Spring 4-26-2022

A Comparison of Models in Video Modeling to Teach Vocal Skills

Caroline Phan
Rollins College, cphan1@rollins.edu

A Comparison of Models in Video Modeling to Teach Vocal Skills

A Thesis
by
Caroline Phan

Submitted to the Faculty of the Department of Health Professions
at Rollins College in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF ARTS IN APPLIED BEHAVIOR ANALYSIS AND CLINICAL SCIENCE

April 2022
Winter Park, FL

Acknowledgements

I cannot even begin to express my sincere gratitude and thanks to the people who have helped me along this journey. First and foremost, I would like to thank my family for always supporting and encouraging me so wholeheartedly in all of my personal and academic endeavors. I would like to give a huge shout out to my boyfriend, Addison Lanpher, for listening to me talk about my thesis for the past school year while only understanding about half of what I'm saying. Your pep talks, home-cooked meals, and encouraging words kept me going. Thank you so much to my amazing cohort for being the best support system! I would not have wanted to go through this program with any other group of people and I'm so happy to have made lifelong friends and colleagues in each of you. I'd especially like to thank my biggest cheerleader, Becca Mischuck, for teaching me more than I could ever describe, for being the best friend, and for always being up for a study or hang out session.

Thank you to my chair, Dr. Kara Wunderlich, for your knowledge and expertise, flexibility, and encouraging words throughout this entire research process. Thank you, especially, for helping me find the right words when I couldn't. I have learned so much from you. Thank you to my committee members, Dr. Michele Williams and Dr. Scott Hewit, for your time, guidance, and support; not only during this research process, but during my time at Rollins.

A special shout out to Sabrina Frank and her incredible team of RBTs. Sabrina, thank you for believing so strongly in the research, letting me come in to run sessions as much and as long as needed, and for your helpful insight throughout. Thank you as well to the RBTs at Acorn Health, who showed great flexibility and kindness, and for their help collecting IOA and treatment integrity data.

Table of Contents

Abstract

Video modeling is a strategy used to teach learners new skills by viewing a video of a model engaging in a target behavior and then having the learner imitate the behavior (Cooper et al., 2019). Much research has been conducted on the effectiveness of video modeling when compared to in vivo modeling, but there are still various limitations to consider, including its effectiveness as a solo intervention, who are the most impactful models to include, and the effectiveness of the intervention for certain populations. The purpose of this study is to evaluate the use of video modeling as a solo intervention to teach tacting letter sounds or letter blends to children with ASD with limited vocal repertoires and to evaluate the effectiveness of different models in the video recordings. Results suggested that the sibling model was the most effective for one participant after the inclusion of a training phase. Results were inconclusive for the second participant.

Keywords: language acquisition, models, peer, sibling, tacting, video modeling

A Comparison of Models in Video Modeling to Teach Vocal Skills

Modeling is a strategy used to teach learners new skills by having them imitate demonstrations of the skills by either live or representative models (Cooper et al., 2019). One of the earliest studies to evaluate the effects of live peer modeling with children with autism spectrum disorder (ASD) was conducted by Egel et al. (1981). Because of the new passage of legislation at the time that mandated the least restrictive learning environment for students with ASD, researchers wanted to assess whether children with ASD should be mainstreamed into classrooms and whether their peers without ASD could serve as role models for appropriate behavior. Egel et al. found that the children with ASD were able to accurately imitate the peer models' correct responses and that the children were able to maintain the correct responding over time as the peer models were removed. This study highlighted that learning through modeling could be an effective tool in teaching skills to students with ASD.

Video modeling is a method of teaching in which learners acquire new skills by viewing a video of a model engaging in a target behavior and then having the learner imitate the behavior (Cooper et al., 2019). Research has shown there are many benefits to using video modeling when compared to traditional live modeling. Charlop-Christy et al. (2000) evaluated the effectiveness of video modeling compared to in vivo modeling to teach developmental skills to children with ASD and the results highlighted the advantages of using video modeling. The authors found that video modeling was more effective and efficient at teaching a wide range of skills and behaviors, such as spontaneous greetings, self-help skills, expressive labeling, and social and cooperative play. The authors also found that video modeling led to more rapid acquisition of the skills for four out of the five children when compared to in vivo modeling. Additionally, only the skills learned through video modeling were able to be generalized and maintained across different stimuli, settings, and people.

Cardon and Wilcox (2011) confirmed the findings by Charlop-Christy et al. (2000). Cardon and Wilcox compared video modeling with reciprocal imitation training (RIT) to evaluate which intervention resulted in more imitation acquisition among six children with ASD between the ages of 20-48 months. Overall, the researchers found that both interventions were effective in increasing object imitation in children with autism. However, the researchers found that subjects in the video modeling group demonstrated a rapid increase in their imitation skills while those in the RIT group showed more of a steady growth throughout the study. Both groups showed that imitation skills were maintained and generalized at 1- and 3-weeks follow-up. Though both conditions showed acquisition of the skill, maintenance, and generalization, the researchers considered video modeling to be the more effective intervention because of its motivating and reinforcing nature. The authors also suggest that the use of a screen in video modeling may help by limiting other distractions while helping the children focus on salient cues.

Charlop-Christy et al. (2000) and Cardon and Wilcox (2011) are just two examples of the methods in which video modeling has been extensively studied for more than two decades. Most recently, the research has focused on using video modeling to teach a wider range of skills to different populations of people. Abadir et al. (2021) evaluated the effects of video modeling on the acquisition of abduction-prevention skills by adolescents with ASD. The subjects were taught how to secure a code word from strangers and known individuals after watching the target behavior modeled on video. The researchers found that the subjects were able to learn the complex safety skill through the use of video modeling combined with vocal instruction and error correction procedures. Two of the subjects were able to learn through video modeling alone. Video modeling has also been used to teach medical residents and specialists how to perform specialized surgical skills. Alkatout et al. (2021) used video feedback and video modeling to teach both beginner and advanced surgeons how to perform

an intracorporeal knot on the vaginal vault, which is considered a complex procedure in gynecologic laparoscopic surgery. The participants highly rated the value of using video modeling and video feedback to develop the surgical skill and reported the training improved their confidence and posture during the procedure.

Although there are many studies that support the use of video modeling with children with autism, there are limitations to the existing literature. Video modeling is often used in conjunction with other interventions as part of a treatment package. For example, Kleeberger and Mirenda (2010) conducted a study to evaluate the effects of using video modeling to teach generalized imitation skills to a preschooler with autism. During the study, a 4-year-old boy was instructed to watch videos for songs that included a model engaging in fine and gross motor skills, finger play, and toy play activities. The authors evaluated the use of video modeling alone and video modeling with highlighting, prompting/fading, and reinforcement on both mastered and not mastered actions. Kleeberger and Mirenda (2010) found that video modeling alone did not have any effect on the child's performance of the motor skills, finger play, or toy play activities on either mastered or not mastered tasks. However, the child was able to complete accurate imitations of the model when video modeling plus highlighting, prompting/fading, and reinforcement was implemented on both mastered and not mastered tasks. It is unknown, therefore, which additional instructional components were necessary to increase the effectiveness of video modeling.

Another limitation to the existing literature is the lack of consistent information about the characteristics and logistics of effective video models. To date, little to no research has been conducted on the logistics of video recordings, including how long videos should be, if there should be narration, inclusion of captions, etc. There have been mixed findings from different studies concerning the most effective models to use in video recordings. Bandura (1994) postulated that peer models that look close in age and that have physical

characteristics similar to the participant are the most effective models to use during in vivo modeling. To date, no studies have evaluated if these findings are also valid in video modeling. In contrast, Hoogerheide et al. (2016) conducted a study evaluating the most effective models to use in video recordings geared towards neurotypical adolescents and concluded that adults were the more effective model when compared to peers. In a separate study, Hoogerheide et al. (2016) also explored whether the gender of the model had a significant impact on the learning of the participants. The researchers found that the gender of the model did not affect the degree to which the participants improved their performance. However, the participants found the learning less effortful when watching a male model compared to a female model for both male and female participants.

Another limitation in the existing research is the lack of consensus for populations and skills for which video modeling would be most effective. Research has shown that video modeling is effective for most learners between the ages of toddler to adult. Extensive research has been conducted using video modeling with participants with ASD, and it has been accepted as an evidence-based practice (National Professional Development Center on Autism Spectrum Disorders, 2010). It is especially an effective intervention when the skill is focused on motor imitation, as numerous studies involving video modeling include motor imitation as the target behavior. However, few, if any studies have evaluated the use of video modeling to teach basic tacting skills when compared to the many studies focusing on teaching intraverbal skills, such as through play comments and conversational speech. Furthermore, limited research has been conducted with those who do not display a strong imitation or verbal repertoire.

Previous research suggests that a peer model both close in age and who looks similar to the learner may be more beneficial when used in live modeling. Therefore, it may be worth investigating whether a sibling may be especially effective as a model in video recordings

because of familiarity, possible closeness in age and in many but not all cases, physical resemblance. Although it is likely more difficult to obtain a sibling model than another type of model (e.g., therapist, adult, or peer), the sibling model may have greater impacts on the learner. For example, in the case of vocal behavior, the model might sound like the learner. The purpose of this study is to evaluate the effectiveness of using video modeling alone to teach tacting skills to children with ASD who have limited vocal skills. In addition, the study will compare the effectiveness of different types of models used in the video recordings.

Method

Participants and Setting

Olivia was a 5-year-old girl with a diagnosis of ASD who received services at a local ABA clinic. The participant was recruited through a recommendation from her Board Certified Behavior Analyst (BCBA). She had prerequisite imitation skills as well as the ability to focus on a video on a screen for varying durations of time (1-3 min). The participant had a limited vocal repertoire and had not mastered tacting letter blends, as reported by her BCBA. All sessions took place in the clinic where she was currently receiving services.

Spencer was a 3-year-old boy with a diagnosis of ASD who received services at the same clinic as Olivia. Spencer was recruited through a recommendation from his BCBA. He had prerequisite imitation skills as well as the ability to focus on a video on a screen for varying durations of time (30 s to 1 min). The participant had a limited vocal repertoire and had not mastered tacting letter sounds, as reported by his BCBA. All sessions took place in the clinic where he was currently receiving services.

Materials

An iPhone XR was used to record the videos that were used during each trial and the videos were played on an Apple MacBook Air during sessions. Index cards (3” x 5”) with each target printed in black Times New Roman font, size 55, were used.

Videotapes

Video recordings were made for each video modeling condition. The videos were made by recording the BCBA or therapist at a table with either a sibling, a parent, or a peer. Videos showed each model displaying the correct response after the S^D (e.g., "What sound does 'J' make?" or "What sound does 'I-T' make?") was delivered by the BCBA or therapist. Individual videos were made for each target and each video was approximately 5 to 7 s in duration.

Sibling modeling video. The therapist presented the S^D to a sibling of the participant and the sibling correctly responded with the target within 5 s. The therapist provided verbal praise (e.g., "Great job!").

Parent modeling video. The BCBA presented the S^D to a parent of the participant and the parent correctly responded with the target within 5 s. The BCBA provided verbal praise (e.g., "Great job!").

Peer modeling video. The BCBA presented the S^D to a peer of similar age (within one year) to the participant and the peer correctly responded with the target within 5 s. The BCBA provided verbal praise (e.g., "Great job!").

Response Measurement and Reliability

An adapted alternating treatment design (Sindelar et al., 1985) was used. Baseline sessions consisted of 12 trials, and each teaching condition session consisted of six trials. The conditions were randomized within each series. Each presentation of the S^D counted as one trial. Data was collected using paper and pencil. During baseline, each target was presented once in random order in each session. During the teaching conditions, each target was presented twice in each session and were presented in random order during sessions. Each trial was scored as correct, incorrect, or no response. Correct responding was defined as the participant responding to the S^D with the accurate target sound within 5 s during the first or

second presentation of the S^D. Incorrect responding was defined as the participant responding to the S^D with an inaccurate target sound or any other sound within 5 s. No response was defined as the participant not engaging in any vocal response.

An independent observer scored at least 30% of the sessions for treatment integrity and interobserver agreement (IOA). Treatment integrity was measured to evaluate whether the experimenter implemented the procedures as described in the study. A task analysis for each condition was created with descriptions of the steps included. Treatment integrity scores were calculated for each session by dividing the number of steps conducted correctly divided by the total number of steps in the task analysis multiplied by 100%. For interobserver agreement, the trained observer collected response data using a trial-by-trial method, and the data were compared to the experimenter's data for percentage agreement for each trial. Percentage agreement was calculated as agreements divided by the number of trials multiplied by 100%. For Olivia, IOA was assessed for 31% of all sessions and was 98% (range, 75% to 100%). Treatment integrity was assessed for 30% of all sessions and was 100%. For Spencer, IOA was assessed for 30% of all sessions and was 99% (range, 92% to 100%). Treatment integrity was assessed for 30% of all sessions and was 100%.

Procedure

Table 1 displays the twelve letter blends that were included in the study for Olivia. The experimenter chose the letter blends from a list of the most common consonant blends from an online list from a Google search. Additionally, the experimenter included letter blends with consonants and vowels at the request of the participant's BCBA. The 12 letter blends were then divided between the three teaching conditions and the control condition. Three specific letter blends were assigned to each condition to ensure that there was minimal similarity in sound or letter shapes within each condition, and to ensure that at least one consonant-vowel blend was included in each condition. The teaching phase ended and targets

were considered mastered when the participant demonstrated 80% accuracy across three consecutive sessions.

Table 2 displays the twelve letters that were included in the study for Spencer. The letters were chosen by random selection through an online generator. The 12 letters chosen by the generator were then divided between the three teaching conditions and the control condition. Three specific letters were assigned to each condition to ensure that there was minimal similarity in letter sound or letter shape within each condition (e.g., the letters “B”, “D”, and “P” were not assigned to the same condition). The teaching phase ended and targets were considered mastered when the participant demonstrated 80% accuracy across three consecutive sessions.

Preference Assessment

A multiple-stimulus-without-replacement (MSWO) preference assessment was conducted (DeLeon & Iwata, 1996) for each participant. The top five items and edibles as recommended by the participants’ BCBA via verbal reporting were included in the assessment. Before starting the assessment, the participant had 30 s of access to each item or a small piece of the edible. The items were then removed and the researcher presented all the items/edibles on the table in a semi-circle in front of the participant. The participant was instructed to “pick one” and received 30 s access to the item or one piece of the edible that they chose. While the participant interacted with the chosen item, the researcher rearranged the order of the remaining items in a semi-circle. After 30 s or item consumption, the researcher removed the chosen item from the participant by saying “my turn.” The researcher continued the above steps until no items remained in the array. If the participant did not pick an item within 15 s of the S^D, the researcher rearranged the items in the semi-circle and presented the S^D again. Attempts to select more than one item were blocked and the S^D was presented again. The researcher recorded the order of the items selected.

Baseline

In each trial, the experimenter presented each of the 12 target cards individually to the participant with the vocal S^D (e.g., “What sound does ‘A’ make?” or “What sound does ‘C-L’ make?”). The experimenter waited 10 s for a response before presenting the next target card. The experimenter provided no feedback or reinforcement throughout the session.

Sibling Video Modeling

During each trial of this phase, the participant watched one of the sibling videos on the computer. The experimenter said, “Let’s watch a video!”, then played the video for the corresponding target for the participant. The experimenter then presented the same S^D from the video. If the participant displayed the correct response, the experimenter provided verbal praise and the highest-preferred item from the MSWO for 30 s. After 30 s, the experimenter said, “My turn,” and took back the reinforcer. If the participant did not display the correct response or did not engage in any response, the experimenter said, “Let’s watch again,” and played the video again. The experimenter then presented the S^D again. If the participant displayed the correct response, the experimenter provided verbal praise and the reinforcer from the MSWO for 30 s. If the participant did not display the correct response or did not engage in a response, the experimenter said, “Let’s watch another video,” and proceeded to the next trial.

Parent Video Modeling

This phase was conducted in the same format as the previous condition, but videos with the parent were used.

Peer Video Modeling

This phase was conducted in the same format as the previous condition, but videos with the peer were used.

Control Condition

During this phase, the experimenter conducted trials in an identical manner to baseline. The participant was presented each card individually with the vocal S^D (e.g., “What sound does ‘C’ make?” or “What sound does ‘B-R’ make?”). The experimenter waited 10 s for a response before presenting the next card until two trials had been conducted for each of the three targets. The experimenter provided no feedback or reinforcement throughout the session.

Training Phase

For Olivia, the experimenter implemented a training phase after incorrect responding occurred for three consecutive series in the initial video modeling phase. Three videos were recorded to target a different vocal skill specifically for teaching purposes. The videos were made by recording the BCBA and the therapist, and arbitrary words were chosen as targets for the vocal S^D, “What’s the magic word?”. In the videos, the BCBA presented the vocal S^D and the therapist responded with the target word (e.g., zebra, banana, or cookie). The experimenter then conducted trials similar to the trials in the initial video modeling phase. However, the experimenter highlighted the therapist’s response in the video by verbally prompting Olivia to listen to what the model was saying while pointing to the model in the video. Each correct response was immediately reinforced with verbal praise and the highest-preferred item from the MSWO for 30 s. After Olivia demonstrated 80% accuracy across three consecutive sessions, procedures from the video modeling phase were reinstated.

Probes

The purpose of this phase was to assess for changes in tacting in the absence of the video model following teaching. During this phase, the experimenter evaluated the targets mastered using procedures similar to baseline sessions.

Results

Figure 1 shows Olivia's percentage of correct responses across sessions. No correct responses occurred across three consecutive baseline sessions. In the video modeling phase, no correct responses occurred across all conditions for three consecutive series, and the experimenter implemented the training phase. After the training phase, Olivia met the mastery criterion with 100% accuracy in the sibling video modeling condition during the first three series. The peer video modeling condition displayed a sharp increasing trend and maintained accurate responding at 67% accuracy for four consecutive series. The parent video modeling condition displayed an increasing trend, and after responding at 83% accuracy, the parent video modeling condition maintained at 67% accuracy. The control condition remained at zero levels throughout the study. During the probes of mastered letter blends, she was able to accurately respond for three consecutive sessions with 100% accuracy.

Figure 2 shows Spencer's percentage of correct responses across sessions. No correct responses occurred across three consecutive baseline sessions. The peer video modeling condition had a high level of correct responding and was the condition closest to reaching mastery criterion. The parent video modeling condition also showed a high level of correct responding during the first two consecutive sessions; however, Session 9 was terminated because the participant engaged in a tantrum. Responding maintained at 67% accuracy during the final two sessions of the parent video modeling condition. The sibling video modeling condition displayed a moderate level of correct responding while the control condition showed a low level of correct responding with some variability.

Discussion

The current study evaluated the use of different models in video modeling to teach vocal tacting skills. Consistent with previous research on video modeling (e.g., Charlop-Christy et al., 2000; Cardon & Wilcox, 2011; Abadir et al., 2021), the study showed that it is

an effective teaching intervention and can aid in skill acquisition. This study added to the video modeling literature by expanding on previous studies by evaluating a new target skill not previously researched. Video modeling studies have primarily focused on evaluating the effectiveness of the intervention to increase motor imitation or intraverbal skills, but to the experimenter's knowledge, this is the first study to evaluate video modeling's effect on tacting vocal skills. The participants were taught to vocally tact letter sounds or letter blends by watching videos of different models engaging in the target skills. Thus, the study also compared the efficacy of using a peer, sibling, or parent as a video model, which had not been compared as directly or extensively in other studies. The study found that the most effective model for Olivia was the sibling model, as it was the only condition that reached mastery criterion. When the videos were removed and probes were conducted for the mastered letter blends, Olivia was able to correctly tact all the letter blends across three consecutive sessions.

As hypothesized by the experimenter, the sibling model was the most effective model used for Olivia. However, it is interesting to note that it was the only condition to meet mastery criterion for her. There may be several reasons for this finding that are idiosyncratic to the participant. First, the participant had two older sisters to choose from to serve as a sibling model. Olivia's mother recommended one sister over the other based on the relationship between the sister and the participant. Second, according to her BCBA, she did not have a peer that she interacted with at the clinic. Instead, she preferred to only interact with her therapists. Third, the peer used in the videos was relatively new to the clinic and may not have had much interaction, if any at all, with Olivia. The peer was chosen to serve as a model because he was the only client in the clinic who was able to tact letter blends. These reasons may warrant future research into whether the reinforcing value of a person affects the efficacy with which a person learns from the video model.

An interesting consideration is the necessity of a training phase after the initial video modeling phase. After Olivia engaged in no correct responding over three consecutive series during the initial video modeling phase, the experimenter created three new videos showing the therapist and BCBA engaging in a different target vocal skill that was unrelated to letter blends. The videos were created to be used exclusively as training videos to increase the saliency of the model's vocal response and to increase the participant's contact with reinforcement. When watching the training videos with Olivia, the experimenter implemented gestural and verbal prompts to highlight the model in the video and to prompt attending to the vocal responses emitted by the model. She was able to meet the training mastery criterion after three consecutive sessions of watching the training videos. After the training videos phase, Olivia immediately engaged in more correct responses in each of the different video modeling conditions. Future practitioners could explore using a training phase as a prerequisite method to help teach video modeling. Previous studies (e.g., Kleeberger & Mirenda, 2010; Abadir et al., 2021) have added other instructional components or procedural modifications to the intervention when video modeling alone was ineffective, but no studies to date have implemented a short training phase, then removed it.

Several limitations in the present study warrant consideration. First, the limited number of participants included in the study may limit the external validity of the results. To further confirm the results of the present study across different individuals, more studies need to be conducted with a wider range of participants.

Second, the environment in which video modeling was used must be considered. The sessions took place in the local clinic where Olivia and Spencer were receiving services. The clinic was consistently busy and often had two or three clients in one room together. Although the participants were habituated to working in an environment with loud noises from peers and instructions from multiple therapists in the background, the noise level of the

room was often louder than the sound from the videos. Therefore, the participants' therapists frequently had to find an empty room or a quieter room, if an empty room was not available. When using video modeling in the future, it is important to consider the influence of the environment on different factors, such as its possible impact on participant attending and responding. The use of video modeling in various environments should also be analyzed, such as comparing the effectiveness of using the intervention in different types of clinics versus in classrooms. In addition, it is important to consider ways to increase the volume and clarity of the videos, such as by having the participant wear headphones.

Lastly, the making of the videos proved to be more difficult and time-consuming than originally considered. As mentioned earlier, scheduling a time for the siblings to come into the clinic to record the videos required specific planning with the parents, as the siblings were typically in school during sessions. Additionally, the experimenter needed sufficient buy-in from the parents to increase their cooperation with making the videos. Some videos had to be recorded multiple times and over a few occasions with the parents as they did not pronounce the target sounds correctly the first time. This could be a serious limitation if the parents are not willing to come in to record the videos again or are not able to come in again in a timely manner. This resulted in more time being taken to record the videos as well as more time between taking baseline data and starting intervention.

The skill being targeted should also be considered as there must be a peer who is able to consistently engage in the correct response for multiple videos. The peer model used in the videos for Olivia engaged in different behaviors during the videos (e.g., spitting, hitting the card, jumping up and down in his seat) which resulted in Olivia imitating those behaviors as well when watching his videos.

Future research can expand on the findings from this study in multiple ways. First, future studies can further investigate the inclusion of a training phase. Because of the limited

amount of participants in the current study, future studies could evaluate whether other participants without prior exposure to video modeling would benefit from a training phase. Researchers could also look into assessment methods that could screen participants and identify if a training phase is needed for specific participants. Evaluating what component(s) of the training phase made the intervention effective should also be considered.

Second, future studies should consider the relationship between the learner and the video models. Olivia's mother described the sister who appeared in the videos as having the best relationship with Olivia and that she was the only person Olivia listened to consistently at home. Spencer's therapists described playing with peers as his biggest reinforcer during sessions and the peer chosen for his videos was considered one of his best friends at the clinic. It would be worthwhile to investigate the correlation between the rate of learning and the reinforcing value of the video model to the learner. No studies have been conducted that evaluated if a person considered a conditioned reinforcer has any impact on the learner's acquisition of different skills.

Third, future studies could extend the current study by evaluating whether the mastered targets were able to be maintained over extended periods of time. In the present study, the probes were conducted with Olivia the day after the video modeling phase ended. Therefore, it is unknown whether learning will maintain over the timespan of a week, month, etc. Generalization of the intervention could also be evaluated in future studies to investigate the practicality and feasibility of using video modeling to teach different skills in other settings, such as in homes or in schools.

Fourth, future researchers could interrogate what other skills would work best with video modeling. Previous research has shown the success of using video modeling with motor imitation skills and intraverbal responses, but future directions could focus on growing

the research literature to include a wider range of tacting or vocal skills, as well as expanding the population with which this intervention could be used with.

Lastly, research could be conducted to analyze methods to simplify and streamline the process of constructing the videos used. Significant time was spent during the present study to record and edit the videos, which consequently took time from beginning the intervention. It would be beneficial for future practitioners to have a more efficient system to create videos for immediate use.

The findings from the study are encouraging as the results suggest that video modeling can be used to target vocal skills in participants who have a limited vocal repertoire. The present study also opens a new avenue of research to explore the effectiveness of different video models based on their value as a conditioned reinforcer. More research needs to be conducted to confirm the possible maintaining effects of the intervention and how to generalize to other skills, populations, and settings.

References

Abadir, C., DeBar, R. M., Vladescu, J. C., Reeve, S. A., & Kupferman, D. M. (2021). Effects of video modeling on abduction-prevention skills by individuals with autism spectrum disorder. *Journal of Applied Behavior Analysis, 54*(3), 1139-1156. https://doi.org/10.1002/jaba.822

Alkatout, I., Dhanawat, J., Ackermann, J., Freytag, D., Peters, G., Maass, N., Mettler, L., & Pape, J. M. (2021). Video feedback and video modeling in teaching laparoscopic surgery: A visionary concept from Kiel. *Journal of Clinical Medicine, 10*, 163-177. https://doi-org./10.3390/jcm10010163

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (ed.), *Encyclopedia of Human Behavior* (Vol. 4, pp. 71-81). New York: Academic Press.

Cardon, T. A., & Wilcox, M. J. (2011). Promoting imitation in young children with autism: A comparison of reciprocal imitation training and video modeling. *Journal of Autism and Developmental Disorders, 41*(5), 654-666. https://doi.org/10.1007/s10803-010-1086-8

Charlop-Christy, M. H., Le, L., & Freeman, K. A. (2000). A comparison of video modeling with in vivo modeling for teaching children with autism. *Journal of Autism and Developmental Disorders, 30*(6), 537-552. https://doi.org/10.1023/A:1005635326276

Cooper, J. O., Heron, T. E., & Heward, W. L. (2019). *Applied Behavior Analysis* (3rd ed.). Pearson Education, Inc.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*(4), 519–533. https://doi.org/10.1901/jaba.1996.29-519

Egel, A. L., Richman, G. S., & Koegel, R. L. (1981). Normal peer models and autistic

children's learning. *Journal of Applied Behavior Analysis, 14*(1), 3-12. https://doi.org/10.1901/jaba.1981.14-3

Hoogerheide, V., Wermeskerken, M., Loyens, S. M., & Gog, T. (2016). Learning from video modeling examples: Content kept equal, adults are more effective models than peers. *Learning and Instruction, 44*, 22-30. https://doi.org/10.1016/j.learninstruc.2016.02.004

Hoogerheide, V., Loyens, S. M., & Gog, T. (2016). Learning from video modeling examples: Does gender matter?. *Instructional Science, 44*(1), 69-86. https://doi.org/10.1007/s11251-015-9360-y

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

National Professional Development Center on Autism Spectrum Disorders. (2010, October). *Evidence-based practice brief: Video modeling*. https://autismpdc.fpg.unc.edu/sites/autismpdc.fpg.unc.edu/files/imce/documents/VideoModeling_Complete.pdf

Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children, 8*(1), 67-76. http://www.jstor.org/stable/42898888

Table 1

Letter Blends Included in Each Condition for Olivia

SIBLING	PARENT	PEER	CONTROL
am	cl	as	at
in	it	br	on
st	sc	sp	sw

Table 2

Letters Included in Each Condition for Spencer

SIBLING	PARENT	PEER	CONTROL
F	A	B	E
J	C	K	T
O	R	S	W

Figure 1

Percent of Correct Responding Across Sessions

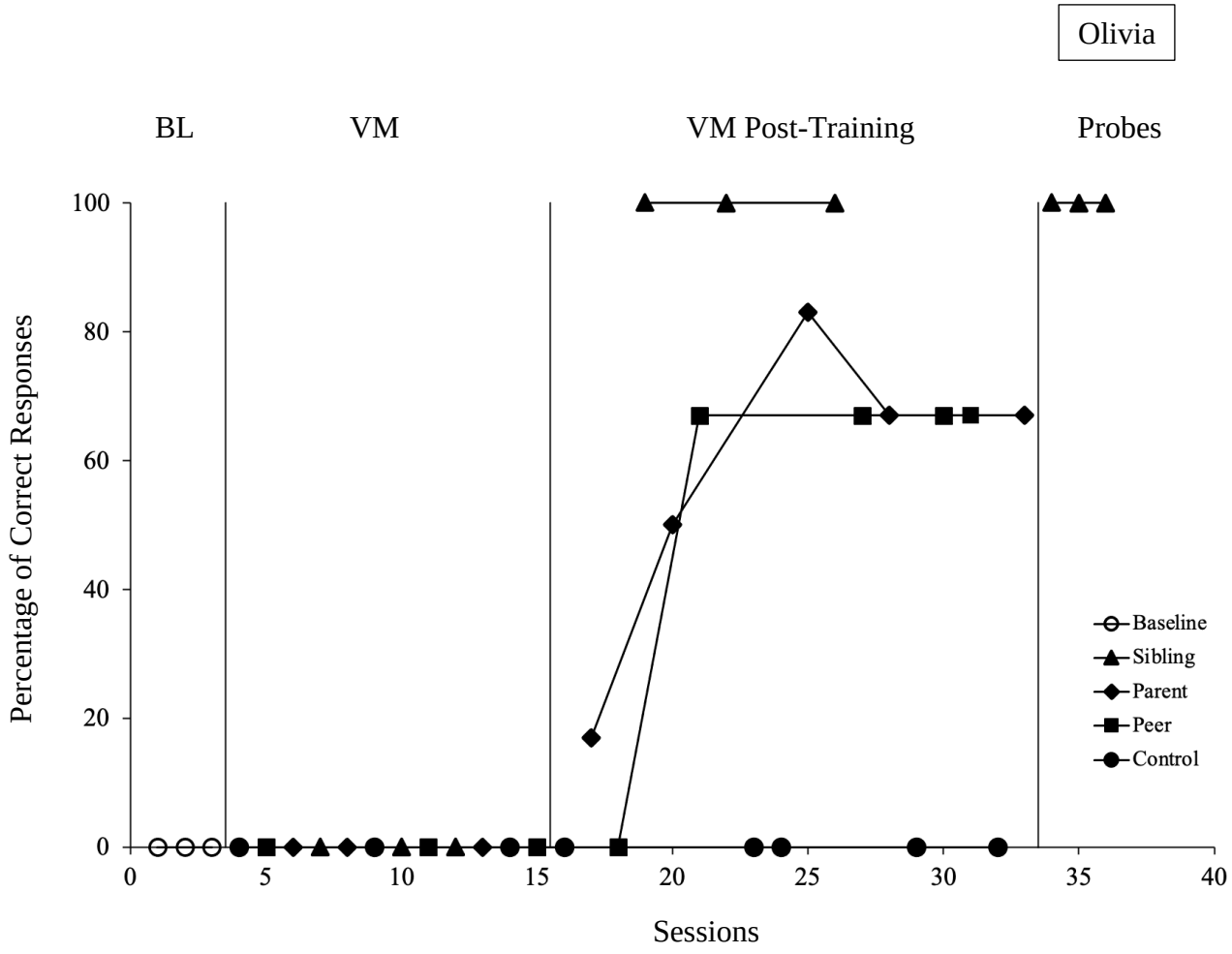
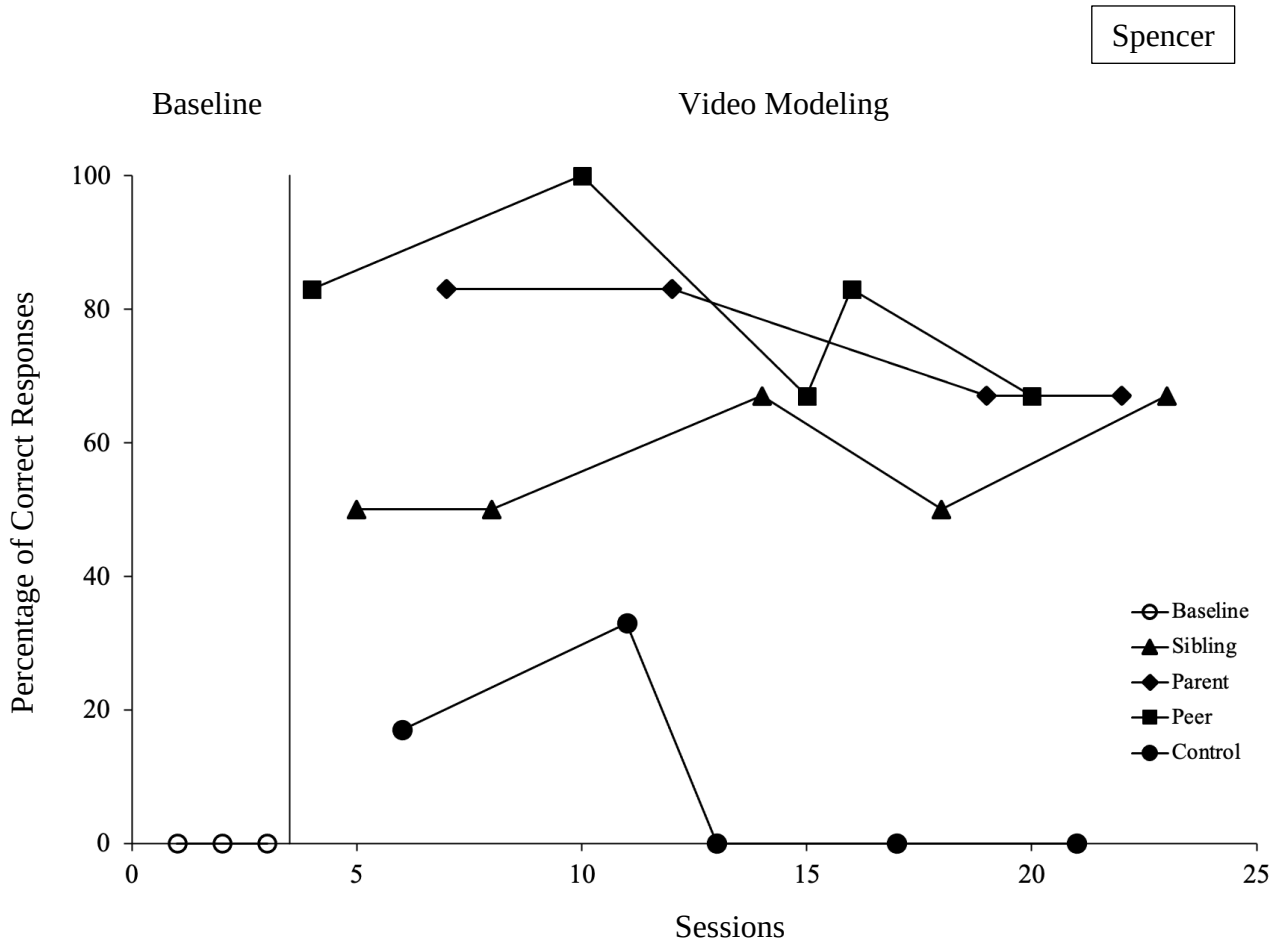


Figure 2

Percent of Correct Responding Across Sessions



Appendix A: Treatment Integrity Task Analysis Sheets

Treatment Integrity Task Analysis (Baseline Sessions)

Participant: _____

Treatment Integrity Assessor: _____

Date: _____

Session #: _____

Circle **Y** for Yes and **N** for No:

1. The experimenter presents the card with the vocal S^D, "What sound does --- make?"	Y	N
2. The experimenter waits 10 s for a response.	Y	N
3. If the participant **displays the correct response**, the experimenter provides no feedback or reinforcement. If the participant **does not display the correct response or does not engage in any response**, the experimenter provides no feedback or reinforcement.	Y	N
4. 12 trials are conducted in the session.	Y	N

Total Y: _____ Total N: _____

Percent: _____

Treatment Integrity Task Analysis
(Teaching Sessions)

Participant: _____

Treatment Integrity Assessor: _____

Date: _____

Condition: _____

Session #: _____

Circle **Y** for Yes and **N** for No:

1. Participant is seated at the table with the computer in front of them.	Y	N
2. The experimenter says, "Let's watch a video!", then plays the video.	Y	N
3. After the video ends, the experimenter presents the card and the same vocal S^D from the video.	Y	N
4. The experimenter waits 5 s for a response.	Y	N
5. If the participant **displays the correct response**, the experimenter provides verbal praise and the reinforcer for 30 s. If the participant **does not display the correct response or does not engage in any response**, the experimenter says, "Let's watch again," and plays the video again.	Y	N
6. The experimenter says, "My turn," after 30 s and obtains the reinforcer from the participant.	Y	N
7. The experimenter moves onto the next letter by saying, "Let's watch another video."	Y	N
8. 6 trials are conducted in the session.	Y	N
9. Each target is presented twice and in random order.	Y	N

Total Y: _____ Total N: _____

Percent: _____

Treatment Integrity Task Analysis
(Control Conditions)

Participant: _____

Treatment Integrity Assessor: _____

Date: _____

Session #: _____

Circle **Y** for Yes and **N** for No:

1. The experimenter presents the card with the vocal S^D, "What sound does --- make?"	Y	N
2. The experimenter waits 10 s for a response.	Y	N
3. If the participant **displays the correct response**, the experimenter provides no feedback or reinforcement. If the participant **does not display the correct response or does not engage in any response**, the experimenter provides no feedback or reinforcement.	Y	N
4. 6 trials are conducted in the session.	Y	N
5. Each target is presented twice and in random order.	Y	N

Total Y: _____ Total N: _____

Percent: _____

Treatment Integrity Task Analysis
(Probes)

Participant: _____

Treatment Integrity Assessor: _____

Date: _____

Session #: _____

Circle **Y** for Yes and **N** for No:

1. The experimenter presents the card with the vocal S^D, "What sound does --- make?"	Y	N
2. The experimenter waits 10 s for a response.	Y	N
3. If the participant **displays the correct response**, the experimenter provides no feedback or reinforcement. If the participant **does not display the correct response or does not engage in any response**, the experimenter provides no feedback or reinforcement.	Y	N
4. 3 trials are conducted in the session.	Y	N

Total Y: _____ Total N: _____

Percent: _____