Spring 5-7-2016

Effects of Modeling Varied Responses and Programming Lag Contingencies on Varied Responding during Discrete-Trial Instruction

Sean P. Peterson
University of Nebraska Medical Center

Effects of Modeling Varied Responses and Programming Lag Contingencies on Varied Responding during Discrete-Trial Instruction

by

Sean P. Peterson

A DISSERTATION

Presented to the Faculty of

the University of Nebraska Medical Center

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

Medical Sciences Interdepartmental Area Graduate Program

(Applied Behavior Analysis)

Under the Supervision of Professor Nicole M. Rodriguez

University of Nebraska Medical Center
Omaha, Nebraska

April, 2016

Supervisory Committee:

Nicole M. Rodriguez, Ph.D.	Wayne W. Fisher, Ph.D.
Kevin C. Luczynski, Ph.D.	Brian D. Greer, Ph.D.

EFFECTS OF MODELING VARIED RESPONSES AND PROGRAMMING LAG CONTINGENCIES ON VARIED RESPONDING DURING DISCRETE-TRIAL INSTRUCTION

Sean P. Peterson, Ph.D.

University of Nebraska, 2016

Supervisor: Nicole M. Rodriguez, Ph.D.

Children with autism often require direct instruction to learn skills (e.g., discrete-trial instruction [DTI]). Despite its advantages, DTI has been criticized for producing rote responding (e.g., Cihon, 2007). Although there is little research supporting this claim, if true, this may be problematic given the propensity of children with autism to engage in restricted and repetitive behavior (American Psychiatric Association, 2013). In Experiment 1, we evaluated the effects of modeling rote versus varied target responses during DTI on producing varied responding and efficiency of skill acquisition in learning intraverbal categorizations. For all four children, all increases in varied responding were temporary, and, for two participants, acquisition was slowed in the variable-modeling condition. In Experiment 2, we evaluated the effects of introducing a Lag-1 schedule to the variable-modeling condition, and varied responding increased for all participants, but only when combined with modeling of varied responding with a progressive-prompt delay.

Keywords: autism, discrete-trial instruction, lag schedule, prompts, variability, rote responding

Table of Contents

Effects of Modeling Varied Responding and Programming Lag Contingencies on Varied Responding during Discrete-Trial Instruction

Whereas children of typical development learn directly from natural interactions with their environment, children diagnosed with autism spectrum disorder (ASD) often require direct and systematic instruction with repeated practice to acquire new skills. Repeated practice opportunities in close temporal proximity is a hallmark feature of discrete-trial instruction (DTI), a method commonly used to teach children with ASD. DTI includes at least four basic components for each trial: (a) the therapist provides a discriminative stimulus (i.e., a stimulus that leads to reinforcement for certain responses in its presence but not in its absence; e.g., an instruction or question), (b) the therapist provides a controlling prompt (e.g., vocal, model, or physical prompt), (c) the learner responds, and (d) the therapist provides a consequence for that response (Smith, 2001). A controlling prompt is defined by the reliability with which it occasions the correct response and is often delivered immediately following the discriminative stimulus during the early stages of DTI. Stimulus control (i.e., the conditions under which a behavior of interest is influenced by the presence of an antecedent stimulus) can then be transferred from the prompt to the discriminative stimulus by, for example, using a progressive-prompt delay in which the delay between the discriminative stimulus and the prompt is systematically increased (e.g., 0 s, 2 s, 5 s; Billingsley & Romer, 1983). For example, the therapist may provide the discriminative stimulus, "How are you?" and then immediately prompt, "Say 'I'm fine.'" If the prompt serves as a controlling prompt, the learner will immediately repeat "I'm fine" every time the prompt is delivered, providing the therapist with an opportunity to reinforce this response. The delay between the discriminative stimulus and the controlling prompt is systematically increased until the learner responds "I'm fine" following the discriminative stimulus. In this way, the stimulus control exerted by the controlling prompt would be transferred to the question "How are you?"

Despite the success of using progressive-prompt delays in DTI to teach a variety of skills such as matching (e.g., Cummings & Williams, 2000), receptive identification (e.g., Gutierrez et al., 2009), expressive identification (Leaf et al., 2013), and responding to social questions (e.g., Wong & Woolsey, 1989), this teaching method has sometimes been criticized for producing rote responding (Cihon, 2007). This criticism may stem from the fact that DTI trials are often arranged such that one correct response is repeatedly prompted and reinforced for each discriminative stimulus. For example, a therapist may prompt and reinforce "I'm fine" following the discriminative stimulus "How are you?" despite the fact that other responses such as "Good, thanks" would also be socially acceptable. Thus, it is possible that the prompting and reinforcing of one correct response narrows the range of responses that (a) are emitted and (b) contact reinforcement, potentially contributing to the development of rote responding.

Given that restricted and repetitive behavior is a diagnostic characteristic of ASD (American Psychiatric Association, 2013), it is important to identify any potential negative effects of commonly used teaching procedures on varied responding as well as strategies for addressing invariant responding in children with ASD. This purpose served as the impetus for the experimental questions addressed in Experiments 1and 2.

Experiment 1

One approach to counter the potential negative effects of DTI procedures may be to prompt and reinforce multiple correct responses during teaching, which is consistent with the response generalization strategies of teaching sufficient exemplars and training loosely (e.g., broadening the contingency class) described by Stokes and Baer (1977). However, simply broadening the contingency class by increasing the range of responses eligible for reinforcement may not be sufficient to promote varied responding. In fact, research has shown that contingencies that allow but do not require varied responding produce invariant (rote) responding over repeated exposure to the contingency. For example, Antonitis (1951) assessed varied responding of the location of rats' nose-thrusting along a 50-cm opening. When reinforcement was delivered for nose-thrusts anywhere along the opening, varied responding decreased within and across sessions.

Neuringer, Kornell, and Olufs (2001, Experiment 2) also showed the importance of reinforcement contingencies on varied responding with rats that were required to complete a three-response sequence across three operanda (two levers and a key) to contact reinforcement. For rats assigned to the variability group (Var), a response sequence was reinforced if it had been emitted on less than or equal to 5% of trials, based on an equation that took into account the recency of the last instance of the response sequence. For rats assigned to the yoked group (Yoke), a response was reinforced following trials in which the rat it was paired with from the Var group contacted reinforcement (e.g., on trials 3, 9, & 13). This procedure yoked the frequency and distribution of reinforcement; further, the contingency allowed, but did not require, varied responding for the rat's behavior to contact reinforcement. Varied responding was significantly higher for the rats in the Var group than for the Yoke group, suggesting that the reinforcement contingency and not extinction-induced variability was responsible for increases in varied responding. In Experiment 3 of the same study, rats were assigned to one of two groups: Var or Repetition (Rep). For rats assigned to the Var group, reinforcement was the same as described above; by contrast, for rats assigned to the Rep group, reinforcement was contingent on a single response sequence. An interesting finding was that a comparison of varied responding across the Yoke group in Experiment 2 and the Rep group in Experiment 3 (Figures 4 and 7 in Neuringer et al., 2001) revealed slightly higher levels of varied responding in the Rep group. This suggests that contingencies that simply allow but do not require varied responding (as arranged in the Yoke group) are likely to result in levels of invariant responding that are comparable to the level of invariant responding observed when reinforcement is made contingent on one correct response.

Taken together, results from basic research on varied responding suggest that varied responding will only increase when contingencies are arranged to reinforce varied responding. However, results from basic research have also shown that varied responding might be observed under contingencies that allow but do not require varied responding to contact reinforcement such as when such contingencies are preceded by a history of reinforcement for varied responding (e.g., Saldana & Neuringer, 1998). Thus, it is possible that supplemental strategies such as modeling varied responding may be sufficient to bring multiple responses in contact with reinforcement when reinforcement is provided for any correct response. To this point, several researchers have suggested the importance of including varied models to increase varied responding during social interactions (Weiss, LaRue, & Newcomer, 2009) and play (Jahr & Eldevik, 2002). The ability to produce varied responding under contingencies that allow but do not require variability has practical implications if we presume that such contingencies are more commonly encountered in the natural environment than contingencies that require varied responding. For example, Rodriguez and Thompson (2015) suggested that it may be considered socially inappropriate to provide corrective feedback for invariant responding (e.g., telling a child that repeating the same joke to the same person is not funny) such that the people interacting with the child may not provide differential consequences for invariant responding. In addition, any social consequences that are delivered following invariant responding may be too subtle or intermittent to function as reinforcers or punishers (e.g., rolling one’s eyes after hearing a repeated joke; Rodriguez & Thompson, 2015).

To our knowledge, no studies have compared how modeling invariability and variability during prompts within DTI affects responding when reinforcement is provided for any correct response, regardless of the level of varied responding. As previously noted, DTI typically involves the repeated prompting of one response and reinforcement for that one response; as such, to determine the potential negative effects of these procedures on varied responding, we evaluated the effects of rote-models versus variable-models of correct responses during a progressive-prompt delay. Because prompting multiple correct responses may interfere with acquisition, we also evaluated the efficiency of each modeling procedure.

Method

To answer our experimental questions, we adopted a translational approach in which the skill we taught was relevant to our participants' early intervention goals (i.e., intraverbal categorization), but the response targeted for variability was more aligned with basic research (i.e., the sequence in which the target exemplars were provided). In this way, we could explore the effects of different types of prompts in a simple and controlled experimental arrangement that allowed us to equate the number of exemplars (three) that participants were exposed to across conditions. Participants were taught to respond to intraverbal categorization questions using a progressive-prompt delay during either a rote-modeling or a variable-modeling condition.

Participants, setting, and materials. Participants included four children, 3 to 8 years old diagnosed with autistic disorder, who were receiving early intervention services at a mid-western medical school. At the start of the experiment, Victor was a 5-year-old boy, Yogi was an 8-year-old boy, Adam was a 5-year-old boy, and Ronda was a 3-year-old girl. Sessions were conducted in either a 2.5-m by 2.5-m private session room in which the participant received early intervention services (Victor, Adam, and Ronda) or a living area of the participant's home (Yogi). The participant sat next to or across from the experimenter at a child-sized table. Each session consisted of 10 trials and was approximately 5 to 10 min. Two to four sessions were conducted per day, up to 5 days per week. Materials included data sheets and reinforcers.

Data collection, dependent variables, and interobserver agreement. During each trial, paper-and-pen data were collected on the participant's vocal-response sequence verbatim (i.e., each word emitted by the participant in the order in which they were stated were recorded) and whether the vocal response occurred prior to or following the controlling prompt (i.e., independent or prompted responding). To evaluate differences in varied responding as well as the rate of acquisition, data were summarized to gather information on the following primary dependent variables for each session: (a) the number of unique response sequences in a session, (b) the percentage of trials in which the response sequence differed from the previous trial, and (c) the percentage of trials with correct responding. A correct response was defined as the participant delivering three appropriate exemplars for the category, independent of a particular response sequence (i.e., the order of exemplars provided; e.g., "dog, cat, horse"; "cat, dog, horse"; or "cat, dog, mouse" as exemplars of animals). The number of unique response sequences was the primary measure of varied responding for Experiment 1 and was calculated by determining the number of trials in which there was a correct response and the order of exemplars differed from the order of exemplars in response sequences in all preceding trials in that session. A response sequence containing a novel exemplar (i.e., an exemplar that was never prompted in the study; e.g., mouse in the previous example) was also scored as unique; this happened only once with Adam. The percentage of trials that differed from the previous trial was the primary measure of varied responding for Experiment 2 and was calculated by determining the number of trials in which a correct response was provided and the response sequence differed from the response sequence provided in the preceding trial; the quotient was then divided by nine (the total number of opportunities for a trial to differ from the previous trial for each session). Thus, if a participant alternated between two correct sequences (e.g., "dog, horse, cat"; "cat, horse, dog"; "dog, horse, cat"; "cat, horse, dog" and so on) in a session, the number of unique response sequences would be two and the percentage of trials that differed from the previous trial would be 100%.

A second, independent observer collected data for 64% to 100% of sessions across participants. Interobserver agreement was calculated for correct responding and response sequence on a trial-by-trial basis. An agreement for correct responding was defined as both observers scoring either correct or incorrect on a trial. An agreement for response sequence was defined as both observers scoring the same three exemplars in an identical order for both independent and prompted responses. Interobserver agreement was calculated by dividing the number of trials with an agreement by the total number of trials (10), and then converting the quotient to a percentage. Interobserver agreement for Victor was 97% (range, 40 to 100%), Yogi was 99% (range, 90 to 100%), Adam was 99% (range, 70 to 100%), and Ronda was 99% (range, 90 to 100%).

Preassessment. We conducted a preassessment to identify unknown target responses to include in the comparison. Three categories (e.g., fruits, animals, and tools) were each presented four times in a session (resulting in 12 trials per session), during which the experimenter provided the discriminative stimulus, “Tell me some [category].” If the participant engaged in an incorrect response or did not respond within 5 s of the discriminative stimulus, we did not provide feedback and began the next trial. Any category for which a participant provided two or more exemplars (e.g., “peach” and “pear” for the category *fruit*) was excluded, and an alternate category was assessed. This occurred only for Adam. This process continued until we identified four categories for which no more than one exemplar was provided for two consecutive sessions. All correct responses were reinforced with praise and a highly preferred edible or tangible (identified via a one-trial multiple stimulus without replacement preference assessment [MSWO] similar to DeLeon et al., 2001).

Experimental design. We used an adapted alternating treatments design (Sindelar, Rosenberg, & Wilson, 1985) within a multiple probe design across sets of categories. There were two categories per set with three exemplars per category. Categories in each set were matched based on the number of syllables for each exemplar and then randomly assigned to either the rote- or variable-modeling condition. A list of categories and exemplars for each condition and each set per participant are in Table 1.

Procedure. Prior to each session, we conducted a one-trial MSWO (similar to DeLeon et al., 2001). The item selected was used throughout the session unless the participant requested an alternative item, in which case the requested item was delivered for the remainder of the session or until a different item was requested.

Differential reinforcement baseline (DSR BL). Each session consisted of 10 trials. During each trial, the experimenter delivered the discriminative stimulus "Tell me some [category]" and provided the participant the opportunity to respond. The participant had 5 s to initiate responding and a response was considered complete if the participant ceased responding for 5 s. For example, if the participant gave two exemplars, they would have 5 s after providing the second exemplar to include a third and final exemplar or the response would be counted as an error. Contingent on a correct response (regardless of response sequence), praise and the reinforcer was delivered. Contingent on an incorrect response or no response within 5 s, the experimenter provided no feedback and proceeded to the next trial. The criterion for moving from baseline to the modeling phase was stable responding across both conditions using visual inspection.

Modeling. This phase was identical to baseline, except that (a) a progressive-prompt delay was used in both conditions and (b) each category was assigned to either the rote- or variable-modeling condition.

Progressive-prompt delay. The progressive-prompt delay (PPD) phase started with at least two sessions at a 0-s prompt delay, during which the experimenter immediately modeled a correct response following the discriminative stimulus during each trial. Following the prompt, the participant had 5 s to initiate responding and the trial ended when there was no response for 5 s or

Table 1

List of Categories and Exemplars for each Condition and Set per Participant

Participant	Set 1		Set 2		Set 3	
	Category	Exemplars	Category	Exemplars	Category	Exemplars
Victor	Little Animals	Frog Dog Mouse	Things in the Fridge	Milk Apple Cheese	Inside Sports	Bowling Wrestling Hockey
	[a]Big Animals	Bear Horse Whale	Things in the Bathroom	Sink Kleenex Soap	Outside Sports	Tennis Fishing Sailing
Yogi	Outside Sports	Tennis Baseball Fishing	[a]Tools in the Garden	Rake Lawn-mower Shovel		
	[a]Inside Sports	Bowling Hockey Wrestling	Tools in the Kitchen	Micro-wave Blender Spoon		
Adam	Rhymes with Snowflake	Earth-quake Headache Cupcake	Rhymes with Haystack	Backpack Kayak Racetrack		
	Rhymes with Batman	Toucan Suntan Dishpan	[a]Rhymes with Allow	Snowplow Eyebrow Meow		
Ronda	Big Animals	Bear Horse Whale	Foods	Eggs Sandwich Cheese		
	Little Animals	Frog Dog Mouse	[a]Drinks	Milk Juice Water		

Note. The top category for each participant and set was assigned to the rote-modeling condition in Experiment 1, and the bottom category was assigned to the variable-modeling condition.

[a]Categories and exemplars that were included in Experiment 2

three exemplars were provided. Following two consecutive sessions with correct prompted responding at or above 90%, the experimenter provided a 2-s delay following the discriminative stimulus before modeling the correct response. The criterion to increase the prompt delay was one session in which 50% or more of the errors consisted of no response (omission errors). The progression of the prompt delays was 2 s, 5 s, 10 s, and 20 s and each condition progressed independent of the other. Throughout this phase, the experimenter provided praise and the reinforcer for a correct independent or prompted response, a vocal model for an incorrect response and no response, and no consequence followed by the next trial for an incorrect response following the model. Because Victor's correct responding was low following the first session at the 2-s prompt delay in Set 1 and the majority of his errors were omission errors, we increased the prompt delay to 5 s. Following the increase, the majority of Victor's errors in the variable-modeling condition were errors of commission. For this reason, we changed the contingencies for Set 1 starting with session 24 (the 12^{th} session of the variable-modeling condition in Set 1) such that only praise was provided for correct prompted responding in this condition.

Experimental conditions. To evaluate the effects of modeling rote versus varied responses, we compared two conditions across which the sequence of prompted exemplars was manipulated. For the category in the rote-modeling condition, the experimenter modeled invariability by modeling the same response during each trial (e.g., the experimenter modeled "apple, orange, banana" on every trial).

For the category in the variable-modeling condition, the experimenter modeled variability by modeling three correct exemplars in a different sequence during each trial (e.g., "dog, cat, horse" for the first trial; "dog, horse, cat" for the second trial; and "cat, horse, dog" for the third trial). To ensure that participants were exposed to a range of possible sequences during the variable-modeling condition while also allowing for a test of generativity (i.e., the occurrence of a response sequence to which the participant had not been exposed to; Axe & Sainato, 2010), four of the six possible sequences were modeled throughout this phase. Prior to each session, the experimenter randomized a list of the four sequences. If a model was necessary, the experimenter modeled the next sequence on the list that differed from the participant's response in the previous trial (such that repeated responding was neither modeled nor reinforced). For instance, if the child responded "dog, cat, horse" on the first trial and engaged in an error or did not respond on the second trial, the experimenter modeled the next sequence on the list that differed from the response on the preceding trial (e.g., "horse, dog, cat"; note that if "dog, cat, horse" was next on the list, it would have been skipped to ensure that the model was different from the correct response provided). Each programmed sequence was checked off after it was modeled, and the experimenter continued with the remainder of the list. To isolate the effects of modeling varied responding, differential consequences were not provided for varied responding across rote- and variable-modeling conditions (i.e., reinforcement was provided for any correct sequence independent of variability).

Results and Discussion

The first purpose of Experiment 1 was to examine whether incorporating multiple models would lead to increased varied responding. Figure 1 shows the number of unique response sequences for Victor, Yogi, Adam, and Ronda. During the DSR BL, none of the participants provided a correct response to any of the intraverbal categories. Following the implementation of the progressive-prompt delay, an initial increase in the number of unique response sequences was observed for at least the variable-modeling condition of the first set for each participant but decreased to one response sequence following repeated exposure to the conditions. This temporary increase in the number of unique response sequences occurred only in the variable-modeling condition in Set 1 for Victor, Set 1 for Yogi, and Sets 1 and 2 for Ronda. Because varied responding was not

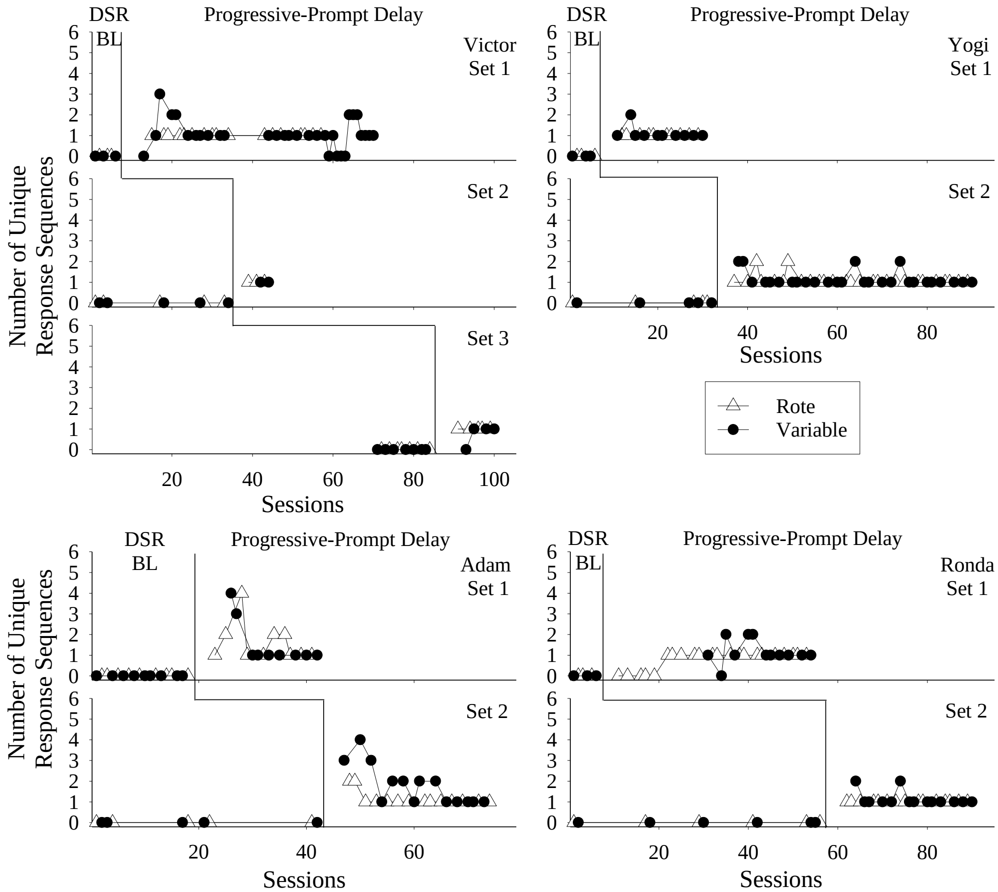

Figure 1. Varied responding as represented by the number of unique response sequences for all participants during Experiment 1. The data that are absent following the introduction of the progressive prompt delay phase represent the 0-s phase (i.e., prompted responding is not depicted in this figure). DSR BL was the differential reinforcement baseline condition.

observed with Set 2 for Victor, we evaluated a third set; however, like in Set 2, Victor emitted only one response sequence across both conditions. Thus, the temporary increase in varied responding was not replicated across sets for Victor. The lack of varied responding in Sets 2 and 3 for Victor may have been the result of sequence effects. That is, Victor's responding during the progressive-prompt delay phase for Sets 2 and 3 may have been affected by the fact that responding did not contact differential contingencies for varied responding in Set 1. It is unclear, however, why similar results were not obtained for other participants.

Whereas a temporary increase in the number of unique response sequences was observed only in the variable-modeling condition for at least one set for Victor, Yogi, and Ronda (Figure 1), we observed an initial increase in the number of unique response sequences across both the rote- and variable-modeling conditions for Set 2 for Yogi and both Sets 1 and 2 for Adam. Figure 2, is an alternative measure of varied responding. It depicts the percentage of trials that differ from the previous trial. Despite efforts to establish stimulus control over the different sets using an adapted alternating treatment design, the fact that (a) the conditions were similar in all regards except for the categories and exemplars targeted and (b) the conditions were conducted in rapid alternation may have led to indiscriminate responding. Said differently, Yogi's and Adam's temporary increases in the rote-modeling condition may have been due to carry over effects; it is possible that temporary increases in varied responding would not have been observed if the rote-modeling condition was not intermixed with the variable-modeling condition. There may have also been carry over effects from the rote-modeling condition to the variable-modeling condition. It is possible that varied responding may have persisted longer in the absence of direct reinforcement contingencies for varied responding had variable-modeling sessions not been intermixed with rote-modeling sessions within an alternating treatment design. That is, exposure to reinforcement for one

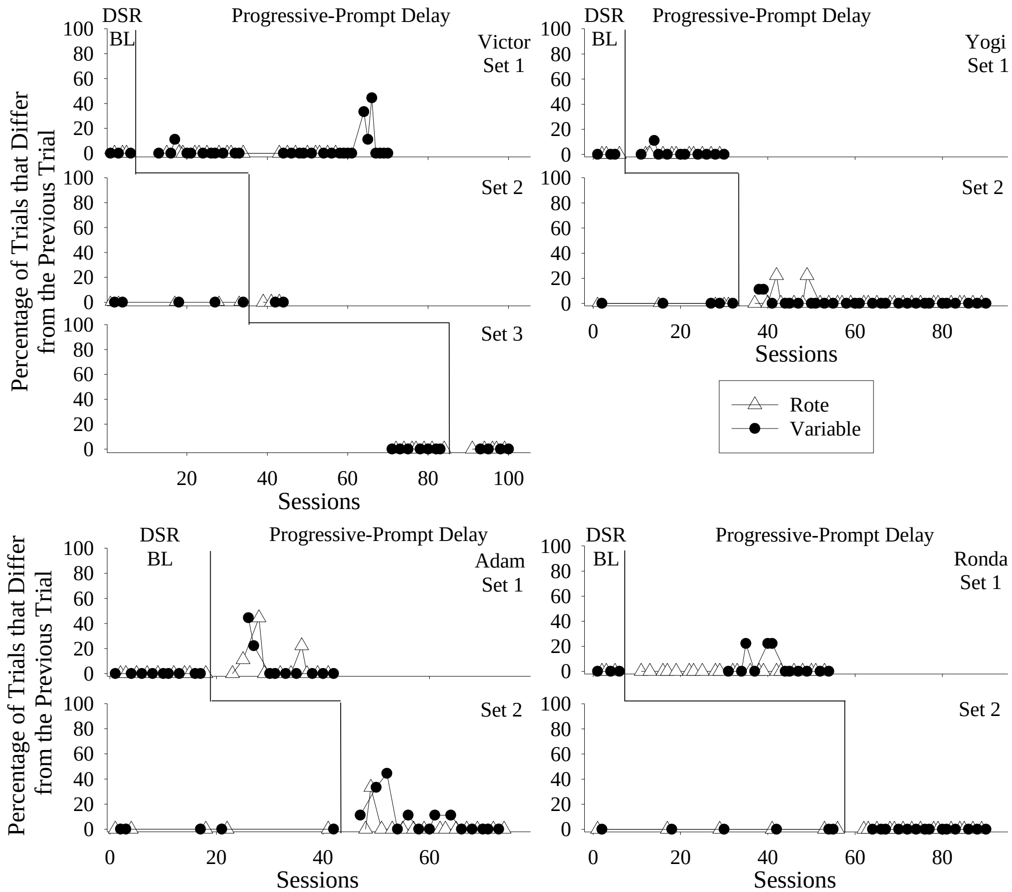

Figure 2. Varied responding as represented by the percentage of trials that differ from the previous trial for Victor, Yogi, Adam, and Ronda in Experiment 1. DSR BL denotes the differential reinforcement baseline condition. The data that are absent following the introduction of the progressive-prompt delay phase represent the 0-s phase (i.e., prompted responding is not depicted in this figure).

response sequence in the rote-modeling condition may have affected responding in the variable-modeling condition. Both of these possibilities (i.e., carryover of the effects of the rote-modeling onto the variable-modeling and carryover of the effects of the varied-modeling onto the rote-modeling) are worthy of consideration in future research. For example, future research might assess the effects of variable-modeling in isolation, when not alternated with a rote-modeling condition.

The second purpose of Experiment 1 was to examine whether incorporating multiple models within prompts would compromise efficiency during acquisition. Figure 3 shows the percentage of correct responses for Victor, Yogi, Adam, and Ronda. Because we selected intraverbal categories that the participants did not know, there were no correct responses during the DSR BL. Following the introduction of the progressive-prompt delay, one of two general patterns of responding were observed across participants: equal rates of acquisition across both phases or slower acquisition in the variable-modeling condition. For Victor, whereas the rate of acquisition was comparable across both conditions for Set 2, it took 7 sessions in the rote-modeling condition and 14 sessions in the variable-modeling condition to reach the mastery criterion in Set 1 and it took only 5 sessions in the rote-modeling condition, but 8 sessions in the variable-modeling condition to reach mastery for Set 3. For Yogi and Adam, acquisition rates were similar across conditions and sets, suggesting that modeling varied responses in the variable-modeling condition did not interfere with acquisition. For Ronda, acquisition was notably slower in the variable-modeling condition. It took 12 sessions in the rote-modeling condition and 21 sessions in the variable-modeling condition to reach the mastery criterion in Set 1, and it took 5 sessions in the rote-modeling condition and 12 sessions in the variable-modeling condition to reach mastery criterion in Set 2[1]. It is worth noting that, unlike other participants who met criteria for increasing the prompt delay from 0 s to 2 s within two to four sessions, numerous sessions (12) were

[1]All 0-s prompt delay sessions were included when calculating the number of sessions it took to reach the mastery criterion in Experiment 1. Sessions conducted at a 0-s prompt delay are not displayed in the figures to facilitate visual inspection of the comparison between rote-modeling and variable-modeling conditions.

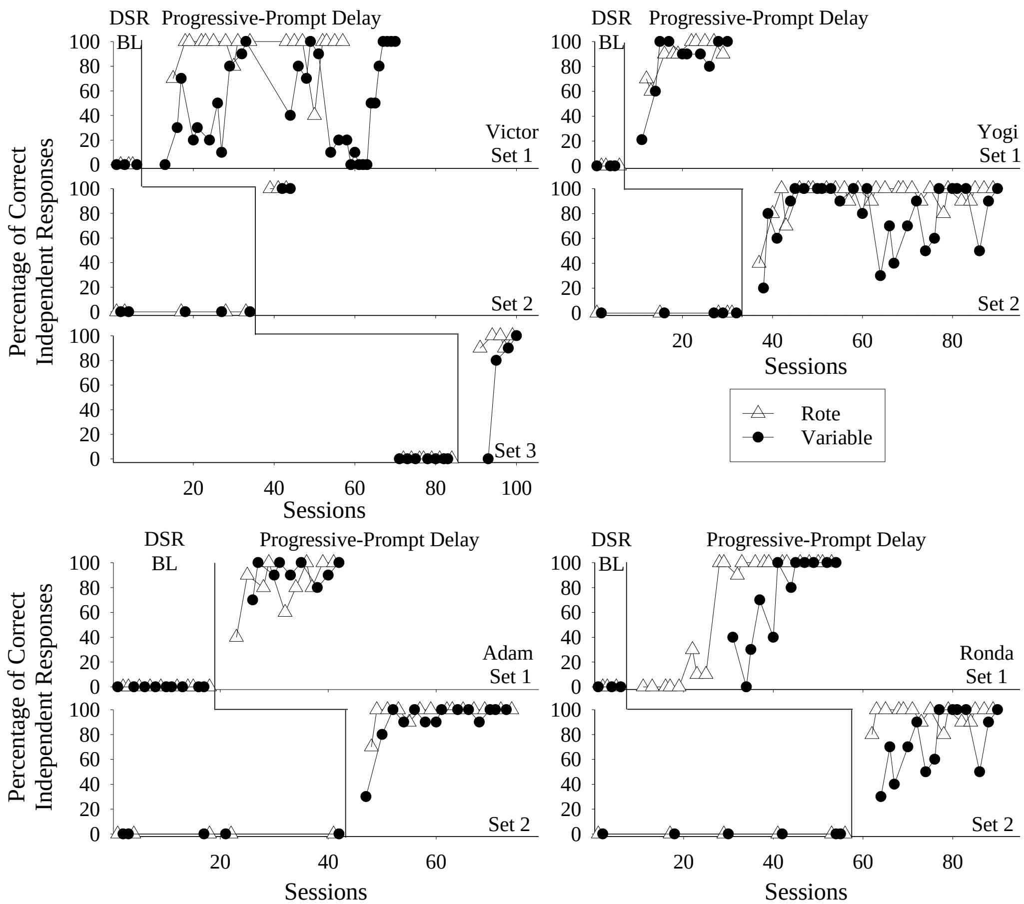

Figure 3. Acquisition data during Experiment 1. The data that are absent following the introduction of the progressive prompt delay phase represent the 0-s phase (i.e., prompted responding is not depicted in this figure). DSR BL was the differential reinforcement baseline condition.

conducted at a 0-s prompt delay in the variable-modeling condition for Set 1 for Ronda before the criterion to increase the prompt delay was met. Initially during the 0-s prompt delay of the variable-modeling condition, Ronda consistently echoed only a proportion of the prompt correctly (e.g., “mouse, frog, mouse” or “dog, mouse, dog”) and did not readily echo the response prompted by the experimenter (e.g., “mouse, frog, dog” when the experimenter modeled “mouse, dog, frog”). Although we assessed participants’ ability to echo during pre-assessments, we did not specifically assess their ability to echo a model consisting of three exemplars. Interestingly, however, Ronda was able to echo three-part response when the prompt remained constant in the rote condition such that an inability to echo a three-part response cannot account for the delay to meeting the criterion in the variable-modeling condition. Instead, delays may have been due to difficulties remembering the three-part response when the model varied. This alternative explanation is consistent with the types of errors that were observed, which often consisted of repeating an exemplar (e.g., “dog, mouse, *dog*”). Similar delays to meeting the criterion for increasing the prompt delay were not observed during Set 2 for Ronda, presumably due to a history of reinforcement for correct responding with a variable three-part response throughout Set 1. Future researchers and clinicians should assess participants’ ability to echo multiple words equal in length to the target response before teaching intraverbal categorization as well as participants’ ability to echo multiple words when the words emitted vary before targeting varied responding during intraverbal categorization.

The temporary effects of the variable-modeling condition on varied responding may not be surprising in light of what we know about the effects of reinforcement. That is, reinforcement increases the future probability of the response that precedes it. When reinforcement is provided for any correct response, that response is likely to occur again in the future, contact reinforcement again, and so on. A within-session analysis of the response sequences provided on each trial (see

Figure 4), revealed that one dominant response sequence emerged both within and across sessions for all participants. The emergence of one dominant response sequence supports the interpretation that the effects of reinforcement may have contributed to the development of invariant responding in the variable-modeling condition. Repeatedly engaging in one dominant response may also represent the least effortful pattern of responding. Thus, it seems unlikely that there would be a sustained increase in varied responding in the absence of contingencies that require, rather than simply allow, varied responding in order to contact reinforcement. Experiment 2 was designed to evaluate the conditions that were sufficient for producing varied responding.

Experiment 2

Varied responding has been shown to be a reinforceable dimension of operant responding (a previously controversial stance; Page & Neuringer, 1985). This discovery led to the development of schedules of reinforcement for varied responding (e.g., Lee et al., 2002; Lee & Sturmey, 2006; Napolitano, Smith, Zarcone, Goodkin, & McAdam, 2010). One such schedule of reinforcement that is particularly amenable to addressing varied responding in applied settings is a lag schedule (Page & Neuringer, 1985). Under a lag schedule, a response is reinforced only if it differs from a set number of preceding responses (e.g., Lag-3 requires that a response differs from the preceding three responses to contact reinforcement). Lag schedules have been used to, for instance, increase varied block building (Napolitano et al., 2010), answers to social questions (e.g., Lee et al., 2002; Lee & Sturmey, 2006), and functional communication (Duker & van Lent, 1991).

Lee et al. (2002) evaluated the effects of a lag schedule on varied responding to a social question (e.g., “What do you like to do?”) with three participants. Differential reinforcement was in place during baseline such that any socially appropriate response (e.g., “I like to watch TV.”) was reinforced and each socially inappropriate response (e.g., “I like to burp.”) was met with a

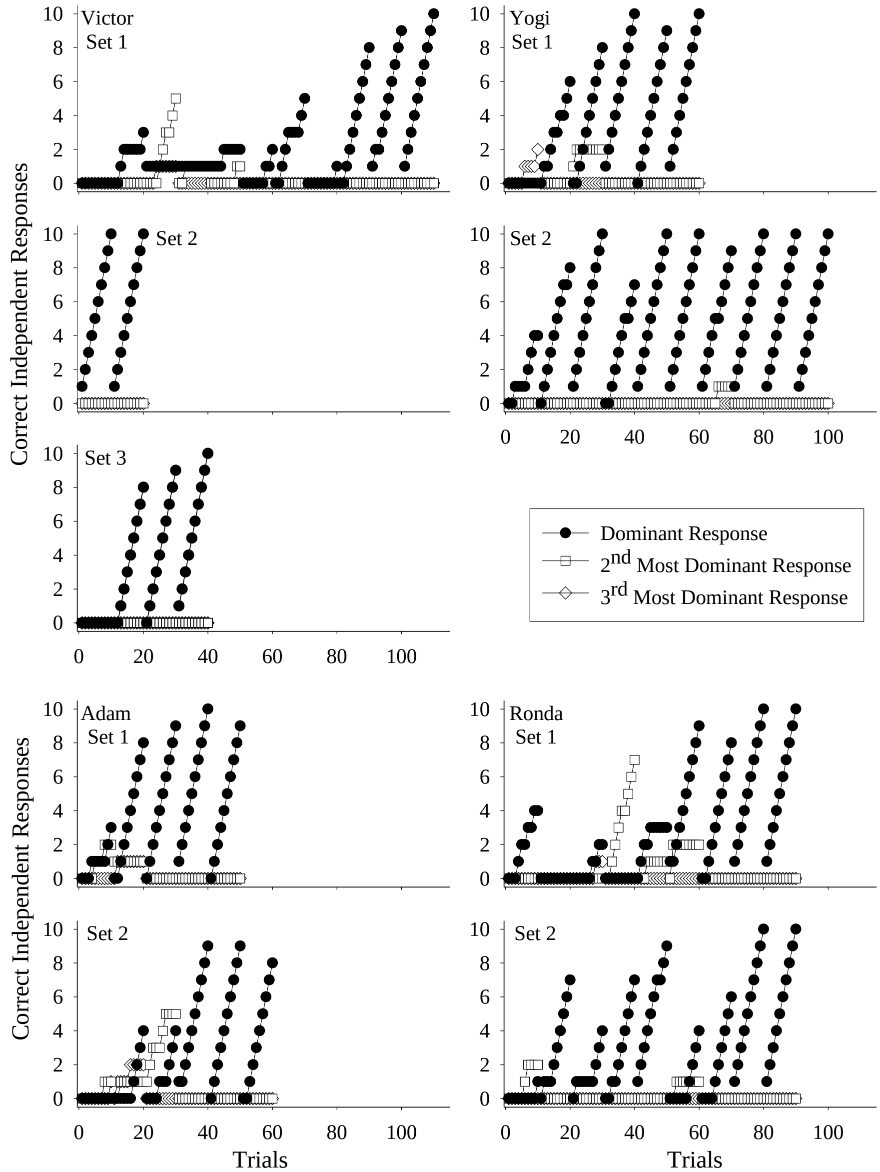

Figure 4. Cumulative record of independent responses during Experiment 1 for each participant. Only a proportion of sessions (beginning with the introduction of the 2-s PPD and ending following three consecutive sessions with only one dominant response) are depicted here. The cumulative record resets every 10 trials (i.e., each session). Only one dominant response occurred for all remaining sessions (not depicted here).

“no” and diverted experimenter attention. During treatment, a Lag-1 schedule was introduced in which reinforcement was provided if the response was appropriate and the content (e.g., the object of the sentence) differed from the preceding response. Differential reinforcement did not produce varied responding for any participant. By contrast, when responding was reinforced on a Lag-1 schedule, there were notable increases in the level of varied responding for two of three participants. This outcome was replicated within a reversal design for both participants. Lee and Sturmey (2006) also observed differential levels of varied responding to the social question “What do you like to do?” with a Lag-1 schedule for two of three participants.

Although a lag schedule has been shown to increase varied responding, results from several studies (e.g., Betz, Higbee, Kelley, Sellers, & Pollard, 2011; Lee et al., 2006; Napolitano et al., 2010) have shown that, in some cases, a lag schedule alone is insufficient to increase varied responding. For example, Napolitano et al. used a Lag-1 schedule to increase block building for six children with ASD. During baseline, all children showed low levels of varied block patterns when intermittent praise was provided for building, regardless of color or form (i.e., for building any block pattern). During the Lag-1 condition, however, varied responding only increased for two of six children. For the other four children for whom there was no effect, modeling of different block structures was combined with the Lag-1 schedule. Specifically, the experimenter had a set of materials identical to those given to the child; following the instruction “build something,” the experimenter modeled building a different block structure and told the participant to build something different. Varied block structures increased for all four participants in this phase, and maintained in the next phase when modeling was removed but the Lag-1 schedule remained in place. Lee et al. (2002) also added a modeling procedure to the lag schedule for a participant for whom the Lag-1 schedule was insufficient to increase varied responding to a social question. During teaching, a second experimenter implemented a progressive-prompt delay to teach 10 responses to the social question. Following three consecutive sessions with 100% correct prompted responses, a delay to providing the model was implemented. However, the introduction of the varied model did not improve varied responding for this participant. Other researchers have simultaneously introduced modeling and the lag schedule. Heldt and Schlinger (2012) sought to increase varied tacts of items in pictures. Participants were shown a page with 10 to 20 images and asked, “What do you see?” During baseline, all correct responses were reinforced, regardless of varied responding. During the intervention phase, varied responding was reinforced on a Lag-3 schedule. During the first session, the experimenter provided a gestural prompt following two trials with an incorrect response. For both participants, varied responding increased under the Lag-3 schedule following the brief history with the prompting procedure. In a similar study, Susa and Schlinger (2012) showed an increase in varied responding to the social question “How are you?” using a changing criterion design in which the lag schedule systematically changed across the criteria. However, because an error-correction procedure in which a vocal model of a varied response was in place during the first session of each change in lag schedule (i.e., first session of Lag-1, Lag-2, and Lag-3), the independent effects of the lag schedule and prompting could not be determined. The results from these studies suggest that, under some conditions, combining a lag contingency with modeling of varied responding may be sufficient to promote varied responding after the models are removed.

Because neither rote- nor variable-modeling conditions was effective in maintaining varied responding in Experiment 1, the purpose of Experiment 2 was to identify the conditions under which increased varied responding would occur. In particular, we evaluated whether a Lag-1 schedule would be sufficient to increase varied responding, and if not, we evaluated the effects of combining the Lag-1 schedule with variable-modeling. Because our participants exhibited low levels of varied responding in Experiment 1, we chose a Lag-1 schedule because the low requirement of varied responding increased the likelihood of contacting reinforcement.

Method

Participants, setting, and materials. All four participants from Experiment 1 participated in Experiment 2. The settings and materials were the same as Experiment 1.

Data collection, dependent variables, and interobserver agreement. Data collection and response definitions were the same as those described in Experiment 1. However, the percentage of trials in which a response differed from the previous trial served as the primary dependent variable in Experiment 2. Response sequences that differed from the previous trial is hereafter referred to as *correct varied responding* and is distinguished from correct responding (which was defined as any three correct exemplars, regardless of the particular response sequence).

A second, independent observer collected data for 59% to 88% of sessions across participants. The procedures for calculating interobserver agreement were the same as Experiment 1. Interobserver agreement for Victor was 98% (range, 70 to 100%), Yogi was 100%, Adam was 99% (range, 90 to 100%), and Ronda was 99% (range, 90 to 100%).

Experimental design. We evaluated the effects of a Lag-1 schedule and Lag-1 schedule plus variable-modeling within a reversal design for Victor, Adam, and Ronda; we used a nonconcurrent multiple baseline design across sets for Yogi.

Procedures. For each participant, we selected one of the four intraverbal categories from Experiment 1 to evaluate in Experiment 2 (see Table 1). Based on patterns of responding for the first target intraverbal category for Yogi, we later included a second target intraverbal category from Experiment 1 to evaluate the effects of our independent variables within a nonconcurrent multiple baseline. We selected responses that the participants were taught previously so we could focus on the conditions necessary to produce varied responding after the target responses had been acquired. Pre-session procedures for selecting a reinforcer were identical to Experiment 1.

DSR BL. This phase was the same as DSR BL in Experiment 1 with two exceptions. First, we examined varied responding for a single intraverbal category (e.g., big animals for Victor), specifically one of the categories mastered in Experiment 1. Second, the response interval was equivalent to the longest prompt delay in Experiment 1 was provided for the participant to respond; a 5-s response interval was used with Victor, Adam, and Ronda, and a 10-s response interval was used with Yogi.

Lag-1 schedule. This phase was identical to the DSR BL except, instead of delivering reinforcement for any correct response, reinforcement (praise and preferred food or item) was delivered according to a Lag-1 schedule. Reinforcement was delivered for the first response sequence of each session, provided that it met the definition of a correct response, because there was no previous response sequence from which to compare varied responding. Following the first trial, the experimenter delivered reinforcement contingent on correct varied responding; that is, reinforcement was provided only if the response sequence was correct and the response sequence on that trial differed from the independent response sequence on the preceding trial. If the independent response sequence on the preceding trial included an error within the exemplars provided (e.g., “frog, dog, chair” in response to “Tell me 3 animals”), the next trial was not eligible for reinforcement because there was not a correct independent response sequence on the previous trial with which to compare. After a response containing three correct exemplars was provided, the lag contingency was reinstated. Contingent on (a) a correct response for which the sequence was the same as the previous trial, (b) an incorrect response, or (c) no response within the prompt delay, the experimenter presented the next trial.

Lag-1 schedule plus variable-modeling PPD (Lag-1 + VAR PPD). Procedures were identical to the Lag-1 schedule, except that contingent on a response sequence that did not meet the Lag-1 schedule, an incorrect response following the prompt (i.e., fewer than three exemplars of the category, or any incorrect exemplar of the category), or no response within the prompt delay, the experimenter modeled a sequence that varied from the previous independent response sequence. The procedures for determining which response to model were identical to Experiment 1. Contingent on a correct varied response following the prompt, the experimenter delivered reinforcement.

Lag-1 schedule plus variable-modeling PPD plus no reinforcement for correct prompted responding (Lag-1 + VAR PPD + No SR P; Victor, Adam, and Ronda). The procedures were identical to the Lag-1 schedule plus variable-modeling except that reinforcement was no longer provided for varied responding that occurred following the prompt.

Results and Discussion

The purpose was to identify the conditions under which increased varied responding would be produced. In particular, we were interested in identifying whether a lag schedule would be sufficient to increase varied responding. We selected one of the categories from Experiment 1 to assess in Experiment 2 so that we could focus on the conditions that were sufficient to observe increased varied responding of a behavior that was already within the participants' repertoire. In addition, because our participants exhibited low levels of varied responding in Experiment 1, we chose a schedule with a low variability requirement (i.e., a Lag-1 schedule as opposed to a Lag-2 schedule) to increase the probability that varied responding would contact reinforcement. Participants' correct responding to the question remained at maintenance levels throughout Experiment 2. More specific, mean correct responding was 99% (range, 80 to 100%) for Victor, 98% (range, 90 to 100%) for Adam, 94% (range, 70 to 100%) for Ronda, and 99% (range, 90 to 100%) for target 1 and 95% (range, 50 to 100%) for target 2 for Yogi (see Figures 5 and 6 for a graphical depiction of correct responding throughout Experiment 2).

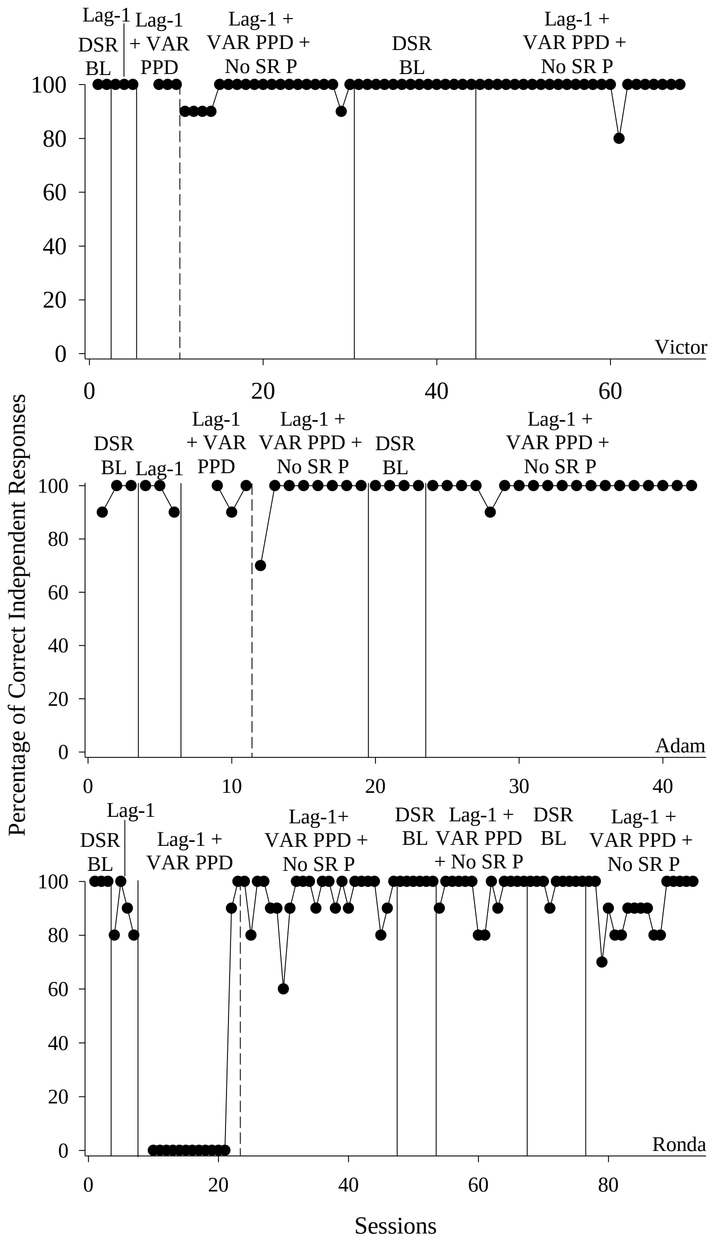

Figure 5. Maintenance data for Victor, Adam, and Ronda in Experiment 2. DSR BL denotes the differential reinforcement baseline condition. Lag-1 + VAR PPD denotes the Lag-1 schedule plus variable-modeling PPD. Lag-1 + VAR PPD + No SR P denotes the Lag-1 schedule plus variable-modeling PPD plus no reinforcement for correct prompted responding. The data that are absent following the introduction of the Lag-1 + VAR PPD phase represent the 0-s phase (i.e., prompted responding is not depicted in this figure).

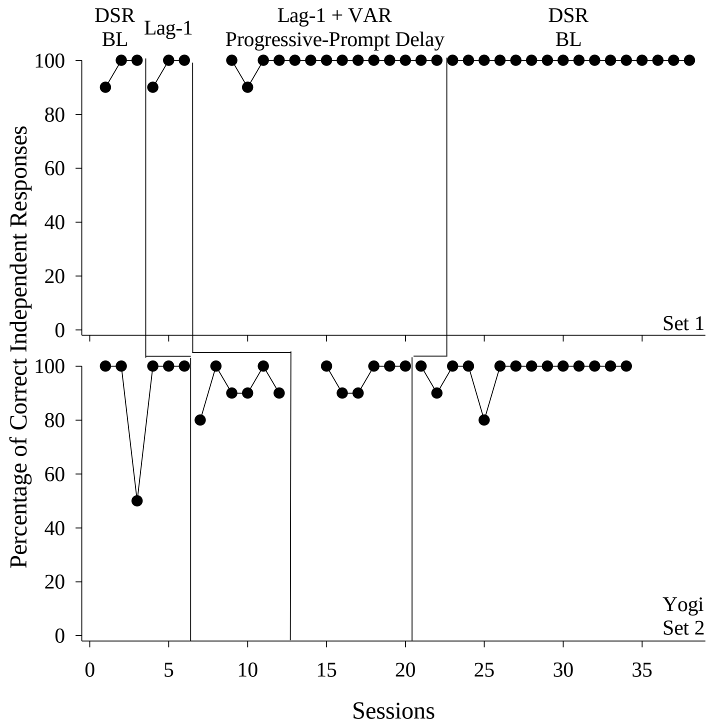

Figure 6. Maintenance data for Yogi in Experiment 2. DSR BL denotes the differential reinforcement baseline condition. Lag-1 + VAR PPD denotes the Lag-1 schedule plus variable-modeling PPD. The data that are absent following the introduction of the Lag-1 + VAR PPD phase represent the 0-s phase (i.e., prompted responding is not depicted in this figure).

Figure 7 shows the percentage of trials that differ from the previous trial for Victor, Adam, and Ronda. Varied responding remained at zero levels during both the DSR BL and the Lag-1 condition. As such, we introduced variable-modeling during which the experimenter modeled a varied response contingent on a response that was the same as the response provided on the previous trial. However, varied responding remained at zero or near-zero levels during this condition, potentially due to the continued availability of reinforcement for correct prompted responding. Karsten and Carr (2009) demonstrated that arranging differential reinforcement for unprompted responses can be used to address prompt-dependent responding during teaching. Thus, in the next phase, all procedures remained the same except that we removed reinforcement for prompted responding; once we made this procedural manipulation, varied responding increased for all three participants. The effect of the Lag-1 schedule plus variable-modeling plus no reinforcement for prompted responding was replicated within a reversal design for Victor, Adam, and Ronda; an additional replication was shown for Ronda.

Figure 8 shows the percentage of trials that differ from the previous trial for Yogi. As with our other three participants, Yogi's responding did not vary under differential reinforcement or with the introduction of the Lag-1 schedule. However, unlike Victor, Adam, and Ronda, varied responding increased for Yogi once the Lag-1 schedule was combined with the variable model; in addition, varied responding did not return to initial baseline levels during the reversal to differential reinforcement. Thus, we adopted a delayed multiple baseline design to identify whether we could replicate (a) the increase in varied responding during the Lag-1 schedule plus variable-modeling phase as well as (b) persistence in varied responding during the reversal to differential reinforcement. With target 2 for Yogi, the effects of each of the phases were replicated, except the final phase, during which varied responding did not persist during the return to DSR BL. The inability to recover baseline levels of responding for Yogi (Figure 8) following a history of reinforcement for varied responding (i.e., following the Lag 1 + VAR PPD phase) may

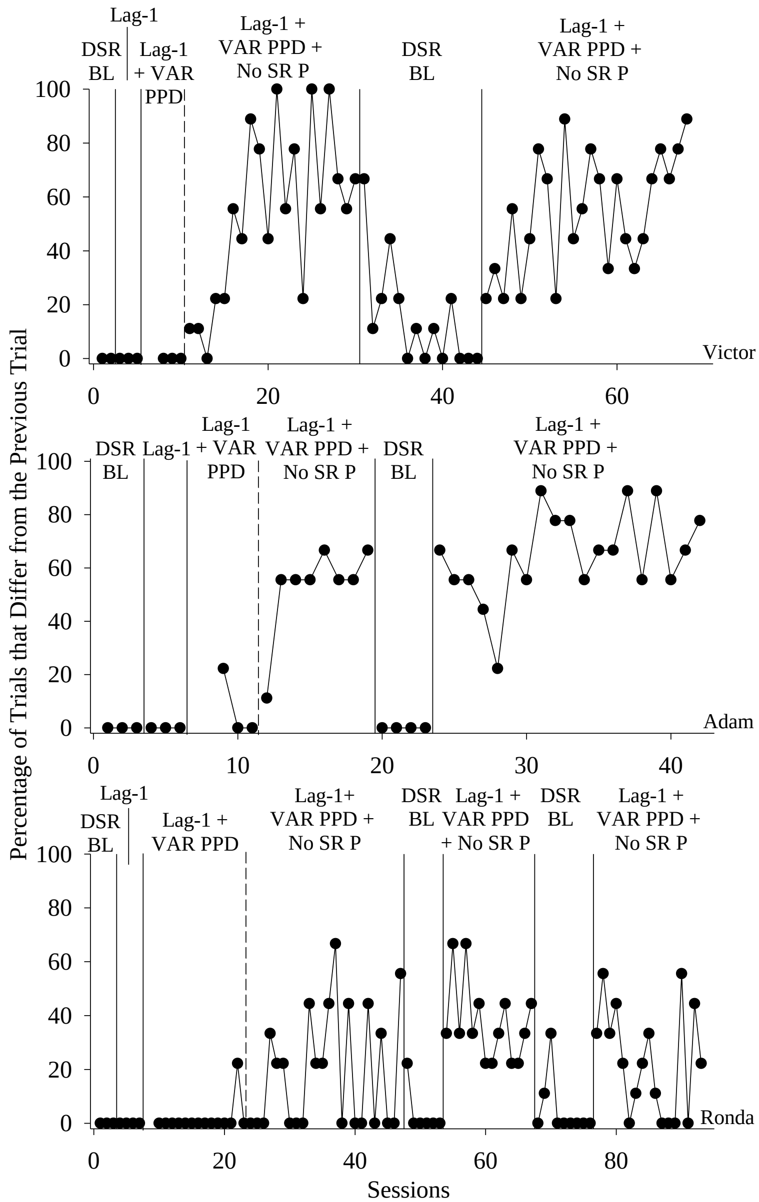

Figure 7. Varied responding as represented by the percentage of trials that differ from the previous trial for Victor, Adam, and Ronda in Experiment 2. DSR BL denotes the differential reinforcement baseline condition. Lag-1 + VAR PPD denotes the Lag-1 schedule plus variable-modeling PPD. Lag-1 + VAR PPD + No SR P denotes the Lag-1 schedule plus variable-modeling PPD plus no reinforcement for correct prompted responding. The data that are absent following the introduction of the Lag-1 + VAR PPD phase represent the 0-s phase (i.e., prompted responding is not depicted in this figure).

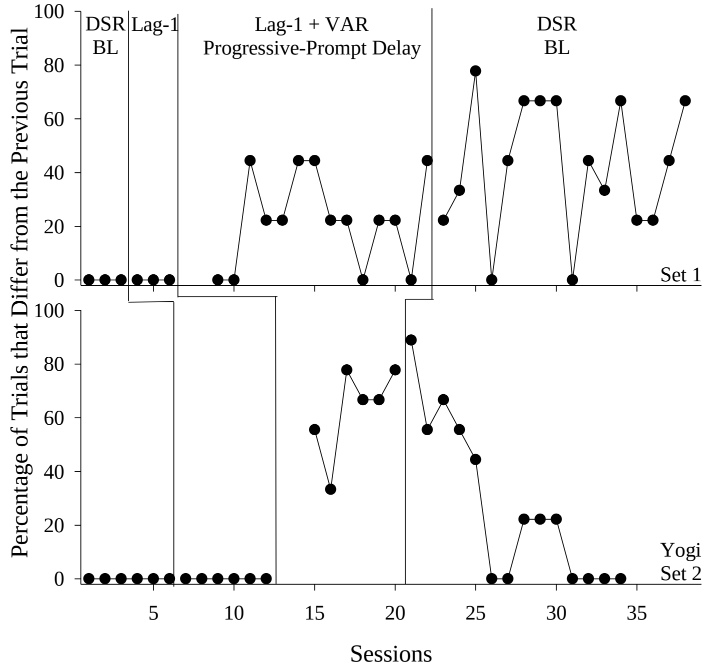

Figure 8. Varied responding as represented by the percentage of trials that differ from the previous trial for Yogi in Experiment 2. DSR BL denotes the differential reinforcement baseline condition. Lag-1 + VAR PPD denotes the Lag-1 schedule plus variable-modeling PPD. The data that are absent following the introduction of the Lag-1 + VAR PPD phase represent the 0-s phase (i.e., prompted responding is not depicted in this figure).

have been due to the fact that varied responding continued to contact reinforcement under variability-independent contingencies. This effect is similar to the results from Miller and Neuringer (2000). Response variability of a computer game with left and right button presses was assessed in a condition that did not require variability to contact reinforcement (i.e., probabilistic contingencies) prior to (PROB1) and following (PROB2) a condition that required response variability. Miller and Neuringer showed that for all five adolescents with autism, four of the five adult controls and three of the four child controls, a recent history of reinforcement for response variability led to higher levels of response variability (i.e., response variability was higher in the PROB2 than in the PROB1 phase). Assuming that contingencies that allow rather than require varied responding are more common in the natural environment (Rodriguez & Thompson, 2015), an understanding of the conditions under which varied responding will persist under contingencies of reinforcement for any correct response has important clinical implications. For example, future research might assess the effects of gradually thinning the schedule of reinforcement for varied responding.

In Experiment 2, extinction (i.e., the Lag-1 schedule alone) was insufficient to increase the varied responding necessary to contact reinforcement under the Lag-1 schedule (Figures 7 and 8). In other words, we did not observe extinction-induced variability. For all participants, varied responding did not increase until variable-modeling (Yogi) or variable-modeling plus no reinforcement for prompted responding (Victor, Adam, and Ronda) was introduced. The effects of the lag schedule in isolation have been mixed in previous applied research. For some participants in previous studies, a lag schedule was sufficient to increase varied responding (e.g., Lee et al., 2002); whereas for others, prompting was necessary prior to observing increased varied responding (e.g., Betz, Higbee, Kelley, Sellers, & Pollard, 2011; Napolitano et al., 2010). In some cases, the lag schedule may have been insufficient to produce varied responding because the participant did not have multiple appropriate responses within their repertoire (e.g., Lee, et al., 2002, see results for Larry; Lee & Sturmey, 2006, see results for Brendan). However, such an explanation is unlikely to apply for our participants as they had exhibited more than one different response sequence at some point during Experiment 1 for the targets in Experiment 2. An alternate explanation for why varied responding was not observed when a lag schedule was introduced is that our participants did not have a history of reinforcement for varied responding. That is, varied responding may be more likely once invariant responding contacts extinction (e.g., during a lag schedule) if the individual has a previous history of reinforcement for varied responding under such conditions.

We also depicted varied responding as the number of unique response sequences to obtain a more complete picture of the type of varied responding that was emitted during Experiment 2. Figure 9 shows the number of unique response sequences for Victor, Adam, and Ronda. The number of unique response sequences ranged from 1 to 3 for Victor, 1 to 4 for Adam, and 1 to 3 for Ronda. Figure 10 shows the number of unique response sequences for Yogi, with a range of 1 to 2 response sequences for target 1 and a range of 1 to 3 response sequences for target 2. Two or three dominant response sequences emerged for all participants (see Table 2 for details). For all participants, two response sequences comprised more than 80% of the response sequences provided throughout Experiment 2. Victor, for example, contacted reinforcement on 100% of trials during sessions 21, 25, and 27 by alternating between two response sequences throughout the session. This may be considered a form of higher-order rote responding, which other researchers have suggested could be addressed with larger lag schedules (e.g., Lag-3; Lee & Sturmey, 2006).

General Discussion

In summary, the results from Experiment 1 suggest that although there may be some initial varied responding when providing variable-models during teaching, such effects are likely to be

Table 2

The Percentage of Sessions during Experiment 2 that each of the two most frequent Response Sequences comprised.

Participant	Response Sequence 1		Response Sequence 2	
	Response Sequence	% of Responses	Response Sequence	% of Responses
Victor	Bear Whale Horse	48.4%	Horse Whale Bear	48.0%
Yogi	Shovel Rake Lawnmower	58.9%	Rake Lawnmower Shovel	40.3%
	Tennis Baseball Fishing	49.4%	Baseball Tennis Fishing	44.7%
Adam	Snowplow Meow Eyebrow	60.0%	Eyebrow Meow Snowplow	22.0%
Ronda	Milk Juice Water	67.5%	Juice Milk Water	16.7%

Note. Yogi had two sets in Experiment 2. The top category was the first set, and the bottom category was the second set.

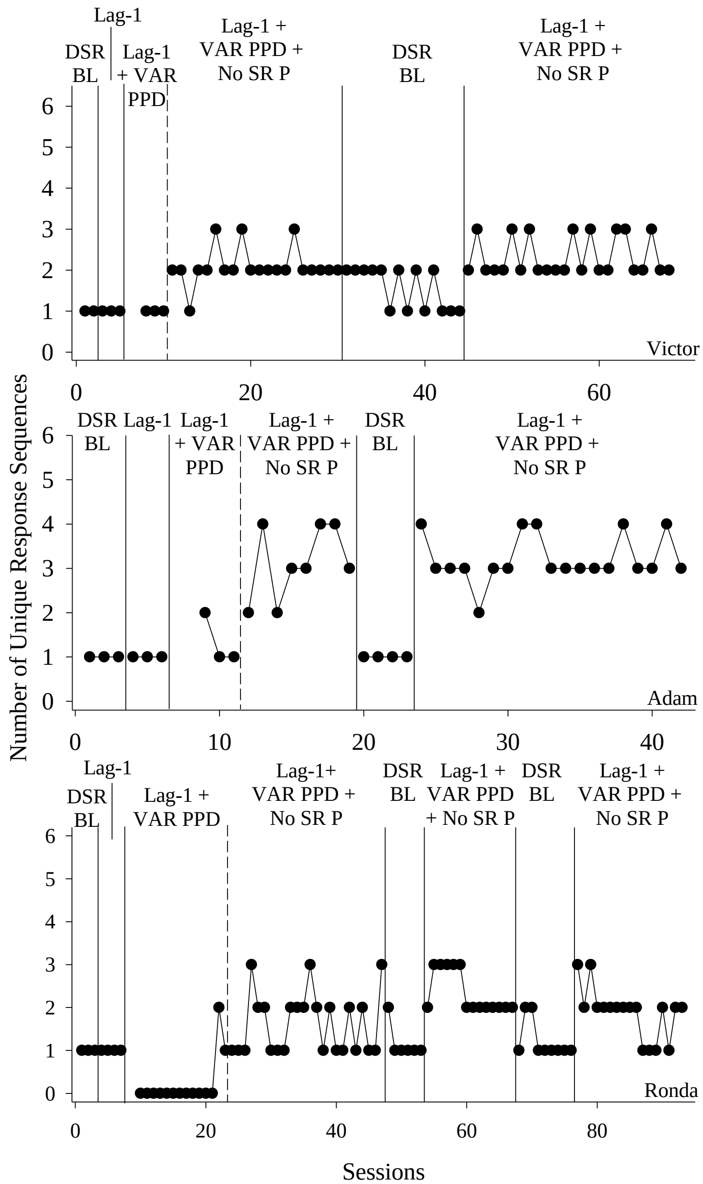

Figure 9. Varied responding as represented by the number of unique response sequences for Victor, Adam, and Ronda in Experiment 2. DSR BL denotes the differential reinforcement baseline condition. Lag-1 + VAR PPD denotes the Lag-1 schedule plus variable-modeling PPD. Lag-1 + VAR PPD + No SR P denotes the Lag-1 schedule plus variable-modeling PPD plus no reinforcement for correct prompted responding. The data that are absent following the introduction of the Lag-1 + VAR PPD phase represent the 0-s phase (i.e., prompted responding is not depicted in this figure).

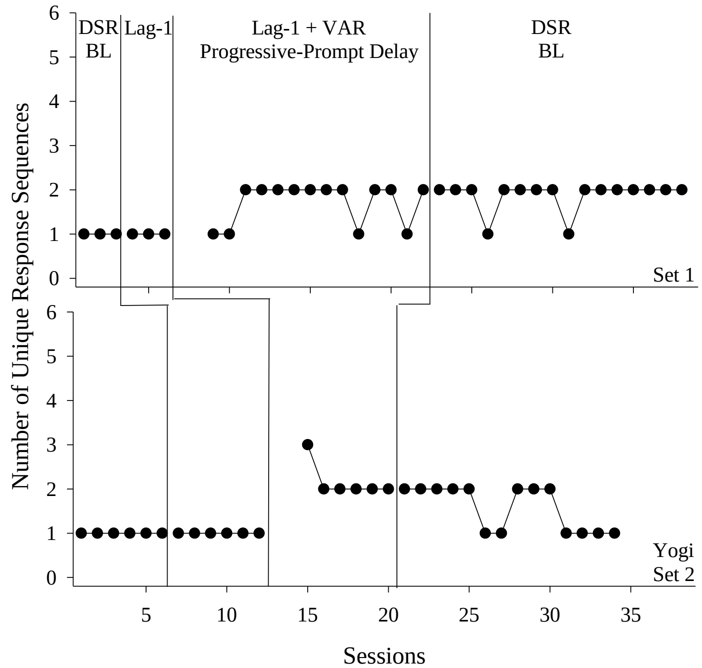

Figure 10. Varied responding as represented by the number of unique response sequences for Yogi in Experiment 2. DSR BL denotes the differential reinforcement baseline condition. Lag-1 + VAR PPD denotes the Lag-1 schedule plus variable-modeling PPD. The data that are absent following the introduction of the Lag-1 + VAR PPD phase represent the 0-s phase (i.e., prompted responding is not depicted in this figure).

temporary. Thus, it seems unlikely that the procedures typically used during DTI (e.g., prompting and reinforcing one target response per discriminative stimulus) uniquely contribute to rote responding in individuals with autism. In other words, invariant responding will likely occur whether or not therapists provide variable-models during teaching. In addition, providing variable-models during teaching may lead to slightly slower acquisition. Because (a) there does not appear to be an advantage to providing variable-models during teaching and (b) there is a potential disadvantage in terms of speed of acquisition of any one correct response, clinicians should consider first teaching correct responding (e.g., in succession) before targeting varied responding. This suggestion is also in line with the results from Experiment 2, which showed that varied responding could be increased by introducing a lag schedule plus variable-models following the acquisition of at least one target response. An alternative approach may be to capitalize on the varied responding observed in the initial stages of teaching by programming differential magnitudes of reinforcement for correct responding (e.g., one edible) versus correct responding that is different from the response in the previous trial (e.g., two edibles). However, the effects of arranging differential magnitudes of reinforcement on varied responding and acquisition when variable-models are provided during teaching have not yet been evaluated. Future research might evaluate this strategy with participants whose behavior has been shown to be sensitive to differential magnitudes of reinforcement (e.g., after completing a preassessment in which participants are presented with a choice between consuming one or two edibles).

As previously noted, we adopted a translational approach in which the target skill (intraverbal categorization) was part of our participants' early intervention goals, but the response targeted for varied responding (response sequence) was more aligned with basic research (e.g., Neuringer, Kornell, & Olufs, 2001). Evaluating the level of varied responding in the order of exemplars provided during intraverbal categorization allowed us to evaluate the effects of our independent variables within a relatively simple experimental arrangement. In addition, using response sequence as our dependent variable for evaluating varied responding allowed for the occurrence of recombinative generalization (Axe & Sainato, 2010). That is, after the participants learned three exemplars for an intraverbal category, they could presumably emit the exemplars in any order without direct teaching of different response sequences. Toward this potential outcome, we purposely modeled only four of the six possible response sequences to be able to evaluate whether participants would emit response sequences that had never been modeled. A within-session analysis of response patterns in Experiment 2 (data available from the first author) showed that recombinative generalization only occurred 3 times for Victor (less than 1% of trials) and never occurred for Yogi or Ronda. By contrast, Adam emitted one of the two response sequences that had never been modeled 240 times and the other response sequence once (for a total of 60% of trials). Thus, despite the fact that we programmed favorable conditions for observing varied responding without extensive exposure to variable-models, lag schedules, or both, relatively few different response sequences were observed in Experiments 1 and 2.

We have depicted multiple measures of varied responding (i.e., number of unique response sequences, percentage of trials that differ from the previous trial) to highlight the different information provided by doing so. We used the number of unique response sequences as our primary dependent variable in Experiment 1 because it gave a summative depiction of varied responding that occurred initially while also showing that children engaged in only a single response sequence as the experiment continued. We included the percentage of trials that differed from the previous trial as a measure of varied responding for Experiment 2 because it corresponded with the contingencies arranged in the Lag-1 schedule. That is, correct varied responding on the Lag-1 schedule was measured as the percentage of trials that differed from the previous trial; if all trials differed from the previous trial during the Lag-1 schedule (100%), then responding would have contacted reinforcement on 100% of trials. In the absence of data on socially acceptable levels of varied responding across types of responding (e.g., initiating a conversation), depicting varied responding in multiple ways may prove useful as researchers continue to refine methods of teaching and reinforcing varied responding, identify treatments that support the persistence of varied responding in the absence of direct reinforcement for varied responding, and identify socially acceptable levels of varied responding across types of responses.

Depicting multiple measures of varied responding may also provide a more complete picture of the type of varied responding that is being emitted, with a cumulative record providing the most detailed measure of varied responding but perhaps the more cumbersome measure given that the number of data points would be multiplied by the number of trials in a session (in our case, 10) and the number of possible response sequences (in our case, at least 6 data paths). As a supplementary measure, we graphed cumulative responding for the three most common response sequences for a proportion of the sessions for all participants in Experiment 1 (Figure 4). Specifically, we graphed cumulative response sequences for a subset of sessions, starting with the introduction of the PPD until there were three consecutive sessions with only one dominant response sequence to show the limited within-session varied responding.

One limitation of the current study was that prompts were never fully removed following the increase of varied responding in Experiment 2. For all four participants, there was an increase in varied responding in the presence of the Lag-1 schedule, but we do not know whether varied responding would have persisted in the absence of variable-modeling. Research in this area has been mixed; Napolitano et al. (2010) removed prompts and observed persistent varied responding with a block building task when only the Lag-1 schedule remained in effect. By contrast, when Lee et al. (2002) removed prompts but maintained a Lag-1 schedule, varied responding to social questions did not persist. In addition, the fact that we did not evaluate maintenance of varied responding in the absence of the lag schedule is a limitation of our study. Regarding clinical utility of the current procedures, it will be important to assess maintenance of varied responding in the absence of the variable model and programmed contingencies, and to identify the natural conditions under which such maintenance will occur.

Yogi's data in Experiment 2 (Figure 8, top panel) suggest that there may be some conditions under which varied responding will persist under conditions that allow but do not require variability. Although baseline levels of varied responding did not recover when baseline contingencies were reinstated in the last phase of the top panel (Figure 8) is a limitation of the current study, it is an important clinical finding that should be considered in future research.

Given the propensity of individuals with ASD to engage in restricted and repetitive behavior and the results of the current study showing that in the absence of reinforcement for varied responding invariant responding is likely to occur, it would be useful to evaluate whether the temporary effects of variable modeling would also be observed in children of typical development. That is, are the fleeting effects of variable modeling during contingencies that allow but do not require variability unique to individuals with ASD? A group comparison across individuals with ASD and peers matched for age and level of functioning would allow for an evaluation of the generality of our findings across populations.

Reference

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

Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology, 42*, 273-281.

Axe, J. B., & Sainato, D. M. (2010). Matrix training of preliteracy skills with preschoolers with autism. *Journal of Applied Behavior Analysis, 43*, 635-652.

Betz, A. M., Higbee, T. S., Kelley, K. N., Sellers, T. P., & Pollard, J. S. (2011). Increasing response variability of mand frames with script training and extinction. *Journal of Applied Behavior Analysis, 44*, 357-362.

Billingsley, L. T., & Romer, F. F. (1983). Response prompting and the transfer of stimulus control: Methods, research, and a conceptual framework. *Journal of the Association for the Severely Handicapped, 8*, 3-12.

Cihon, T. (2007). A review of training intraverbal repertoires: Can precision teaching help? *The Analysis of Verbal Behavior, 23*, 123-133.

Cummings, A. R., & Williams, W. L. (2000). Visual identity matching and vocal imitation training with children with autism: A surprising finding. *Journal on Developmental Disabilities, 7*, 109-122.

DeLeon, I. G., Fisher, W. W., Rodriguez-Catter, V., Maglieri, K., Herman, K., & Markhefka, J. M. (2001). Examination of relative reinforcementeffects of stimuli identified through pretreatment and daily brief preference assessments. *Journal of Applied Behavior Analysis, 34*, 463-473.

Duker, P. C., & van Lent, C. (1991). Inducing variability in communicative gestures used by severely retarded individuals. *Journal of Applied Behavior Analysis, 24*, 379-386.

Gutierrez, A., Hale, M. N., O'Brien, H. A., Fischer, A. J., Durocher, J. S., Alessandri, M. (2009).

Evaluating the effectiveness of two commonly used discrete trial procedures for teaching receptive discrimination to young children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 3*, 630-638.

Heldt, J., & Schlinger, H. D. (2012). Increased variability in tacting under a lag 3 schedule of reinforcement. *The Analysis of Verbal Behavior, 18*, 131-136.

Jahr, E., & Eldevik, S. (2002). Teaching cooperative play to typical children utilizing a behavior modeling approach: A systematic replication. *Behavioral Interventions, 17*, 145-157.

Karsten, A. M., & Carr, J. E. (2009). The effects of differential reinforcement of unprompted responding on the skill acquisition of children with autism. *Journal of Applied Behavior Analysis, 42*, 327-334.

Leaf, J. B., Tsuji, K. H., Lentell, A. E., Dale, S. E., Kassardjian, A., Taubman, M., McEachin, J., Leaf, R., & Oppenheim-Leaf, M. L. (2013). A comparison of discrete trial teaching implemented in a one-to-one instructional format and in a group instructional format. *Behavioral Interventions, 28*, 82-106.

Lee, R., McComas, J. J., & Jawor, J. (2002). The effects of differential and lag reinforcement schedules on varied verbal responding by individuals with autism. *Journal of Applied Behavior Analysis, 35*, 391-402.

Lee, R., & Sturmey, P. (2006). The effects of lag schedules and preferred materials on variable responding in students with autism. *Journal of Autism and Developmental Disorders, 36*, 421-428.

Miller, N., & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. *Journal of Applied Behavior Analysis 33*, 151-165.

Napolitano, D. A., Smith, T., Zarcone, J. R., Goodkin, K., & McAdam, D. B. (2010). Increasing response diversity in children with autism. *Journal of Applied Behavior Analysis, 43*, 265-271.

Neuringer, A. Kornell, N. & Olufs, M. (2001). Stability and variability in extinction. *Journal of*

Experimental Psychology: Animal Behavior Processes, 27, 79-94.

Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 429-452.

Rodriguez, N. M., & Thompson, R. H. (2015). Behavioral variability and autism spectrum disorder. *Journal of Applied Behavior Analysis, 48*, 167-187.

Saldana, R. L., & Neuringer, A. (1998). Is instrumental variability abnormally high in children exhibiting ADHD and aggressive behavior? *Behavioural Brain Research, 94*, 51-59.

Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children, 8*, 67-76.

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

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Susa, C., & Schlinger, H. D. (2012). Using a lag schedule to increase variability of verbal responding in an individual with autism. *The Analysis of Verbal Behavior*, *28*, 125-130.

Weiss, M. J., LaRue, R. H., & Newcomer, A. (2009). Social skills and autism: Understanding and addressing the deficits. In J. Matson (Ed.) *Applied Behavior Analysis for Children with Autism Spectrum Disorders* (pp. 129-144). New York: Springer.

Wong, S. E., & Woolsey, J. E. (1989). Re-establishing conversational skills in overtly psychotic, chronic schizophrenic patients. *Behavior Modification*, *13*, 415-430.