correct"):	__	__	__	__	__	__	__	__	__	__

Key:	+: Successful Completion	-: Unsuccessful	P: Problem behavior

Write Name/Copy Task Analysis and Datasheet (1st Person)

Task:	**Trial:**									
Task Analysis:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1. Pick up utensil										
2. Write first letter										
3. Write second letter										
4. Write third letter										
5. Write fourth letter										
6. Write fifth letter										
7. Write sixth letter										
Totals (only + counts as "correct"):	__	__	__	__	__	__	__	__	__	__

Key:	+: Successful Completion	-: Unsuccessful	P: Problem behavior

APPENDIX B

Data Collector: ________________ Date:_________

Procedural Integrity Checklist

Task:	**Trial:**									
Task Analysis:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1. Researcher gains learner's attention										
2. Task direction (watch this/do this, etc…) is given										
3. Video was played following task direction without additional prompting or cues										
4. After video, additional cue (your turn/now you, etc…) is given										
5. Following post video cue, instruction "do ___________________" is given										
6. Learner is given the opportunity to attempt the behavioral chain (refer to TA)										
7. During attempt, researcher provides correct mild praise (good/yes) for correct steps and "no" for incorrect steps										
8. If attempt is unsuccessful, researcher ends the trial, and gives praise for effort. A short break is allowed after.										
9. If attempt was successful, effusive praise is given along with access to reinforcing item										
Totals (only + counts as "correct"):	__	__	__	__	__	__	__	__	__	__

Key:	+: Successful Completion	-: Unsuccessful

References:

Aldi, C., Crigler, A., Kates-McElrath, K., Long, B., Smith, H., Rehak, K., & Wilkinson, L. (2016). Examining the effects of video modeling and prompts to teach activities of daily living skills. *Behavior Analysis in Practice*, 9(4), 384–388. https://doi.org/10.1007/s40617-016-0127-y

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology*, 1(6), 589–595. https://doi.org/10.1037/h0022070

Bandura, A., Ross, D., & Ross, S. A. (1963). Vicarious reinforcement and imitative learning. *The Journal of Abnormal and Social Psychology*, 67(6), 601–607. https://doi.org/10.1037/h0045550

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(3), 264–287. https://doi.org/10.1177/001440290707300301

Cannella-Malone, H., Fleming, C., Chung, Y., Wheeler, G., Basbagill, A., & Singh, A. (2011). Teaching daily living skills to seven individuals with severe intellectual disabilities: A comparison of video prompting to video modeling. *Journal of Positive Behavior Interventions*, 13(3), 144–153. https://doi.org/10.1177/1098300710366593

Charlop-Christy MH, Le L, & Freeman KA. (2000). A comparison of video modeling with in vivo modeling for teaching children with autism. *Journal of Autism & Developmental Disorders*, *30*(6), 537. https://doi.org/10.1023/a:1005635326276

Charlop, M., Schreibman, L., & Tryon, A. (1983). Learning through observation: The effects of peer modeling on acquisition and generalization in autistic children. *Journal of Abnormal Child Psychology*, 11(3), 355–366. https://doi.org/10.1007/BF00914244

Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis* (3rd ed.). Hoboken, NJ: Pearson Education.

Cotter, C. (2010). Evaluating the effects of camera perspective in video modeling for children with autism: Point of view versus scene modeling. *ProQuest Dissertations Publishing.*

Delano, M. (2007). Video modeling interventions for individuals with autism. *Remedial and Special Education*, 28(1), 33–42. https://doi.org/10.1177/07419325070280010401

Goldsmith, T., & LeBlanc, L. (2004). Use of technology in interventions for children with autism. *Journal of Early and Intensive Behavior Intervention*, 1(2), 166–178. https://doi.org/10.1037/h0100287

Kleeberger, V., & Mirenda, P. (2010). Teaching generalized imitation skills to a preschooler with autism using video modeling. *Journal of Positive Behavior Interventions*, *12*(2), 116–127. https://doi.org/10.1177/1098300708329279

Ledford, J., & Gast, D. (2018). *Single case research methodology: applications in special education and behavioral sciences* (Third edition.). Routledge.

Quill, K. (1997). Instructional considerations for young children with autism: The rationale for visually cued instruction. *Journal of Autism and Developmental Disorders*, 27(6), 697–714. https://doi.org/10.1023/A:1025806900162

Shipley-Benamou, R., Lutzker, J., & Taubman, M. (2002). Teaching daily living skills to children with autism through instructional video modeling. *Journal of Positive Behavior Interventions*, 4(3), 165–175.

Spriggs, A., Gast, D., & Knight, V. (2016). Video modeling and observational learning to teach gaming access to students with ASD. *Journal of Autism & Developmental Disorders*, 46(9), 2845–2858. https://doi.org/10.1007/s10803-016-2824-3