A COMPARISON OF IN VIVO AND VIDEO MODEL PROMPTS ON TACT ACQUISITION

By

Kenzie Gatewood

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Applied Behavior Analysis—Master of Arts

2018

ABSTRACT

A COMPARISON OF IN VIVO AND VIDEO MODEL PROMPTS ON TACT ACQUISITION

By

Kenzie Gatewood

Children with autism spectrum disorder (ASD) often have difficulty with social communication (American Psychiatric Association, 2013). The tact is often a vital component of many social interactions (Marchese, Carr, LeBlanc, Rosati, & Conroy, 2012); thus, teaching this skill to children with ASD is an important prerequisite skill for many social and academic skills. To date, vocal-verbal tacts have been taught using an *in vivo* verbal prompt via discrete trial instruction. The current study compared two prompting procedures, *in vivo* verbal models and video model prompts to determine their effect on tact acquisition and problematic behaviors using a parallel treatments design. Three pre-school aged children with a diagnosis of ASD who received 30 hours of applied behavior analysis therapy per week participated in the study. Results of the study indicate that video model prompts are effective in teaching young children with ASD to tact, however the extent to which one prompt led to quicker acquisition of the target stimuli varied across and within participants. Implications for clinicians are summarized, and potential areas for future research are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO SYMBOLS

® Registered Trademark

KEY TO ABBREVIATIONS

BCBA-D	Board Certified Behavior Analyst- Doctoral Level
BCBA	Board Certified Behavior Analyst
EESA	Early Echoic Skills Assessment
VB-MAPP	Verbal Behavioral Milestones and Assessment Placement Program
MSWO	Multiple Stimulus without Replacement

Introduction

The tact is a verbal response that is "evoked by a particular object or event or property of an object or event" (Skinner, 1957, p. 82). Tacts are reinforced by generalized social reinforcement, making the tact a vital component of many social interactions (Marchese, Carr, LeBlanc, Rosati, & Conroy, 2012) and academic skills (e.g., identifying letters and numbers as prerequisites for reading and mathematical skills) (Sundberg & Sundberg, 2011; Michael, Palmer, Sundberg, 2011; Sundberg & Partington, 1998). For typically developing children, the acquisition of tacts and the ability to tact various internal and external stimuli often occur without explicit instruction (Smith, 2001) whereas children with autism spectrum disorder (ASD) often require explicit instruction to learn the same skills.

Discrete trial instruction has been used to effectively teach children with ASD to tact objects (Pistoljevic & Greer, 2006), actions (Williams, Carnerero, & Perez-Gonzalez 2006), and emotions (Conallen & Reed, 2016). Discrete trial tact instruction includes artificially placing a nonverbal discriminative stimulus (S^D) within the individuals line of sight, giving an *in vivo* verbal model of the nonverbal S^D name, providing generalized reinforcement contingent on the child echoing the verbal model, and presenting the next nonverbal S^D after a short intertrial interval (ITI). Over time the verbal model is faded which results in a transfer of stimulus control from an echoic response to a tact.

One reason children with ASD may not acquire tacts as readily as their same-aged peers may be due to antecedent variables such as failing to attend to the nonverbal stimulus to be tacted (Ploog, 2010; Partington, Sundberg, Newhouse, & Spengler, 1994; Marchese et al., 2012). When a visual nonverbal S^D is presented, individuals with ASD may not look at the target stimulus or attend to the verbal model. Failure to attend to one or both of these antecedent variables would likely impede instruction, lead to extended instructional time, or lead to faulty stimulus control. Previous research has evaluated the effects of using a verbal supplementary stimulus (e.g., "what is it"?) to increase the saliency of the nonverbal S^D (Marchese et. al., 2012 & LaLonde, Duenas, Neil, Wawrzonek, & Plavnick, 2017). The verbal supplementary stimulus is intended to draw the participant's attention to the nonverbal S^D to ensure that the stimulus acquires proper stimulus control. Although the addition of the supplemental stimulus may have increased the saliency of the stimulus to be tacted for some participants, there were no differential effects of the supplementary stimulus across or within participants. For the participants who did not learn to tact the stimulus it is possible that despite the addition of the verbal supplementary stimulus, the nonverbal S^D did not acquire proper stimulus control, suggesting that participants developed prompt dependency, and when the verbal model was faded, the participant was unable to correctly tact the stimulus.

Another antecedent variable that may affect the efficacy and efficiency of discrete trial tact instruction may be the type of prompt provided during instruction. Plavnick and Vitale (2016) compared the effects of *in vivo* mand training to video modeling. A video model involves an individual viewing a video that includes a demonstration of the antecedent conditions, the target behavior, and the consequences of the target behavior (Bellini & Akullian, 2007). Plavnick and Vitale (2016) found that participants acquired mands more quickly when a video model was used. Plavnick and Vitale propose two possible explanations of these findings. First, it is possible that participants learned more efficiently with a video-based prompt because on each trial they were continually exposed to the relevant contingency when observing the video, whereas during *in vivo* sessions, they were only exposed to the relevant contingency when they made a correct response (Plavnick & Vitale, 2016). Second, the video-based prompt, compared to the *in vivo* prompt, may more clearly distinguish the experimenter's role. Within traditional *in vivo* prompting scenarios, the experimenter alternates between being the speaker (i.e., providing the verbal model in which the participant is expected to echo), and then immediately switches to the role of the listener (i.e., the experimenter is then the audience to which the participant mands). The alternation of roles by the experimenter likely leads to complications in the transfer of stimulus control, as the participant is tasked with correctly echoing the target response, and also with discriminating between the two roles of the experimenter (Plavnick & Vitale, 2016). Conversely, within the context of a video-based prompt, the experimenter does not alternate between speaker and listener, which may mitigate issues of discrimination and stimulus control. The development of clear distinctions of the experimenter's role have also been found to be effective in interventions such as the picture exchange communication system (PECS; Frost & Bondy, 2002), and script fading (McClannahan & Krantz, 2005).

The purpose of the present study is to compare the effects of *in vivo* verbal model and a video model on the acquisition of tacts in young children with ASD. In addition, problem behavior will be measured during instructional sessions to evaluate the collateral effects of these interventions on problem behavior in children with ASD.

Method

Setting, Participants, and Materials

Experimental sessions were conducted in quiet rooms located within two early intensive behavioral intervention (EIBI) centers housed within preschool settings. The first room (used for participants one and two) included a child sized table and chairs, a large changing table, a sink, a large office wall divider, and two tables. The second room (participant three) included four cubicles with large office desks and computers, copy machines, a large meeting table, a couch, a child sized table and chairs, a refrigerator, bathrooms, and a sink. During all sessions participants sat at the child sized table in child sized chairs across from the experimenter.

Participants were three preschool aged children including two boys, Otis and Ross, and one girl, Lyla, all four-years-old with a diagnosis of ASD that were enrolled at the EIBI program. Each participant received individualized ABA for 30 hours per week. To be included in the study each participant demonstrated having a generalized echoic repertoire as measured by the Early Echoic Skills Assessment (EESA; Esch, 2008) and demonstrated mastery on a video model prompt pre-test. These criteria were included to ensure that the participants could echo a verbal model, a prerequisite skill for the interventions, and to ensure that the participants could attend to and imitate a video model. The video model prompt pre-test included having each participant watch two videos, the first was a video of a same aged peer dumping out a star stacker and placing the pieces back on. The children were shown this video and presented with an array of three toys. The second video was presented in the same format but involved the same age peer picking up a hat and placing it on a potato head toy. During the video model prompt pretest each participant made a correct response, by choosing the correct toy and imitating the action shown in the video. In addition, all participants could sit appropriately at a table (i.e., hands on table or lap, feet on floor, and sitting upright), attend to instructional materials for at least three minutes, and were independently using a conditioned reinforcement system (e.g., point card or token board). Participants were excluded from the study if they had demonstrated complete scores in level two of the Verbal Behavioral Milestones Assessment and Placement Program (VB-MAPP; Sundberg, 2008) for tacts, or had more than one month of tact instruction using *in vivo* or video model prompts in an attempt to eliminate previous exposure to either prompt for tact acquisition as a confounding variable.

Materials required for the sessions included data sheets, target stimuli, a Canon® video camera for recording sessions, an iPad air® to display the video models, child specific reinforcers (e.g., cheese puffs, trains, books) and a token board or point card. Otis and Ross used token economies that consisted of a laminated piece of white paper divided equally into 12 boxes. Contingent upon a correct response, participants were given a small, plastic coin that was attached to the board. Each box and coin had a small piece of Velcro® so that the coins didn't move. For Lyla, a point card was used. The point card consisted of a laminated strip of white paper that was divided into a table with two rows and twelve columns numbered 1-12. Contingent upon a correct response, the experimenter drew a tally in a box using a dry erase marker. When all 12 boxes had a coin or tally, the experimenter provided the participant with a backup reinforcer.

Dependent Variables

There were two dependent variables in this study. The primary dependent variable was the frequency of independent, correct tacts during instructional sessions. Data were collected on the responses emitted by the participants during each session. Participant responses were coded as one of the following: (a) correct response, defined as a participant emitting the correct tact within six seconds of the presentation of the nonverbal discriminative stimulus; (b) an incorrect response, defined as the participant emitting a response that does not match the nonverbal discriminative stimulus presented (e.g., said “gear” in the presence of a goat); (c) a prompted correct response, defined as the participant echoing the experimenter’s verbal model within six seconds (*in vivo*) (e.g., the participant echoed “goat” after the experimenter presented a goat while giving the verbal model “goat”) or within six seconds of the experimenter presenting the video model (e.g., the participant said “goat” in the presence of a goat after viewing a video model in which a goat was tacted), and (d) a prompted incorrect response, defined as the participant emitting a response that did not match the experimenter’s verbal model (*in vivo*) (e.g., the participant said “kite” after the experimenter presented a goat while giving the verbal model “goat”) or the child said something other than the nonverbal discriminative stimulus named after viewing the video model (e.g., the experimenter placed a goat on the table and displayed a video model of a child tacting a goat, after viewing the video the child looks at the goat and says “bee”). A frequency of independent, correct responses was obtained by counting the number of trials in which the participant engaged in independent, correct responses completed after each instructional session.

The second dependent variable was the percentage of 10 s intervals in which a participant engaged in problem behavior. Individual response topographies were selected and defined for each participant. Following sessions, the experimenter scored the session from video using a partial interval recording system (Cooper, Heron, & Heward, 2007). The experimenter watched an instructional session and for each 10 s interval recorded a "+" if the participant engaged in problem behavior at any point during the interval, or a "-" if no problem behavior was observed during the interval. Data were converted to a percentage by dividing the number of intervals in which the target behavior occurred by the total number of intervals and multiplying by 100%, this process was completed for each topography of problem behavior per participant.

Discrete trial instruction is often completed in a highly structured teaching arrangement which can evoke problematic behaviors such as stereotypy, physical non-compliance, and eloping (Roxburgh & Carbone, 2012) which can impede learning. Although both types of prompts (*in vivo* and video model) were completed in the same instructional arrangement, proponents of video modeling suggest that video models may be an abolishing operation to engage in problematic behaviors (Charlop-Christy, Le, & Freeman, 2000). To that end, we wanted to evaluate the effects of each prompt type on problematic behavior.

Problem behavior definitions for each participant were as follows:

Otis:

- Physical non-compliance: Any instance in which the participant was presented with stimuli to tact or an instruction to get ready and he physically (i.e., runs from the area, slides out of his chair onto the floor, lays on the floor) does not complete the task. Examples include, when a stimulus was placed on the table and Otis tipped his body sideways and slid out of his chair, when told to “get ready” and Otis stood up and ran from the table, and when told to “get ready” and Otis began to push the table toward the experimenter or pull the table to himself without letting go of the table.

- Vocal Stereotypy: Any instance in which the participant engaged in inappropriate (i.e., vocalizations that do not relate to the target stimulus or the intervention) and/or repetitive vocalizations. Examples include, when Otis began to sing nursery rhymes, when Otis began to emit unintelligible vocalizations repeatedly such as “tickatickaticka”.

Ross:

- Physical non-Compliance: Any instance in which the participant was presented with stimuli to tact or an instruction to get ready and he physically (i.e., runs from the area, slides out of his chair onto the floor, lays on the floor, lays on the table) does not complete the task. Examples include, when instructed to “get ready” and Ross would stand up and lie across the table, running from the table when presented with a stimulus or told to “get ready”, or when Ross would stand up, pick up his chair and hold it near the table when told to “get ready” or presented with a stimulus.
- Motor Stereotypy: Any instance in which the participant bounced in his chair, his bottom leaving the seat for more than two consecutive instances, the participant pointed to or reached for objects in the room (non-target items), flapped his

hands, or placed his hands on his face on either side of his mouth while engaging in vocal stereotypy. Examples include, bouncing up and down his chair four consecutive times, flapping his hands in the air, and wringing his hands together with his arms outstretched.

- Vocal Stereotypy: Any instance in which the participant engages in inappropriate (i.e., vocalizations that do not relate to the target stimulus or the intervention) or repetitive vocalizations. Examples include, making a loud "ee" noise in the back of his throat and making unintelligible vocalizations that did not match previously taught tacts or mands.

Lyla:

- Physical non-Compliance: Any instance in which the participant was presented with stimuli to tact or instructed to get ready and she physically (i.e., runs from the area, slides out of her chair onto the floor, lays on the floor) does not complete the task. Examples include, standing up on top of her chair when told to "get ready" or when presented with a stimulus, standing up and walking around the table and experimenter when told to "get ready" or presented with a stimulus, and turning backwards in her chair when instructed to "get ready" or when presented with a stimulus.
- Stereotypy: Any instance in which the participant engages in vocalizations that are unrecognizable or unrelated to the target stimulus or engages in ear plugging or hand flapping. Examples include plugging her ears with her thumbs, hand flapping, and standing up and dancing while repeating the name of a previously presented stimulus.

Interobserver Agreement

Interobserver agreement (IOA) was collected during 31%, 32%, and 34% of sessions by an independent data collector for each dependent variable for Otis, Ross, and Lyla, respectively. The secondary data collectors were a doctoral level board certified behavior analyst (BCBA-D), two lead behavior technicians with three years of experience working at the EIBI center, and a behavior technician who worked at the EIBI center for two years. The secondary data collectors were trained by reviewing the data collection codes as well as examples of each code via video. The secondary data collector was considered reliable once the two data collectors demonstrated 80% reliability across three instructional sessions for each dependent variable. An agreement was recorded if the primary and secondary data collectors both scored a trial or interval using the same response code. A disagreement was recorded if the primary and secondary observer recorded different response codes per trial or interval. Data were compared, trial-by-trial, and interobserver agreement was calculated using the point-by-point method (Ayres & Ledford, 2014). The average IOA percentage for the first dependent variable (i.e., tact acquisition) for Otis was 93% (range: 75-100%), average IOA for the second dependent variable (i.e., problematic behavior) for physical non-compliance was 92% (range: 67-100%) and for vocal stereotypy was 79% (range: 63-100%). The average IOA percentage for the first dependent variable for Ross was 96% (range: 75-100%), average IOA for the second dependent variable was 95% (range: 79-100%), 90% (range: 61-100%), and 91% (range: 71-100%) for physical non-compliance, motor stereotypy, and vocal stereotypy, respectively. The average IOA percentage for the first dependent variable for Lyla was 97% (range: 80-100), average IOA for the second dependent variable was 85% (range: 67-100%) and 80% (range: 67-100%) for physical non-compliance and stereotypy, respectively.

Selection of Discriminative Stimuli and Response

To ensure all discriminative stimuli were of equal difficulty, the logistical method described by Ledford and Gast (2014) was used. Figure 1 provides visual representation of the discriminative stimuli selection process. Sets of discriminative stimuli were determined by generating a list of 18 one syllable nouns (Otis and Ross), and two syllable nouns (Lyla). These 18 stimuli were numbered, and then randomly assigned to one of the two intervention conditions (i.e., *in vivo* or video model prompts) using a random number generator. Once stimuli were assigned to an intervention, they were numbered again, and then randomly assigned to sets. Once potential sets were made, they were reviewed by a BCBA-D who is dual certified as a speech-language pathologist who provided recommendations to make the sets of equality difficulty. A primary recommendation was to ensure that the targets were predominantly composed of consonant-vowel-consonant (CVC) words when possible, and if consonant-vowel words were included, to ensure that they were distributed equally between the two prompt types (e.g., bee and tie were included as targets for one participant, it was recommended that one of these words be assigned to video model prompts, and the other to the *in vivo* prompt training condition). These recommendations were followed, however during baseline some of the participants tacted items requiring the items to be removed and replaced with new stimuli from the target list, at which time new sets were determined based on previous recommendations.

Experimental Design

A parallel treatment (Wolery, Gast, & Ledford, 2014) across sets was used to compare the effects of video model and *in vivo* prompts on tact acquisition and problem behavior during instructional sessions. The application of the design in the present study was as follows: the baseline condition remained in place until visual inspection revealed no trend in the frequency of correct tacting of nonverbal discriminative stimuli, at which time treatment was implemented for the first set of stimuli (i.e., Set A), while the remaining sets remained in baseline. The two independent variables were alternated across sessions. The alternation of interventions continued until the participant reached mastery-level responding for the set under both interventions, or until the participant completed one and a half times the number of sessions that it took for mastery-level responding to be achieved under the more effective intervention (e.g., if the participant reached mastery-level responding on the first set of stimuli under the video model prompt in five sessions, experimental sessions continued for both prompt types until the participant also reached mastery-level responding under *in vivo* prompts or when a total of eight sessions were completed, which ever came first) (Wolery, et. al., 2014). Probes were then conducted for each set, in random order (i.e., the probe for set A was not always conducted first, followed by sets b and c). Intervention was then implemented for set B, while set C remained in baseline. Once a participant demonstrated mastery-level responding for set B, probes were conducted again for all sets (i.e., sets A-C). This continued until a participant mastered each set of stimuli, at which point maintenance probes were conducted at two and four weeks after the post probes for set C.

This experimental design was selected for multiple reasons. First, the parallel treatments design is useful when comparing the efficiency of interventions across multiple sets of stimuli. Additionally, the staggering of sets for each participant allows for intra-subject replication (Wolery, et. al., 2014). The baseline condition built into the design allows for the assessment of threats to internal validity. Further, the repetition of probe conditions also assesses threats to internal validity prior to sets being placed into intervention, while also permitting an analysis of maintenance for each intervention on the target stimuli for previously mastered sets.

Procedure

Baseline. The primary goal of the baseline condition was to establish that the participants were not able to tact nonverbal discriminative stimuli prior to intervention. Prior to the start of each baseline session a brief multiple stimulus without replacement preference assessment (MSWO; Higbee, Carr, & Harrison, 2000) was conducted to determine a putative reinforcer that the participant received following the session. All sessions consisted of 12 trials (six target stimuli, three from each intervention presented two times each) and lasted approximately five min. One to three sessions were completed each day, up to four days each week (depending on participant attendance). Prior to the start of the study, each participant had learned to make a ready response in the presence of the verbal instruction “ready.” The ready response was defined as the participant sitting with their hands on the table or in their lap and looking at the therapist. To create an establishing operation to respond during baseline sessions, the experimenter would provide the instruction “ready” in-between presenting tact trials. Tokens or points were provided contingent on the participant making a correct ready response. The experimenter ensured that the 12 tact trials were presented before participants traded in their points or token for the backup reinforcer.

A session began with the experimenter providing an instruction of what would happen during the session. Specifically, the experiment said, “I am going to put some things on the table, and I want you to tell me what they are”. Following the statement, trials consisted of the experimenter saying “ready” to obtain the participant’s attending. The experimenter then presented the nonverbal S^D by placing the stimulus on the table in front of the participant and waited 6 s for the participant to make a response. After 6 s, the experimenter removed the stimulus and presented the next trial after a three to five second ITI. If a participant correctly tacted a stimulus during baseline, social praise and a token or point were provided. That stimulus was removed and replaced with an item from the target list and another baseline session was conducted until there were three sessions in which the participant emitted zero correct responses.

***In Vivo* Prompts.** *In vivo* prompt sessions were identical to baseline except for the addition of verbal model. A progressive prompt delay was used across treatment sessions (Walker, 2008). Teaching began at a 0-s prompt delay in which the experimenter simultaneously presented the nonverbal S^D and the verbal model (e.g., held up a goat and said, “goat”). If the participant correctly echoed the verbal model within six seconds the experimenter provided a token and social praise (e.g., “yeah it is a goat!”). The time delay was increased once a participant correctly echoed the verbal model on 11 of 12 trials for one session. During the 3-s delay, the experimenter presented the nonverbal S^D, waited three seconds, and then presented the verbal model if the participant did not make a response. The participant had 6 s to echo the verbal model and doing so was reinforced with a token or point and social praise. The same criteria were used to move to a 6-s time delay, and independent. During any sessions, if the participant did not make a response within 6 s of the verbal prompt (during prompted sessions) or of the presentation of the nonverbal S^D (unprompted sessions), the experimenter tacted the item, removed the stimulus, and presented the next trial after a three to five second ITI. If the participant echoed the experimenter's tact, no differential outcome was provided. Mastery-level responding was defined as the participant independently tacting the target stimuli with 80% accuracy for two consecutive sessions under one of the intervention conditions. This ensured that participants tacted each discriminative stimulus correctly at least two times within a teaching session.

Video Model Prompts. Video model sessions were identical to baseline except for the addition of video models. A progressive prompt delay was used across treatment sessions (Walker, 2008). Teaching began at a 0-s prompt delay in which the experimenter simultaneously presented the nonverbal S^D, an iPad®, and the instruction "watch the video". The iPad® showed a brief video of a same-aged peer tacting the target stimulus. Each video clip included the same sequence which started with a zoomed-in view of the stimulus to be tacted on a table, the video then zooms out to show a same-aged peer and the experimenter sitting at a child sized table at which time the child tacts the stimulus, the experimenter says, "That is a [stimulus name]" and provides a token (videos or Otis and Ross) or a point (Lyla). When the clip was finished, the experimenter removed the iPad and waited 6 s for the participant to make a response. If the child correctly tacted the stimulus the experimenter provided social praise (e.g., “yeah it is a ____!”) and delivered a token or point. The time delay increased once the participant correctly tacted 11 of 12 trials for one session after watching the video. During the 3-s delay, the experimenter presented the stimulus, waited three seconds for the participant to tact the stimulus, if the participant did not make a response, the experimenter presented the iPad®, said “Watch the video” and played the clip. The same criteria were used to move to a 6-s time delay, and independent. The child mastered a set once they had independently tacted the target stimuli with 80% accuracy for two consecutive sessions. This ensured that participants tacted each discriminative stimulus correctly at least two times within a teaching session.

During all prompt delays, incorrect responses and no responses were followed by the experimenter providing the tact of the item, removing the stimulus, and then conducting the next trial with a 6-12 second ITI. If the participant echoed the experimenter’s tact, no differential outcome was provided. The ITI interval was longer during video model prompts because of the time it took for the experimenter to close out of the previously played clip and select the clip to be played for the next trial.

Probes. Probe sessions were identical to baseline except for the addition of intermittent reinforcement for correct responses of target tact stimuli in addition to reinforcement for “ready” behavior. This was done to encourage correct responding and minimize non-responding (Grow and LeBlanc, 2013; & Reichow & Wolery, 2009). For example, when set A was mastered (after a participant mastered three stimuli following *in vivo* prompts and video model prompts, or if one set of stimuli were mastered under one independent variable and one and a half times the number of sessions were completed for the other independent variable) all three sets were assessed.

Procedural Integrity

Procedural integrity was collected during 39%, 39%, and 37% of sessions for Otis, Ross, and Lyla, respectively. Specific experimenter behaviors for *in vivo* model sessions included, (a) providing the general instruction about what was going to happen during the session (this only happened prior to the first trial); (b) gaining the participant's attending by saying "ready"; (c) presenting the correct verbal model at the appropriate time delay; (d) providing social praise and a token within one to three seconds if the participant engaged in a correct response; (e) providing a correct tact of the stimulus and withholding token reinforcement if an error or no response occurred; (f) affirming the correct response and withholding token reinforcement if a participant engaged in a self-correction; (g) presenting the next trial within three to five seconds; and (h) providing the backup reinforcer identified through the MSWO within three seconds of the participant trading in their filled token board.

Specific experimenter behaviors for video modeling sessions included, (a) providing the general instruction about what was going to happen during the session (this only happened prior to the first trial); (b) gaining the participant's attending by saying "ready"; (c) presenting the correct video model at the appropriate time delay; (d) providing social praise and a token within one to three seconds if the participant engaged in a correct response; (e) providing a correct tact of the stimulus and withholding token reinforcement if an error or no response occurred; (f) affirming the correct response and withholding token reinforcement if a participant engaged in a self-correction; (g) presenting the next trial within six to twelve seconds; and (h) providing the backup reinforcer identified through the MSWO within three seconds of the participant trading in their filled token board.

The secondary observer scored each of the experimenter's discrete behaviors on each trial. The observer recorded a ("+") if a behavior was implemented correctly and a ("-") if a behavior was implemented incorrectly. Following a session, treatment integrity was calculated by dividing the number of correctly implemented experimenter behaviors by the total number of experimenter behaviors within the session and multiplying by 100%. The average PI percentages were 99% (range: 83-100%), 98% (range: 92-100%), and 99% (range: 98-100%) for Otis, Ross, and Lyla, respectively.

Results

Figures 2, 3, and 4 depict the results for correct independent tacts during baseline, intervention, post probes, and two and four-week maintenance probes for Otis, Ross, and Lyla respectively. Figures 5, 6, and 7 depict the number of tacts emitted during probe and maintenance sessions for Otis, Ross, and Lyla, respectively. The bar graphs identify the number of correct responses made that were taught under each prompt type. Figures 8, 9, 10, 11, 12, 13, and 14 depict the results for problem behavior across the two independent variables for Otis, Ross, and Lyla respectively, during baseline, intervention, post probes, and two and four-week maintenance probes.

Tact Acquisition

During baseline, Otis did not emit correct tacts for any set. For set A, he met mastery-level responding after five sessions under the *in vivo* prompt; whereas the 0-s time delay under the video model prompt could not be faded until after five sessions, at which time an upward trend was observed, but mastery-level responding was not observed before the termination criterion was met. During the first probe he tacted 9 (of 12) trials correctly (the twelve trials were composed of three stimuli taught under video model prompts, and three stimuli taught under *in vivo* prompts, presented two times each) for set A; six of which were stimuli taught under the *in vivo* prompt (meaning he tacted each of the three stimuli correctly on both presentations), and three were stimuli taught under the video model prompt (however he only tacted two of the three stimuli taught). For Sets B and C, he did not emit correct tacts during the first post probe. When intervention was implemented for set B, an upward trend was observed under both prompts and mastery-level responding was achieved under the *in vivo* prompt within six sessions. One session of mastery was achieved under the video model prompt however, the termination criterion was met before a second demonstration was observed. During the second post probe he responded correctly on eight trials; six of which were stimuli taught under the *in vivo* prompt, and two that were taught under the video model prompt (he tacted two of the three stimuli). For set A he tacted 9 (of 12) trials correctly; six of which were stimuli taught under the *in vivo* prompt, and three were stimuli taught under the video model prompt (he tacted two of the three stimuli). He did not correctly tact any stimuli in set C. When intervention was implemented for set C, an immediate upward trend was observed under the *in vivo* prompts, and mastery-level responding was achieved within four sessions. Although an upward trend was observed under the video model prompt, the termination criterion was met after five sessions. Although mastery-level responding was not met under the video model prompt, during the third post probe, he tacted 12 of 12 trials correctly. For set A he tacted 6 (of 12) trials correctly; four of which were stimuli taught under the *in vivo* prompt (he tacted two of the three stimuli), and two were stimuli taught under the video model prompt (he tacted one of the three stimuli). For set B he correctly tacted on 5 (of 12) trials; four of which were stimuli taught under the *in vivo* prompt (he tacted two of the three stimuli), and one was a stimulus taught under the video model prompt. During two-week maintenance probes he tacted eight, five, and 10 trials correctly for sets A, B, and C respectively. For set A, five of the correct responses were taught under the *in vivo* prompt and he tacted each stimulus correctly at least once, the remaining three correct responses were stimuli taught under the video model prompt, and he tacted two of the three stimuli taught. For set B all of the correct responses were taught under the *in vivo* prompt and he tacted each stimulus at least once. For set C six of the correct responses were taught under the *in vivo* prompt, the remaining four correct responses were stimuli taught under the video model prompt, and he tacted two of the three stimuli taught. During four-week maintenance probes he tacted one, two, and six trials correctly for sets A, B, and C, respectively. For set A the correct response was a stimulus taught under the video model prompt. For set B all of the correct responses were stimuli taught under the *in vivo* prompt and he tacted two of the three stimuli taught. For set C three of the correct responses were stimuli taught under the *in vivo* prompt and he tacted two of the three stimuli taught, the remaining three correct responses were stimuli taught under the video model prompt, and he tacted two of the three stimuli taught.

During baseline, Ross did not emit correct tacts for sets A and B, for set C, he emitted one correct tact during the third session of baseline, this stimulus was removed, and three more baseline sessions were conducted in which he did not emit any correct tacts. For set A, he met mastery-level responding after five sessions under the *in vivo* prompt. Whereas for the video model prompt the 0-s time delay could not be faded before termination criteria were met. During the first probe he tacted 7 (of 12) trials correctly for set A; six of which were stimuli taught under the *in vivo* prompt, and one was a stimulus taught under the video model prompt. For Sets B and C, he did not emit correct tacts during the first post probe. When intervention was implemented for set B, an upward trend was observed under both prompts, mastery-level responding was achieved under both the video model and *in vivo* prompt within nine and ten sessions, respectively. During the second post probe he tacted 11 (of 12) trials correctly; five of which were stimuli taught under the *in vivo* prompt (he tacted each of the three stimuli taught once), and six were stimuli taught under the video model prompt. For set A he emitted correct responses on six trials, all of which represented stimuli taught under the *in vivo* prompt. He did not correctly tact any stimuli in set C. When intervention was implemented for set C, an immediate upward trend was observed under the *in vivo* prompt, and mastery-level responding was achieved within five sessions. Although an upward trend was also observed under the video model prompt, termination criteria was met after eight sessions. Despite mastery-level responding not being met under the video model prompt, during the third post probe for set C, he tacted 9 (of 12) trials correctly; six of which were stimuli taught under the *in vivo* prompt, and three were stimuli taught under the video model prompt (he tacted each of the three stimuli taught correctly once). During the post probe for set A he tacted 6 (of 12) trials correctly, all of which were stimuli taught under the *in vivo* prompt. For set B he tacted 11 (of 12) trials correctly; five of which were stimuli taught under the *in vivo* prompt (he tacted each of the stimuli correct once), and six were stimuli taught under the video model prompt. During two-week maintenance probes he tacted four, three, and three trials correctly for sets A, B, and C respectively. For set A, all correct responses were taught under the *in vivo* prompt and he tacted two of the three stimuli taught. For set B all of the correct responses were taught under the video model prompt and he tacted two of the three stimuli taught. For set C two of the correct responses were taught under the *in vivo* prompt and he tacted two of the three stimuli taught, the remaining correct response was taught under the video model prompt. During four-week maintenance probes he tacted four, seven, and eight trials correctly for sets A, B, and C respectively. For set A, all correct responses were taught under the *in vivo* prompt and he tacted two of the three stimuli taught. For set B four of the correct responses were taught under the *in vivo* prompt and he tacted two of the three stimuli taught, the remaining three stimuli were taught under the video model prompt and he tacted two of the three target stimuli taught. For set C five of the correct responses were taught under the *in vivo* prompt and he tacted all three stimuli taught correctly one time, the remaining three correct responses were taught under the video model prompt, and he tacted two of the three stimuli taught.

During baseline, Lyla did not emit correct tacts for sets A and B, for set C, she emitted one correct tact during the third session of baseline, this stimulus was removed, and three more baseline sessions were conducted in which she did not emit any correct tacts. For set A, she met mastery-level responding after four sessions under the *in vivo* prompt. Whereas for the video model prompt the 0-s time delay could not be faded until the fifth session, at which time an upward trend was observed, but mastery-level was not achieved before termination criteria were met. During the first probe she tacted 8 (of 12) trials correctly for set A; six of which were stimuli taught under the *in vivo* prompt, and two were stimuli taught under the video model prompt (she tacted one of the three stimuli taught). For Sets B and C, she did not emit correct tacts during the first post probe. When intervention was implemented for set B, an upward trend was observed under both prompts, mastery-level responding was achieved under both the video model and *in vivo* prompt within four and five sessions, respectively. During the second post probe she tacted 9 (of 12) trials correctly; five of which were stimuli taught under the *in vivo* prompt (she tacted each of the three stimuli taught once), and four were stimuli taught under the video model prompt (she tacted two of the three stimuli taught). For set A she tacted 9 (of 12) trials correctly; five of which were stimuli taught under the *in vivo* prompt (she tacted each of the three stimuli taught correctly one time), the remaining four stimuli were taught under the video model prompt (she tacted two of the three stimuli taught). She did not emit any correct responses for set C. When intervention was implemented for set C, an immediate upward trend was observed under both the *in vivo* and video model prompt, and mastery-level responding was achieved within five and six sessions, respectively. During the third post probe for set C, she tacted 12 (of 12) trials correctly. For set A she tacted 10 (of 12) trials correctly; six of which were stimuli taught under the *in vivo* prompt, and four of which were taught under the video model prompt (she tacted two of the three stimuli taught). For set B she tacted 11 (of 12) trials correctly; six of which were stimuli taught under the *in vivo* prompt, and five were stimuli taught under the video model prompt (she tacted each of the stimuli correct once). During two-week maintenance probes she tacted 9, 12, and 12 trials correctly for sets A, B, and C respectively. For set A six of the correct responses were taught under the *in vivo* prompt, and three were taught under the video model prompt (she tacted two of the three stimuli taught). During four-week maintenance probes she tacted 10, 9, and 12 trials correctly for sets A, B, and C respectively. For set A, six correct responses were taught under the *in vivo* prompt and four stimuli were taught under the video model prompts, and she tacted two of the three stimuli taught. For set B five of the correct responses were taught under the *in vivo* prompt and she tacted each of the three stimuli taught correctly on the first presentation, the remaining four stimuli were taught under the video model prompt and she tacted two of the three target stimuli taught.

Problem Behavior

During baseline Otis engaged in on average, 8% (range: 0-24%), 21% (range: 0-50%), and 19% (range: 0-57%) of intervals with physical non-compliance, for sets A, B, and C, respectively. Under *in vivo* prompts Otis, on average, engaged in physical non-compliance for 8% (range: 0-38%), 24% (range: 0-46%), and 8% (range: 0-25%) of intervals for sets A, B, and C, respectively. Under the video model prompt Otis, on average, engaged in 6% (range: 0-19%), 13% (range: 0-32%), and 3% (range: 0-11%) of intervals with physical non-compliance, for sets A, B, and C, respectively. During all post and maintenance probes, low levels of physical non-compliance were observed.

During baseline Otis engaged in, on average, 7% (range: 0-21%), 6% (range: 0-13%), and 56% (range: 40-83%) of intervals with vocal stereotypy, for sets A, B, and C, respectively. Under *in vivo* prompts Otis engaged in on average, 5% (range: 0-19%), 20% (range: 0-50%), and 11% (range: 0-31%) of intervals with vocal stereotypy, for sets A, B, and C, respectively. Under the video model prompt Otis engaged in on average, 9% (range: 0-25%), 8% (range: 0-18%), and 6% (range: 3-14%) of intervals with vocal stereotypy, for sets A, B, and C, respectively. During all post probes low levels of vocal stereotypy were observed, however during maintenance probes increased levels of vocal stereotypy occurred.

Ross engaged in on average, 8% (range: 0-23%), 10% (range: 0-29%), and 34% (range: 0-76%) of intervals with physical non-compliance, for sets A, B, and C, respectively during baseline. Under the *in vivo* prompts, Ross engaged in on average, 0%, 0% and 4% (range: 0-32%) of intervals with physical non-compliance, for sets A, B, and C, respectively. Under the video model prompt Ross engaged in on average, 13% (range: 0-31%), 8% (range: 0-23%), and 35% (range: 13-85%) of intervals with physical non-compliance, for sets A, B, and C respectively. During all post and maintenance probes, low levels of physical non-compliance were observed.

During baseline Ross engaged in on average 28% (range: 13-38%), 27% (range: 8-60%), and 10% (range: 0-25%) of intervals with motor stereotypy, for sets A, B, and C, respectively. Under *in vivo* prompts Ross engaged in on average 9% (range: 0-38%), 5% (range: 0-11%), and 37% (range: 0-82%) of intervals with motor stereotypy, for sets A, B, and C, respectively. Under the video model prompt Ross engaged in on average 22% (range: 0-81%), 24% (range: 14-36%), and 37% (range: 18-74%) of intervals with motor stereotypy, for sets A, B, and C, respectively. Trends of motor stereotypy were similar under both prompt types, for sets B and C, however for set A, there was an increasing trend under video model prompts, but not under *in vivo* prompts. During all post and maintenance probes low levels of motor stereotypy were observed, however during the two-week maintenance probe for set A, he engaged in motor stereotypy 44% of intervals, whereas during the same probe for sets B and C he engaged in motor stereotypy for nearly zero percent of intervals.

During baseline Ross engaged in on average 20% (range: 13-31%), 21% (range: 8-40%), and 12% (range: 0-25%) of intervals with vocal stereotypy, for sets A, B, and C, respectively. Under *in vivo* prompts Ross engaged in on average 4% (range: 0-31%), 10% (range: 0-22%), and 10% (range: 0-45%) of intervals with vocal stereotypy, for sets A, B, and C, respectively. Under the video model prompt Ross engaged in on average 35% (range: 0-69%), 26% (range: 0-54%), and 39% (range: 11-77%) of intervals with vocal stereotypy, for sets A, B, and C, respectively. For set A, he engaged in higher levels of vocal stereotypy during video model prompts, whereas under *in vivo* prompts he never engaged in vocal stereotypy save session five, in which he engaged in vocal stereotypy for 31% of intervals. During set B, higher rates of vocal stereotypy were observed during the first two sessions under video model prompts, however not during subsequent sessions. For set C he engaged in higher levels of vocal stereotypy during video model prompts, whereas under *in vivo* prompts he engaged in low levels, save session five, in which he engaged in vocal stereotypy for 45% of intervals. During all post and maintenance probes low levels of vocal stereotypy were observed.

During baseline Lyla engaged in on average, 38% (range: 23-67%), 83% (range: 76-96%), and 54% (range: 18-87%) of intervals with physical non-compliance, for sets A, B, and C, respectively. Under *in vivo* prompts Lyla engaged in on average, 12% (range: 0-50%), 62%

(range: 50-81%), and 24% (range: 0-60%) of intervals with physical non-compliance, for sets A, B, and C, respectively. Under the video model prompt Lyla engaged in on average 22% (range: 0-67%), 27% (range: 7-52%), and 6% (range: 0-11%) of intervals with physical non-compliance, for sets A, B, and C, respectively. Trend and average level of physical non-compliance were similar under both prompt types, for sets A and C, however for set B, trend was the same, but the average level of physical non-compliance was higher under the *in vivo* prompts. During all post and maintenance probes, percentages of non-compliance were similar to that of baseline and intervention levels.

During baseline Lyla engaged in on average 58% (range: 47-64%), 53% (range: 0-94%), and 40% (range: 0-91%) of intervals with stereotypy, for sets A, B, and C, respectively. Under *in vivo* prompts Lyla engaged in on average 70% (range: 45-100%), 67% (range: 50-95%), and 97% (range: 88-100%) of intervals with stereotypy, for sets A, B, and C, respectively. Under the video model prompt Lyla engaged in on average 40% (range: 24-56%), 77% (range: 56-95%), and 85% (range: 57-100%) of intervals with stereotypy, for sets A, B, and C, respectively. Trends of stereotypy were flat under both prompt types for all sets, however there was a difference in average level of stereotypy, although one prompt was not observed to consistently lead to higher levels of stereotypy. With the exception of the first set A post probe, high levels of stereotypy were observed during post probes and maintenance probes.

Discussion

The present study sought to compare the efficacy of video model and *in vivo* prompts on the acquisition of tacts for young children with a diagnosis of ASD, as well as to compare the collateral effects of each prompt type on problematic behavior. For Otis, *in vivo* prompts led to quicker acquisition of mastery-level responding across all sets. For Ross, *in vivo* prompts led to quicker acquisition of mastery-level responding for sets A and C, and video model prompts led to quicker acquisition of mastery-level responding for set B. Lastly, for Lyla, *in vivo* prompts led to quicker acquisition of mastery-level responding for set A, but video model prompts led to quicker acquisition of mastery-level responding for sets B and C. Taken together, the results across each set and post probes demonstrate that video model prompts are an efficacious instructional practice to teach young children with ASD to tact.

Across all participants, *in vivo* prompting procedures led to quicker acquisition of mastery-level responding for the first set of stimuli. There are two possible explanations for this pattern across participants. First, it is possible that although each participant demonstrated the ability to imitate a video prior to the start of the instruction, their limited experience with video models as a prompt for verbal behavior may have led to the need for repeated exposure of the contingency via the video prompts. Second, although the participants did not have prior exposure to *in vivo* prompts for tact training, they each had experience with *in vivo* verbal prompts to acquire mands for tangibles and edibles.

Despite quicker acquisition of mastery-level responding during the first set, results of subsequent sets demonstrate that participants required fewer trials at more intrusive prompt levels (i.e., 0-s time delay) prior to prompts being faded. This allowed for the participants to engage in independent responding more quickly than what was demonstrated in the first set. For Ross and Lyla, video model prompts led to quicker acquisition of mastery-level responding during their second set, suggesting that as participants contacted the contingency and received reinforcement for correctly echoing the video model prompts, the effectiveness of the prompt increased.

In addition, the results of post probes indicate that each participant acquired tacts under a specific prompting procedure even when they did not achieve mastery-level responding. For Otis, mastery-level responding was not met under the video model prompts for sets A-C, however, during post probe sessions he emitted correct tacts of stimuli taught under video model prompts. Similarly, Ross emitted tacts from the less effective prompting procedure during sets A and C during post probes, despite termination criteria being met prior to mastery-level responding for *in vivo* prompts and video model prompts for sets A and C, respectively. In the same way, Lyla emitted correct tacts of stimuli taught under video model prompts during the post probe for set A, despite the termination criterion being met prior to mastery-level responding. These results suggest that both prompt types were effective for tact acquisition even in the absence of mastery-level responding for all participants.

Results of problematic behavior indicate some modest differentiation between the two prompting procedures, however the separation of data paths does not appear significant across any participant or set to indicate that one prompting procedure functioned as a motivating operation for problematic behavior. The lack of separation may have occurred for a few reasons, first, the teaching arrangement for both *in vivo* and video model prompts were identical, with the exception of the modality through with the prompt was provided. A second possible explanation for the lack of differentiated problem behavior could also be a result of environmental factors outside of the experimenters control. For example, for Lyla's non-compliance data set for set B, there appears to be some separation in data, however, a closer analysis reveals that the two data paths for *in vivo* and video model prompt sessions increase and decrease together, indicating that the changes in non-compliance may have been affected by extraneous environmental variables. Given these data, definitive conclusions regarding the effects of the independent variables on problematic behavior cannot be drawn.

Limitations and Future Research

Potential limitations of the present study should be considered. First, the participant's history with *in vivo* prompts during mand training may have had a significant role in the effectiveness of *in vivo* prompts, particularly during set A across participants. Had the participants been completely naïve to both prompting procedures for verbal behavior, it is possible that the rate of acquisition following the *in vivo* prompt may have more closely mimicked that of the video model prompt. This is further supported by the increase in correct prompted and independent responding following video model prompts in subsequent sets. These results may suggest that as participants contacted the contingency associated with the video model prompt, and thus established a history of reinforcement for echoing the video model, they then learned more quickly following this prompt. Future research could replicate the present study using learners who are completely naïve to *in vivo* prompts for the acquisition of verbal behavior. This would allow for more definitive data about the efficacy of the two prompting procedures, allowing for an unbiased comparison of the two prompts.

A second limitation of the present study is the differential ITI time between the two prompt types. The absence of technology for *in vivo* prompting procedures led to the experimenter's ability to quickly present trials, with ITIs of approximately three to five seconds, whereas the need to switch videos prior to presenting stimuli during video model prompt sessions increased the ITI during these sessions to approximately six to twelve seconds. Previous research has compared the effects of short and progressive ITIs to long ITIs on skill acquisition and maintenance of skill (e.g., intraverbal word associations) in children with ASD. This research shows that shorter ITIs typically led to quicker acquisition of target stimuli (Cariveau, Kodak, & Campbell, 2016; Koegel, Dunlap, & Dyer, 1980). The results of the present study may further support this research as mastery-level responding for all participants during set A, as well as sets B and C for Otis, and set C for Ross were met more quickly under the *in vivo* prompt condition in which ITI times were shorter. Future research could equate the ITI times for each prompt type to eliminate differential ITI times as a confounding variable.

A third limitation of the present study was limited maintenance of acquired targets during two and four-week post assessments. Despite high performance on initial post probes for each participant, maintenance of stimuli as demonstrated on two and four-week post assessments was lower than post probes across the first two participants. Future research could be conducted to evaluate how maintenance would be affected if the termination criterion were removed (i.e., if sessions continued until the participant reached mastery-level responding under both prompt types). Given the documented success of video-based prompts in promoting maintenance of acquired targets (MacDonald, Sacramone, Mansfield, Wiltz, Ahearn, 2009; Gena, Couloura, Kymissis, 2005; & Charlop-Christy & Daneshvar, 2003) it is plausible that if taught to mastery-level responding, video-based prompts may lead to maintenance of acquired tacts. Furthermore, future research could compare the effects of video model prompts to *in vivo* prompting procedures on the generalization of learned targets to the natural environment. Video modeling has been shown to lead to generalization (Charlop-Christy, Le, & Freeman 2000). The addition of generalization probes could assess whether video model prompts lead to participants emitting tacts outside of the training setting. Should video model prompts promote better generalization than *in vivo* prompts, researchers and clinicians would have to weigh the increased instructional time of video models against the potential long-term benefits of a generalized tact repertoire.

Overall, the results of the current study provide information about the efficacy of an additional prompting procedure for tact acquisition for young children with a diagnosis of ASD. Further research is needed to determine the extent to which prior histories with *in vivo* prompts for mand repertoires may influence rate of acquisition, though the present study presents data that demonstrates that both *in vivo* and video model prompts are effective in teaching tact repertoires to children with ASD. Given these results, clinicians should consider the use of video model prompts to teach tact repertoires.

APPENDIX

Table 1.

Otis Target Stimuli by Set

Intervention	Set A	Set B	Set C
In Vivo Prompts	Bed	Top	Boat
	Goat	Net	Cake
	Hat	Kite	Gear
Video Model Prompts	Pot	Map	Dime
	Bee	Hen	Tie
	Bat	Fan	House

Table 2.

Ross Target Stimuli by Set

Intervention	Set A	Set B	Set C
In Vivo Prompts	Pot	Can	Deer
	House	Gear	Phone
	Duck	Lime	Gum
Video Model Prompts	Map	Goat	Tape
	Fan	Bat	Top
	Dime	Kite	Net

Table 3.

Lyla Target Stimuli by Set

Intervention	Set A	Set B	Set C
In Vivo Prompts	Shampoo Doorbell Wallet	Napkin Pretzel Dresser	Cabbage Honey Chicken
Video Model Prompts	Band-Aid Towel Feather	Candle Bucket Ketchup	Taco Waffle Peanut

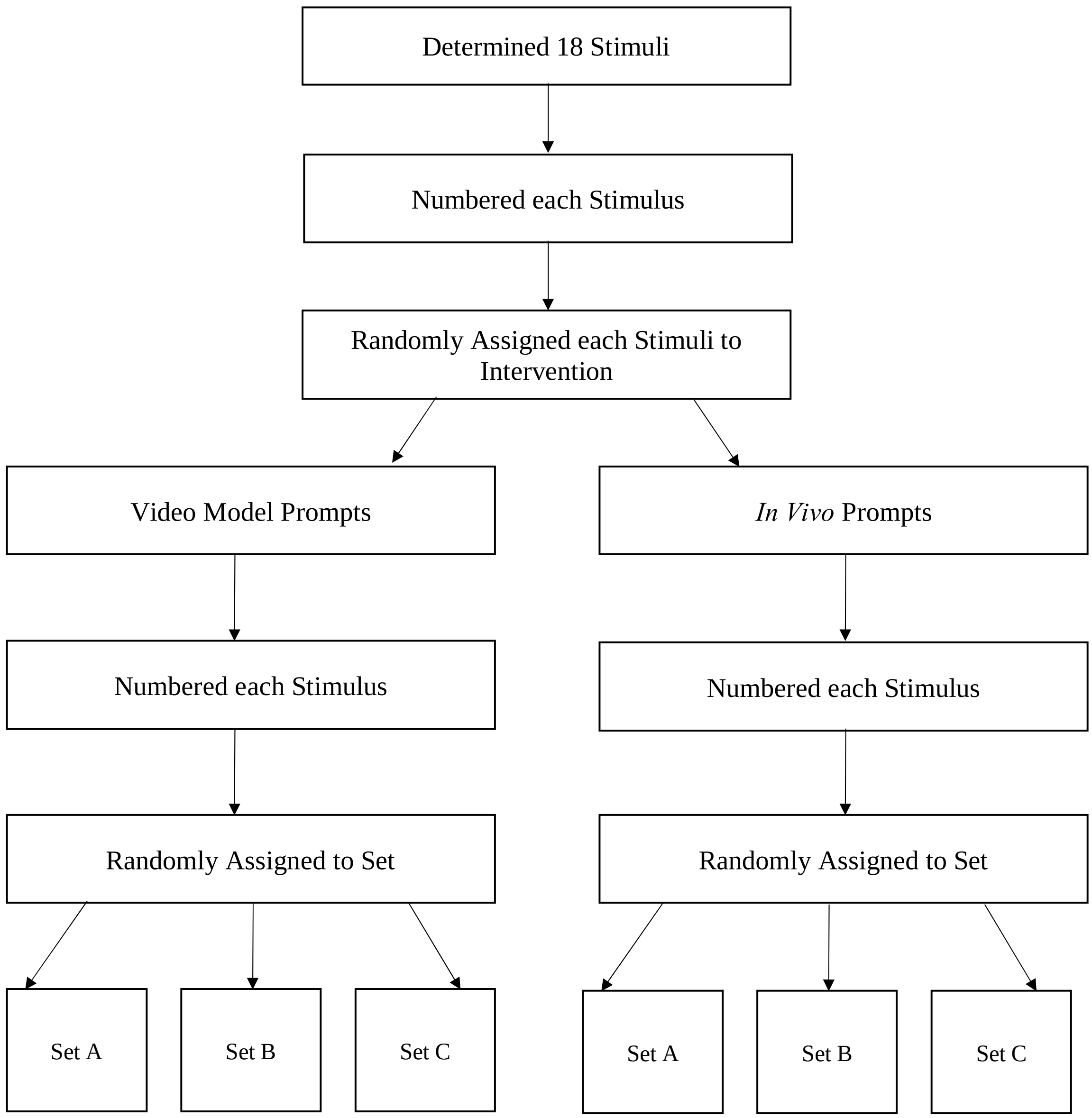

Figure 1. **Target Stimuli Selection Flow Chart**.

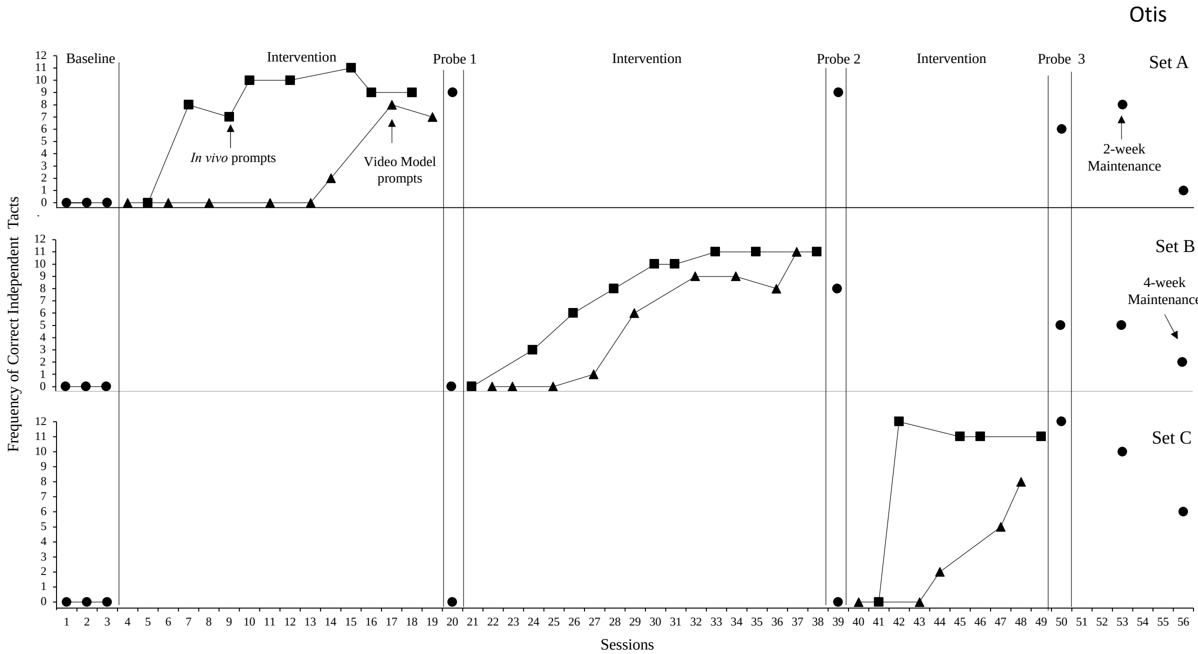

Figure 2. **Otis tact acquisition**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

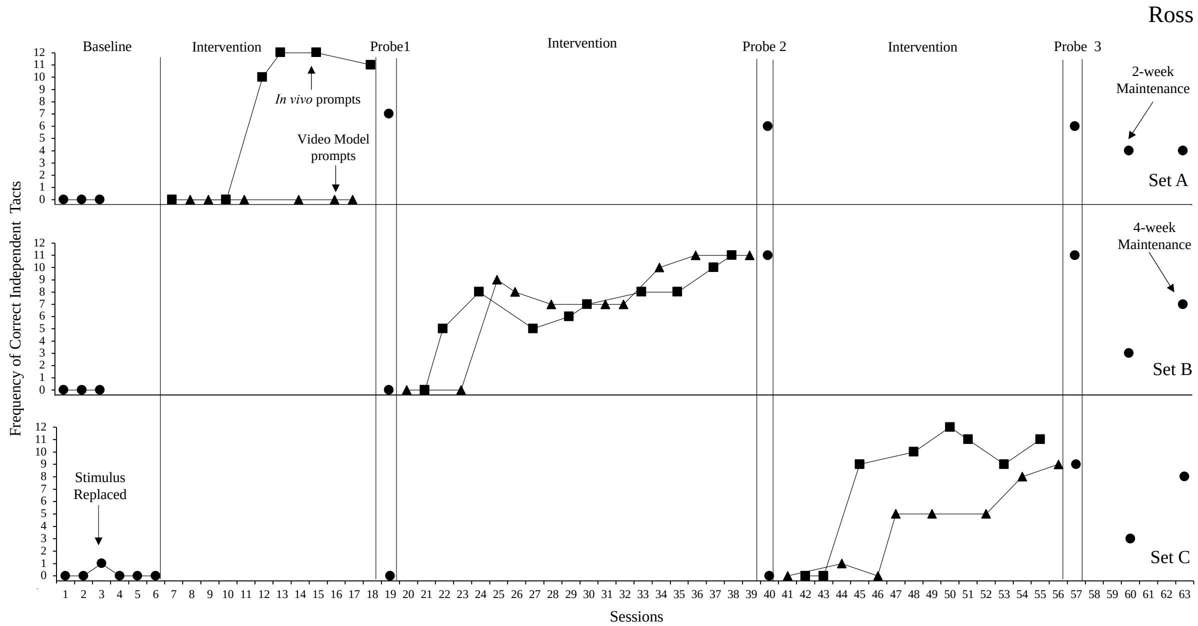

Figure 3. **Ross tact acquisition**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

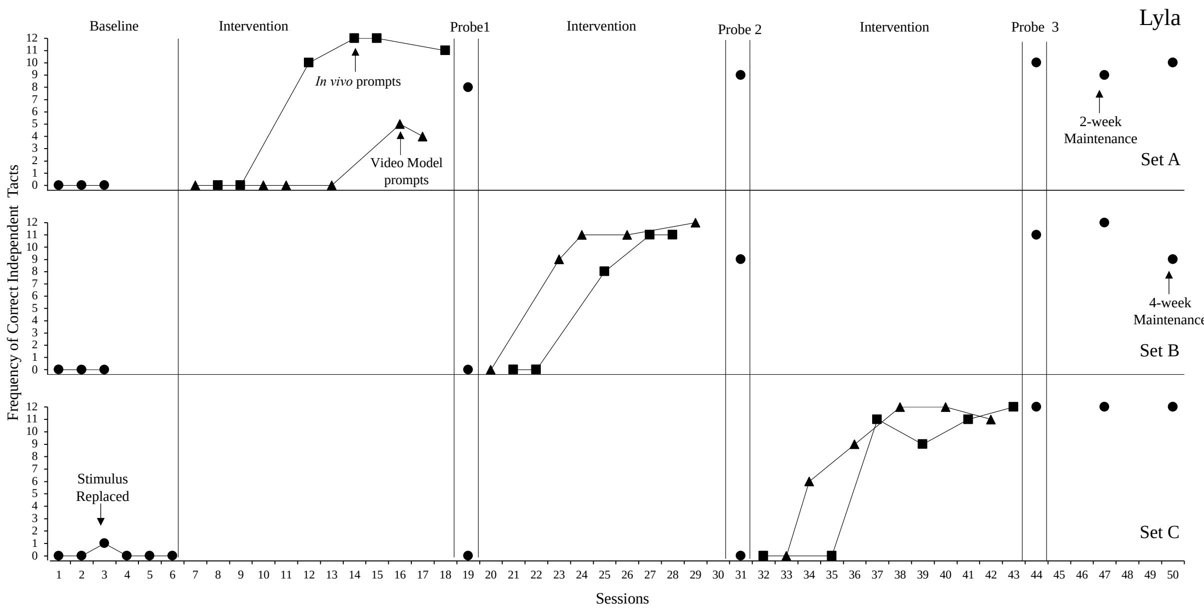

Figure 4. **Lyla tact acquisition**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

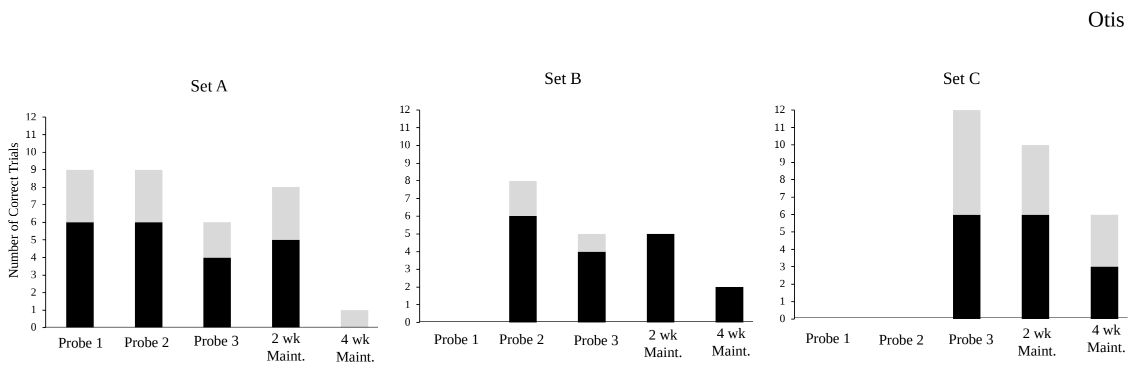

Figure 5. **Otis results by probe**. The black bars represent trials of stimuli taught under *in vivo* prompts, the grey bars represent trials of stimuli taught under video model

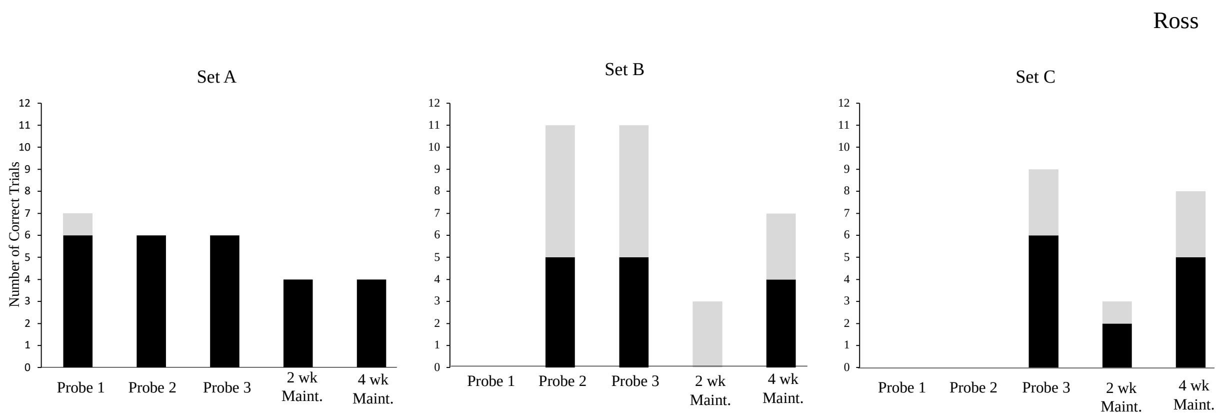

Figure 6. **Ross results by probe**. The black bars represent trials of stimuli taught under *in vivo* prompts, the grey bars represent trials of stimuli taught under video model

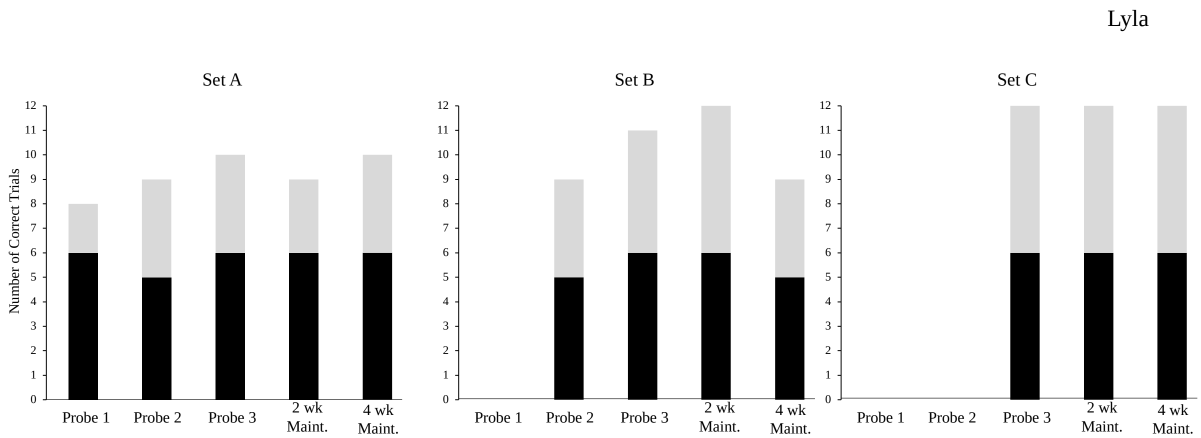

Figure 7. **Lyla results by probe**. The black bars represent trials of stimuli taught under *in vivo* prompts, the grey bars represent trials of stimuli taught under video model

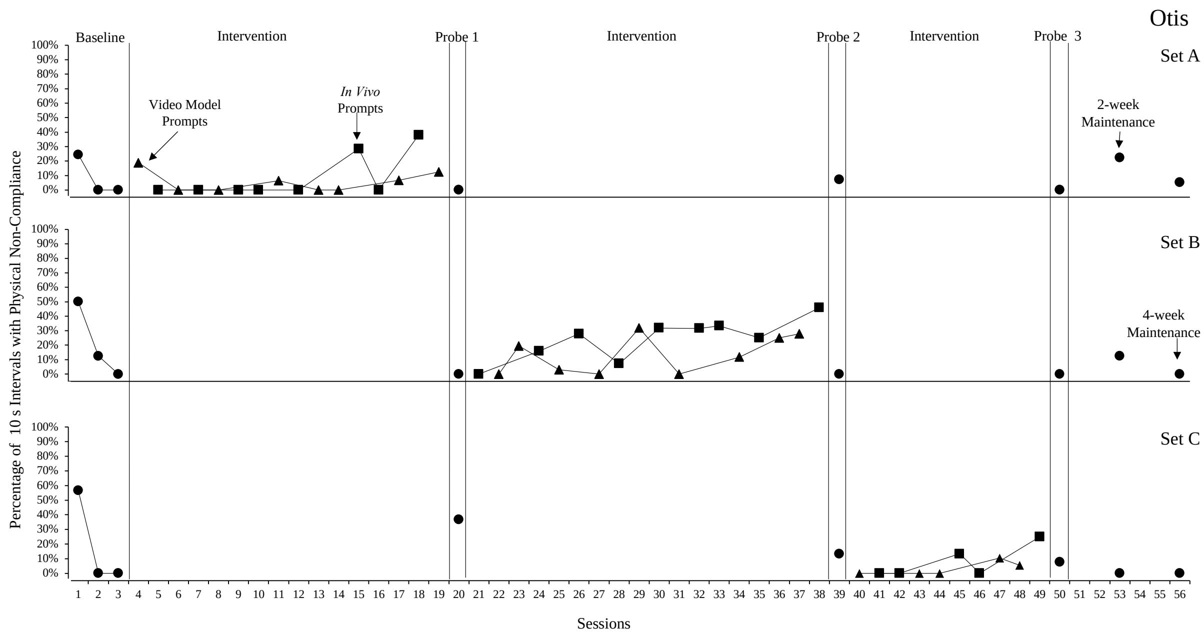

Figure 8. **Otis physical non-compliance**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

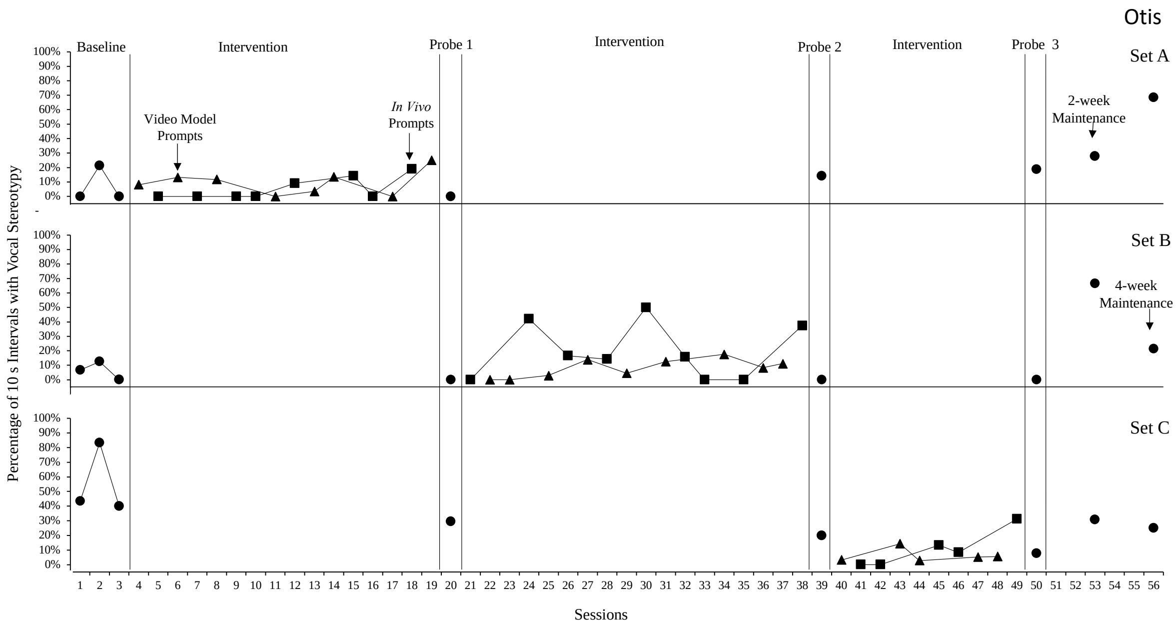

Figure 9. **Otis vocal stereotypy**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

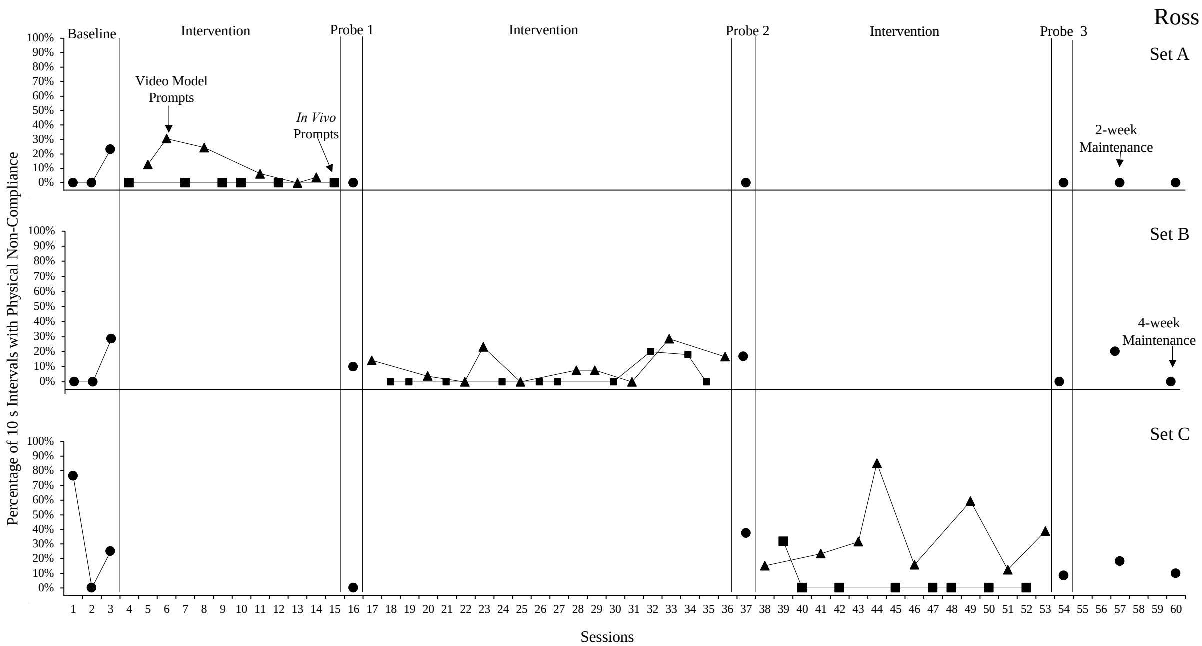

Figure 10. **Ross physical non-compliance**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

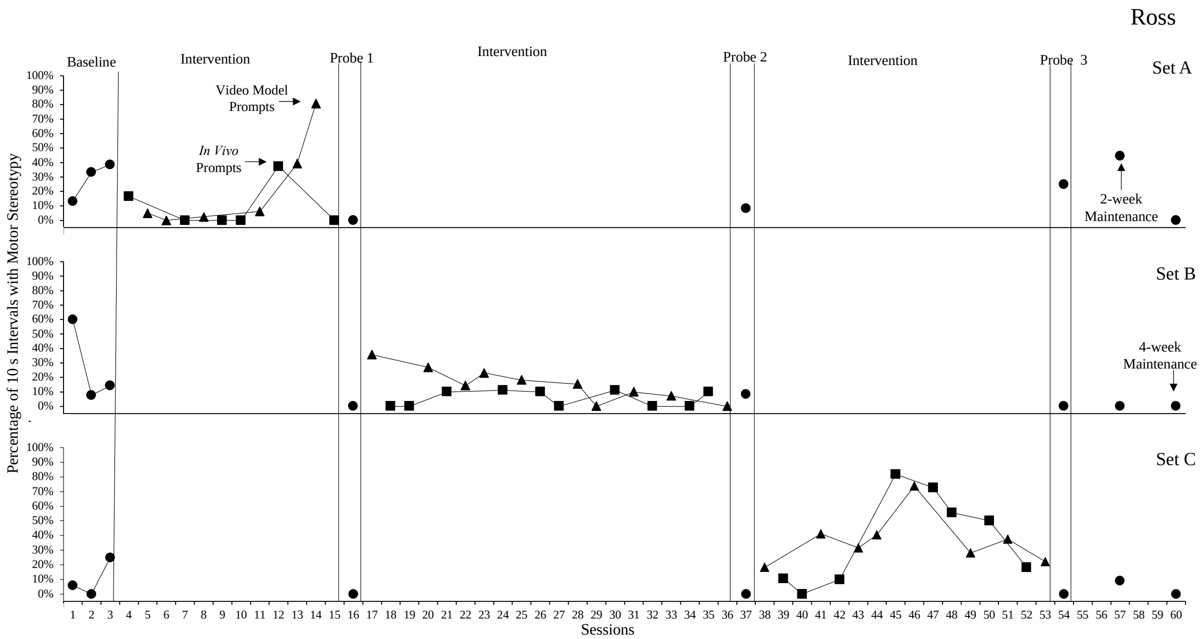

Figure 11. **Ross motor stereotypy**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

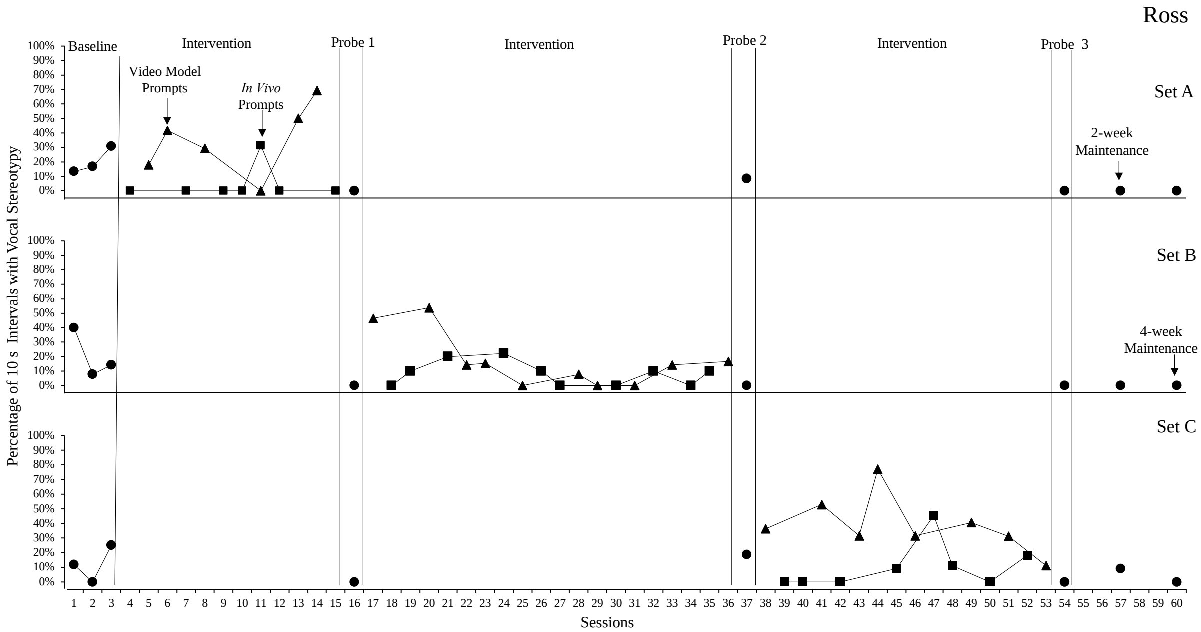

Figure 12. **Ross vocal stereotypy**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

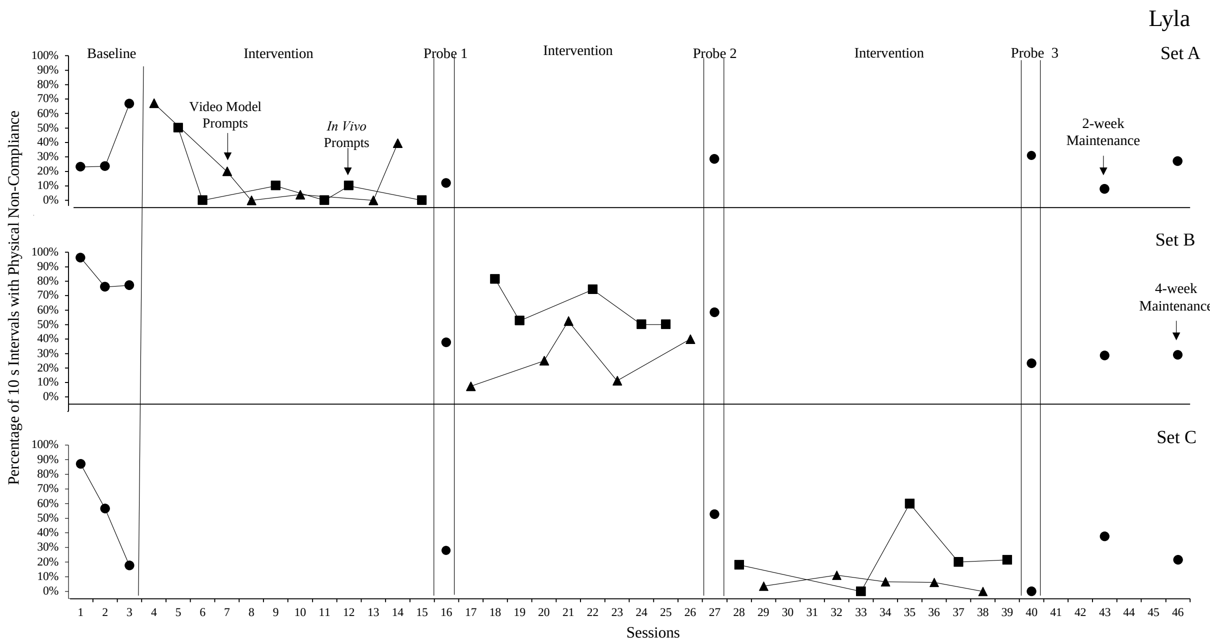

Figure 13. **Lyla physical non-compliance**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

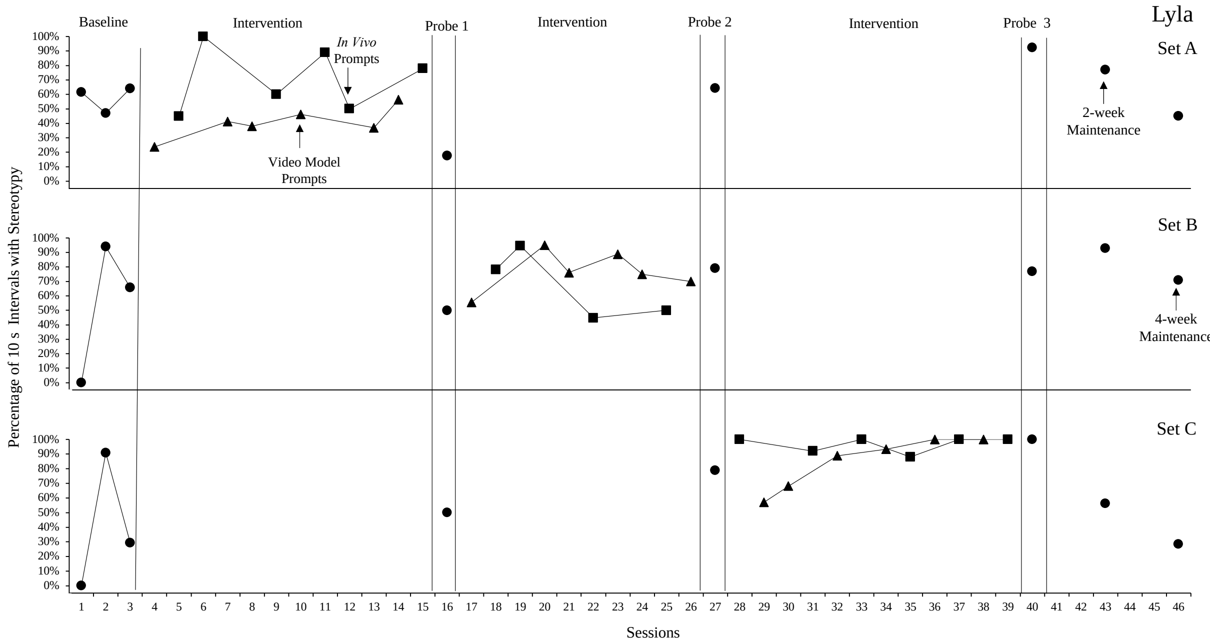

Figure 14. **Lyla stereotypy**. Square markers represent *in vivo* prompts and triangles represent video model prompts.

REFERENCES

REFERENCES

Ayres, K., & Ledford, J. R. (2014). Dependent measures and measurement systems. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed., pp. 124-153). New York, NY: Routledge.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling Interventions for children and adolescents with autism spectrum disorders. *Exceptional Children*, *73*, 264-287. retrieved from https://search-proquest-com.proxy2.cl.msu.edu/docview/201148251?accountid=12598

Cariveau, T., Kodak, T., & Campbell, V. (2016). The effects of intertrial interval and instructional format on skill acquisition and maintenance for children with autism spectrum disorders. *Journal of Applied Behavior Analysis, 49,* 809-825. doi: 10.1002/jaba.322

Charlop-Christy, M. H., Daneshvar, S. (2003). Using video modeling to teach perspective taking
to children with autism. *Journal of Positive Behavior Interventions, 5*, 12-21. doi:10.1177/10983007030050010101

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

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Measuring behavior. In J.O Cooper, T. E. Heron, & W. L. Heward, *Applied Behavior Analysis* (2nd ed. pp.72-101). Columbus, OH: Prentice Hall.

Esch, B. (2008). Early echoic skills assessment. Concord, CA: AVB Press.

Conallen, K., & Reed, P. (2016). A teaching procedure to help children with autistim spectrum disorder to label emotions. *Research in Autism Spectrum Disorders, 23*, 63-72. doi: 10.1016/j.rasd.2015.11.006

Frost, L. A., & Bondy, A. S. (2002). *The picture exchange communication system training manual* (2nd ed.). Newark, DE: Pyramid Educational Consultants.

Gena, A., Couloura, S., & Kymissis, E. (2005). Modifying the affective behavior or preschoolers with autism using in vivo or video modeling and reinforcement contingencies. *Journal of Autism and Developmental Disorders, 35,* 545-556. doi: 10.1007/s10803-005-0014-9

Grow, L., & LeBlanc, L. (2013). Teaching receptive language skills: Recommendations for instructors. *Behavior Analysis in Practice, 6*, 56-75. doi: 10.1007/BF03391791

Higbee, T. S., Carr, J. E., & Harrison, C. D. (2000). Further evaluation of the multiple stimulus preference assessment. *Research in Developmental Disabilities, 21*, 61–73.

Koegel, R. L., Dunlap, G., & Dyer, K. (1980). Intertrial interval duration and learning in autistic children. *Journal of Applied Behavior Analysis, 13,* 91–99. doi: 10.1901/jaba.1980.13-91

LaLonde, K., Duenas, A. D., Neil, N., Wawrzonek, A., & Plavnick, J. B. (2017). An evaluation of two tact training procedures on acquired tacts and tacting during play. *Manuscript submitted for publication.*

LeBlanc, L. A., Coates, A. M., Daneshvar, S., Charlop-Christy, M. H., Morris, C. & Lancaster, B. M. (2003). Using video modeling and reinforcement to teach perspective-taking skills to children with autism. *Journal of Applied Behavior Analysis*, *36*, 253-57. doi: 10.1901/jaba.2003.36-253

Ledford, J. R., & Gast, D. L. (2014). Combination and other designs. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed., pp. 346-376). New York, NY: Routledge.

MacDonald, R., Sacramone, S., Mansfield, R., Wiltz, K., & Ahearn, W. H. (2009). Using video modeling to teach reciprocal pretend play to children with autism. *Journal of Applied Behavior Analysis, 42*, 43-55. doi: 10.1901/jaba.2009.42-43

Marchese, N. V., Carr, J. E., LeBlanc, L. A., Rosati, T. C., & Conroy, S. A. (2012). The effects of the question “what is this?’ on tact-training outcomes of children with autism. *Journal of Applied Behavior Analysis, 45*, 539-47. doi: 10.1901/jaba.2012.45-539

McClannahan, L. E., & Krantz, P. J. (2005). *Teaching conversation to children with autism: Scripts and script fading*. Bethesda, MD: Woodbine House.

Michael, J., Palmer, D. C., & Sundberg, M. L. (2011). The multiple control of verbal behavior. *Analysis of Verbal Behavior, 27*, 3-22. Retrieved from https://search-proquest-com.proxy1.cl.msu.edu/docview/896406233?accountid=12598

Partington, J. W., Sundberg, M. L., Newhouse, L., & Spengler, S. M. (1994). Overcoming an autistic child’s failure to acquire a tact repertoire. *Journal of Applied Behavior Analysis*, *27*, 733–734. doi: 10.1901/jaba.1994.27-733

Pistoljevic, N., & Greer, R. D. (2006). The effects of daily intensive tact instruction on preschool students’ emission of pure tacts and mands in non-instructional settings. *Journal of Early and Intensive Behavioral Interventions*, *3*, 103-120. doi: 10.1037/h0100325

Plavnick, J. B., & Vitale, F. A. (2016). A comparison of vocal mand training strategies for children with autism spectrum disorders. *Journal of Positive Behavior Interventions*, *18*, 52-62. doi: 10.1177/1098300714548800

Ploog, B. O. (2010). Stimulus overselectivity four decades later: A review of the literature and its implications for current research in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 40*, 1332-1349. doi: 10.1007/s10803-010-0990-2

Reichow, B. & Wolery, M. (2009). Comparison of everyday and every-fourth-day probe session with the simultaneous prompting procedure. *Topics in Early Childhood Special Education, 29,* 79-89. doi: 10.1177/0271121409337885

Roxburgh, C. A., & Carbone, V. J. (2013). The effect of varying teacher presentation rates on responding during discrete trial training for two children with autism. *Behavior Modification, 37*, 298-323. doi: 10.1177/0145445512463046

Skinner, B. F. (1957). Verbal behavior. New York, NY: Appleton-Century-Crofts.

Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities, 16*, 86-92. doi: 10.1177/108835760101600204

Sundberg, M. L. (2008). Verbal behavior milestones assessment and placement program. Concord, CA: AVB Press.

Sundberg, M. L., & Partington, J. W. (1998). Teaching language to children with autism or other developmental disabilities. Pleasant Hill, CA: Behavior Analysts, Inc.

Sundberg, M. L., & Sundberg, C. A. (2011). Intraverbal behavior and verbal conditional discriminations in typically developing children and children with autism. *The Analysis of Verbal Behavior, 27*, 23-43. Retrieved from https://search-proquest-com.proxy1.cl.msu .edu/docview/896406236?accountid=12598

Walker, G. (2008). Constant and progressive time delay procedures for teaching children with autism: A literature review. *Journal of Autism and Developmental Disorders*, *38*, 261–275. doi:10.1007/s10803-007-0390-4

Williams, G., Carnerero, J. J., & Pérez-González, L. A. (2006). Generalization of tacting actions in children with autism. *Journal of Applied Behavior Analysis, 39*, 233-237. doi: 10.1901/jaba.2006.175-04

Wolery, M., Gast, D. L., & Ledford, J. R. (2014). Comparison designs. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed., pp. 297-345). New York, NY: Routledge.