Journal of
Applied Behavior Analysis

Journal of Applied Behavior Analysis 2020, **9999,** 1–19 NUMBER 9999 ()

Effects of time delay and requiring echoics on answering questions about visual stimuli

Olga Meleshkevich

Department of Behavior Analysis, Simmons University and ABA Consulting, Inc., Westborough, MA

Judah B. Axe

Department of Behavior Analysis, Simmons University

Francesca degli Espinosa

Department of Medicine and Surgery, University of Salerno, Italy and ABA Clinic, United Kingdom

An important communication skill for children with autism is answering multiple questions about visual stimuli (e.g., "What is it?" "What color is it?"). We targeted answering "What number?" and "What shape?" in the presence of numbers inside shapes, and "What is it?" and "What color?" in the presence of colored objects (e.g., a yellow cat) with 3 preschoolers with autism. In addition to a progressive time delay, we required the participants to answer the questions by echoing a keyword from the questions. For example, we taught them to answer, "What color?" with "color blue." In the context of a multiple-probe design across behaviors within a multiple-probe design across participants, the procedure was effective in increasing trained responses and producing within- and across-category generalization. The echoic may have facilitated the responses by increasing the salience of the auditory stimuli and strengthening intraverbals within autoclitic frames.

Key words: autoclitic frame, echoic, intraverbal, multiply controlled verbal behavior, tact

Multiply controlled verbal behavior is an integral part of language development (Michael et al., 2011; Pence Turnbull & Justice, 2012). Examples include tacting multiple visual stimuli (Miguel et al., 2005), emitting intraverbals under the control of multiple verbal stimuli (i.e., convergent multiple control; DeSouza et al., 2019; Devine et al., 2016), and emitting multiple verbal responses under the control of a verbal stimulus (i.e., divergent multiple control; Glodowski & Rodriguez, 2019). Multiply controlled verbal behavior occurs routinely when parents tact and ask questions about object features while holding them up to their young children (e.g., "This is a doll. It has two eyes. What else does it have?"; Floor & Akhtar, 2006; Miguel, 2016). Sometimes parents wait to initiate verbal interactions about an object until they notice their child paying attention to it by pointing to or playing with it (Horne & Lowe, 1996). For children with autism, this language learning does not happen purely incidentally (Tager-Flusberg et al., 2011), and there is limited research on teaching multiply controlled verbal behavior to children with autism (DeSouza et al., 2017; Stauch et al., 2017).

Two behavior analytic accounts of language learning suggest that echoic responses facilitate

We thank Melanie McCarthy-Pepin, Jason Pepin, and Taylor Murphy of Behavioral Connections for their invaluable on-site support of this study. We also thank Gretchen Dittrich and Philip Chase for their feedback on earlier versions of this manuscript. We have no conflicts of interest relevant to this article to disclose.

Address correspondence to: Judah B. Axe, Department of Behavior Analysis, Simmons University, Room W304A, 300 The Fenway, Boston, MA 02115. Email: judah.axe@simmons.edu

doi: 10.1002/jaba.790

© 2020 Society for the Experimental Analysis of Behavior (SEAB)

the acquisition of multiply controlled verbal behavior. In the *naming* account, when an adult names an object, the child echoes the word, which is part of the explanation for tacting the object later without direct training (Horne & Lowe, 1996). In the analysis of *joint control*, selecting the correct comparison in an auditory–visual conditional discrimination task (i.e., receptive discrimination) is facilitated by a match between the product of the echoic of the spoken sample and the product of the tact of the correct comparison (Lowenkron, 2006). Researchers have examined the role of the echoic in facilitating multiply controlled verbal behavior, including receptive discrimination (Charlop, 1983; Petursdottir et al., 2014) and intraverbals (Kisamore et al., 2013, 2016).

Receptive discrimination is multiply controlled by the verbal stimulus (e.g., "Find dog") and the nonverbal stimulus (e.g., the picture of the dog). Charlop (1983) demonstrated the efficacy of requiring children with autism and echolalia to echo the "keyword" before the selection. For example, a participant heard, "Find boat," repeated "boat," and then pointed to the boat. Causin et al. (2013) presented 10-12 picture cards and asked children with autism to hand over three pictures (e.g., "Give me the horse, pig, and sheep"). The procedure, based on joint control, required the participants to repeatedly echo and self-echo the three keywords (e.g., "horse, pig, sheep").

Researchers have targeted multiply controlled intraverbals that involve discriminating questions such as, "What's an animal that's red?" and "What's a vehicle that's red?" (Kisamore et al., 2016), as well as "What do you sweep?" and "What do you sweep with?" (Ingvarsson et al., 2016). Correct intraverbal responding to such questions requires the response to come under the multiple control of two verbal stimuli (e.g., "animal" and "red," "sweep" and "with"; Axe, 2008; Eikeseth & Smith, 2013; Sundberg & Sundberg, 2011). For example, when teaching intraverbal responses to question pairs such as, "Name the opposite of tall" and "Name the same as tall" to typically developing preschoolers, Kisamore et al. (2013) first showed that constant time delay, differential reinforcement, error correction, and errorless teaching were ineffective. They then required the differential observing response (DOR; Farber et al., 2017; Fisher et al., 2019) of echoing the instruction. For example, the instructor stated, "Name the opposite of tall" and then, "You say it." Echoing "opposite tall" facilitated correct intraverbal responding. Kisamore et al. (2016) extended these results with a DOR of echoing keywords (e.g., saying, "fruit green" after "Tell me a fruit that's green"). Echoing the keywords may have enhanced their salience and led to correct intraverbal responses.

For another type of multiply controlled verbal behavior, consider a child who is asked "What is it?" in the presence of a blue car and "What color?" in the presence of a red ball. In each case, a correct response must be under the stimulus control of a verbal (i.e., the question) and a nonverbal (i.e., the object) stimulus or stimulus property (e.g., the color [i.e., a conditional discrimination]). The question, "What color?" evokes "red" only if both the verbal stimulus (i.e., "color") and the nonverbal property of the ball (i.e., red) multiply control the response (Michael et al., 2011). This multiply controlled response is part intraverbal and part tact (i.e., an *intraverbal-tact*; Bondy et al., 2004) and is addressed in Milestone 11 in the tact section of the *Verbal Behavior Milestones Assessment and Placement Program* (VB-MAPP; Sundberg, 2008). If a child responds incorrectly, such as saying "ball" when asked "What color?", the verbal stimulus may not be exerting stimulus control over the response, a situation indicative of *restricted stimulus control* or *stimulus overselectivity* (Dube et al., 2010, 2016).

degli Espinosa et al. (2020) evaluated a three-step procedure for teaching intraverbal-tacts to two children with autism. They targeted three groups of questions: number versus color (i.e., "What number?" and "What color?" when presenting different colored numbers), object versus color (i.e., "What is it?" and "What color?" when presenting different colored objects), and animal versus animal sound (i.e., "What is it?" and "What does it say?" when presenting animals). The first step of the intervention was echoic training with relevant verbal stimuli (e.g., echoing "color green," "number two," "it's a cat," "it says meow"). This procedure aimed to establish intraverbal control between the fixed class terms (e.g., "color") and variable terms (e.g., "green"). The second step was simultaneous simple discrimination training in which each session contained one question in the presence of multiple pictures and prompts to answer with an echoic of the keyword in the question and the tact (e.g., saying "color blue" in the presence of "What color?" and a blue card). The purpose of this step was to establish control between the verbal antecedent (e.g., colors) and the non-verbal property of the visual stimulus (e.g., blue). The third step was intraverbal-tact training in which the experimenter asked two questions in the same session (e.g., "What color?" and "What number?"), randomly rotated across compound stimuli. As demonstrated in a multiple-baseline design across behaviors, this three-step training procedure resulted in the acquisition of intraverbal-tact responses, as well as generalization within (e.g., novel colors) and across (e.g., novel visual stimuli) stimulus classes.

Responding with the echoic and the tact, such as "color blue," may be described as responding with an autoclitic frame that has formal structure and temporal contiguity of words (Palmer, 2007; Skinner, 1957). The structure is the functional feature of an autoclitic frame that provides additional intraverbal control (i.e., temporal contiguity) for answering questions. In other words, reinforcement for echoing a *class word* (e.g., "number") in temporal contiguity with *members-of-that-class words* (e.g., "four," "seven") strengthens the intraverbal relations and future probability of emitting member words matching stated class words (Palmer, 2007).

Considering the potential facilitative effects of echoic behavior on listener and intraverbal behavior, the purpose of the current study was to examine the effects of requiring echoic responses in the context of time delay on trained and novel intraverbal-tacts with children with autism. This study extended degli Espinosa et al. (2020) in several ways. First, we omitted the initial echoic training phase and prompted echoics in the context of intraverbal-tact training. Second, we used the conditional-only method of conditional discrimination training rather than the simple-conditional method (Grow et al., 2011). Third, we used a multiple-probe design across behaviors within a multiple-probe design across participants.

Method

Participants

The three participants were diagnosed with autism and received center-based behavior analytic services for 4-6 hr per day, 5 days per week, for at least a year prior to the study. All participants had strong attending skills, such as orienting to the instructor, looking at materials, and shifting gaze between materials and the instructor. Mark was 5 years 1 month and scored at the beginning of Level 3 (score of 122) of the VB-MAPP (Sundberg, 2008). Because strong listener responding was required in this study, we report that Mark mastered Milestone 12 (M-12) for Listener Responding (LR) and M-8 for Listener Responding by Feature, Function, and Class (LRFFC). Mark was multiracial, spoke English at home, and lived with his mother, father, older sister, and older brother. His mother was a business owner, and his father was a ground agent for an airline.

David was 5 years 8 months and scored at the beginning of Level 3 (score of 133) of the VB-MAPP. He mastered M-14 for LR and M-13 for LRFFC. David was African American, spoke English at home, and with his older sister, spent equal time with his separated mother and father. David's mother had attended college and worked for an insurance company and his father held a master's degree and taught at a community college. Jack was 3 years 7 months and scored at Level 2 (score of 75) of the VB-MAPP. He mastered M-14 for LR and M-8 for LRFFC. Jack was Caucasian, spoke English at home, and lived with his mother, father, older brother, and paternal grandmother. Jack's parents held master's degrees and were special education teachers. Part of Mark's and Jack's education was spent in an integrated preschool with assistance from a Registered Behavior Technician.

Preassessments

To be included in the study, participants had to (a) echo two-word phrases (e.g., "love mommy," "shape circle"), as well as one four-word phrase (i.e., "it is a cat"), with 100% accuracy; (b) tact the colors, numbers, shapes, and objects targeted in the study with 100% accuracy; (c) emit approximately 150-200 tacts; (d) demonstrate at least a "moderate problem" on the "Failure to Make Conditional Discriminations" barrier of the VB-MAPP; and (e) demonstrate below 75% correct responses during two 12-trial preassessment intraverbal-tact sessions with visual stimuli not targeted in the study (see Figures 1 and 2).

Preassessments of these repertoires consisted of the following methods. First, the experimenter assessed echoic behavior by asking each participant to repeat 15 phrases similar in length to the responses targeted in the study (e.g., "number five," "it is a cat"). Second, because we did not want to expose the participants to the questions used in the experiment (e.g., "What color?"), we assessed tact repertoires using a pointing procedure. To test tacts of colors, the experimenter placed six colored cards (brown, yellow, green, purple, blue, and red) on the table in front of the participant and modeled tacting the first two cards while pointing to them. Then the experimenter pointed to the third card (and so on) and waited for the participant to tact the color. The mastery criterion was tacting each stimulus correctly twice (i.e., 12 trials; did not include modeled trials). The experimenter used the same procedure to test tacting six numbers by presenting cards with black numbers (1-6) on a white background, shapes by presenting cards with black shapes (circle, diamond, star, square, triangle, and oval) on a white background, and objects by presenting two-dimensional visual stimuli (pictures of: purple cup, blue cat, red flower, green spoon, yellow hat, brown car).

Third, to assess intraverbal-tacts, the experimenter presented one picture at a time in front of the participant (marked PA in Figures 1 and 2) and asked, "What shape?" and "What number?" in the first 12-trial session, and "What is it?" and "What color?" in the second 12-trial session. All procedures were the same as the probe procedures (see below), except there were no interspersed mastered trials. The experimenter did not deliver prompts and said, "Good job" after correct and incorrect responses.

Setting and Materials

The study took place at the participants' center-based program, a primarily insurance-funded, for-profit organization providing ABA services located in a suburb in Southeast Massachusetts. The center has six ABA classrooms with 5-6 students in each class, as well as a state-approved inclusive preschool program with two classrooms. Sessions were conducted

Figure 1
Distribution of Stimuli Across Conditions for Category 1 (Shape/Number)

Shape / Number	Circle	Triangle	Square	Oval	Star	Diamond
1	Trained		Gen1			PA
2	Gen1	Trained			PA	
3		Gen1	Trained	PA		
4	PA			Gen2		
5		PA			Gen2	
6			PA			Gen2

Note. Trained = Trained stimuli; Gen1 = generalization stimuli with one component trained; Gen2 = generalization stimuli with both components untrained; PA = preassessment stimuli.

in a quiet room containing a table with chairs, tangible reinforcers (e.g., toy train, iPad), edible reinforcers (Goldfish, gummy bears), picture cards (10 x 10 cm), a video camera, and data sheets with the prescribed sequence of questions and visual stimuli.

Category 1 stimuli (shape/number) were organized into a matrix table (see Figure 1) and consisted of six numbers (1-6) written on six shapes (circle, triangle, square, oval, star, and diamond; see Figure 3). Category 2 stimuli (object/color) were organized into a matrix table (see Figure 2) and consisted of colored objects (e.g., yellow cat) created from six objects (cup, cat, flower, spoon, hat, and car) and six colors (brown, yellow, green, purple, blue, and red; see Figure 4). Cross-category stimuli included three sets of stimuli with visual components combined from Categories 1 and 2 (see Figure 5).

As displayed in Figures 1 and 2, we distributed stimuli in Categories 1 and 2 evenly into three subsets: (a) both visual components trained (Trained); (b) generalization with one visual component trained (Gen1); and (c) generalization with no visual components trained (Gen2).

Figure 2
Distribution of Stimuli Across Conditions for Category 2 (Object/Color)

Object / Color	Cup	Cat	Flower	Spoon	Hat	Car
Brown	Trained		Gen1			PA
Yellow	Gen1	Trained			PA	
Green		Gen1	Trained	PA		
Purple	PA			Gen2		
Blue		PA			Gen2	
Red			PA			Gen2

Note. Trained = Trained stimuli; Gen1 = generalization stimuli with one component trained; Gen2 = generalization stimuli with both components untrained; PA = preassessment stimuli.

Figure 3
Category 1 Stimuli Across the Trained, Gen1, and Gen2 Subsets

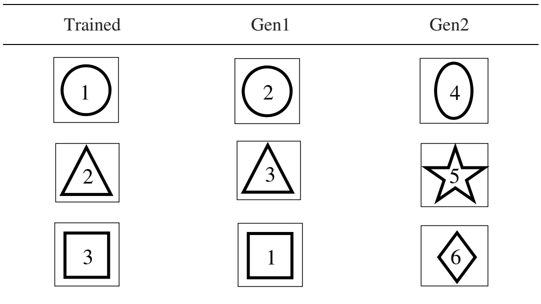
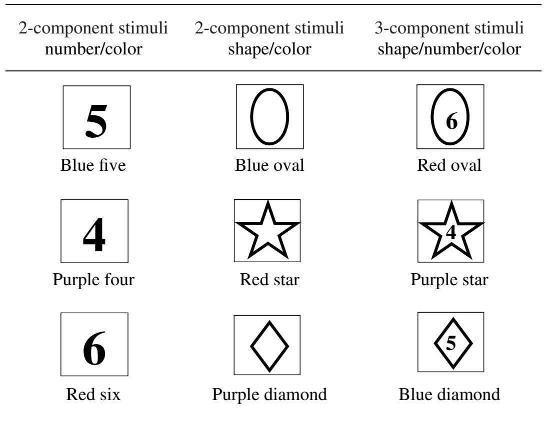


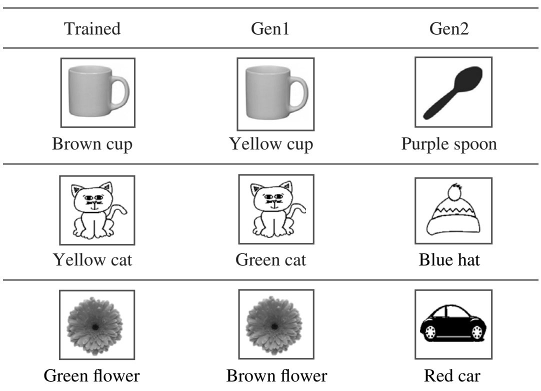
Figure 5
Cross-Category Stimuli



Dependent Variable

The primary dependent variable (DV1) was the percentage of correct intraverbal-tacts during probes. A correct response was defined as saying a word corresponding to the feature of the visual stimulus indicated by the question within 5 s. For example, when shown a picture of a yellow cat and asked, "What color?" the participant responded, "yellow." An incorrect response would be saying "blue" or "cat." Even though in training we required the echoic and tact (e.g., "color yellow"), a response in the probe with or without the echoic was considered correct. A secondary dependent variable (DV2) was the percentage of responses with correct question discrimination. For example, when shown a picture of a yellow cat and asked, "What color?" the participant responded, "green." This was correct for this dependent variable because the participant discriminated the question by saying a color, whether or not the stated color matched the visual stimulus. An incorrect response for DV2 would be saying, "cat" because there was incorrect question discrimination (i.e., saying the object instead of the color). The instructor conducting sessions collected the data.

Figure 4
Category 2 Stimuli Across the Trained, Gen1, and Gen2 Subsets



Interobserver Agreement

To assess interobserver agreement, the first author collected data on the dependent variables for 100% of probes across all participants and phases. Trial-by-trial agreement averaged 95% (range, 75%-100%) during baseline probes and 100% during posttraining probes for Mark, 97% (range, 83%-100%) during baseline probes and 98% (range, 83%-100%) during posttraining probes for David, and 96% (range, 58%-100%) during baseline probes and 100% during posttraining probes for Jack.

Procedural Integrity

The first author measured procedural integrity from videotapes during a portion of the training sessions for all training conditions and participants. For Mark, scores averaged 98% (range, 90-100%) during 41% of training sessions, 97% (range, 92-100%) during 57% of massed trial training sessions, 100% during 50% of tact training sessions, and 99% (range, 98-100%) during 32% of retention training sessions. For David, scores averaged 97% (range, 91-100%) during 34% of training sessions, 99% (range, 98-100%) during 36% of massed trial training sessions, 100% during 40% of tact training sessions, and 100% during 42% of retention training sessions. For Jack, scores averaged 100% during 40% of training sessions and 38% of retention training sessions. Jack did not receive massed trial or tact training.

Experimental Design

We used a multiple-probe design across behaviors embedded in a concurrent multiple-probe design across participants. We conducted probes in clusters of the three subsets: Trained, Gen1, and Gen2; always in that order. After a stable baseline with Category 1 (shape/number) for Mark, we trained Category 1 with Mark. After an increase in posttraining probes for Category 1 with Mark, we demonstrated a stable baseline, training, and an improvement in posttraining probes with Category 2 (object/color) with Mark and Category 1 with David. Then we did the same with Category 2 with David and Category 1 with Jack. This design allowed us to attempt to demonstrate experimental control across participants (across behaviors, or categories), as well as within participants in case responding generalized from one category to the other. In addition, baseline cross-category probes occurred after the first Category 1 and 2 baseline probes. Posttraining cross-category probes occurred after the posttraining probes of Category 1 and Category 2, as well as intermittently with subsequent posttraining probes. We conducted two consecutive posttraining cross-category probes at the end of David's analysis due to unresolved tact errors. See Table 1 for the general sequence of conditions.

Table 1

General Sequence of Procedural Steps for All Participants

Steps	Procedures	Stimulus Sets
1	Baseline probes	Trained-C1 (Shape/Number) Gen1-C1 Gen2-C1 Trained-C2 (Object/Color) Gen1-C2 Gen2-C2
2	Cross-category probes	Shape/Number/Color Shape/Color Number/Color
3	Training	Trained-C1
4	C1 probes	Trained-C1 Gen1-C1 Gen2-C1
5	Training	Trained-C2
6	Retention training	Trained-C1
7	C2 probes	Trained-C2 Gen1-C2 Gen2-C2
8	Cross-category probes	Shape/Number/Color Shape/Color Number/Color

Instructor Training

The first author implemented behavioral skills training (Parsons et al., 2012) with a Board Certified Behavior Analyst®-level instructor who conducted all sessions. The training procedures were: (a) providing a written protocol describing the experimental steps, (b) vocally describing the procedures, (c) asking the instructor questions about the procedures (e.g., "When would you move the participant from a 2-s delay to a 1-s delay?"), (d) demonstrating the skills, and (e) asking the instructor to practice the steps while providing corrective feedback. Once the instructor showed proficiency in the procedures, she was prepared to start baseline probes.

Procedures

Probes

There were 12 trials in each probe session. There were two questions and three visual stimuli; each question was asked twice with each visual stimulus. The trials were arranged in a quasirandom order (i.e., random, but no more than two identical visual stimuli or questions were presented consecutively). The instructor gained the participant's attention and put a visual stimulus (i.e., picture) on the table in front of the participant. The instructor asked a question in the presence of the picture and provided 5 s to respond. Correct and incorrect responses resulted in the presentation of the next picture and question. After every trial, the instructor presented a mastered instruction (e.g., motor imitation, simple intraverbal, listener task) and delivered praise and a tangible reinforcer contingent on a correct response. Praise and tangible items were identified as reinforcers based on direct observations of clinical sessions.

The cross-category probes were identical to the probe procedures above, with one exception. For shape/number/color, there were 12 trials per session presented in a quasirandom order. The instructor asked the three questions (i.e., "What shape?" "What number?" and "What color?") once for each of the three cards and then conducted three trials with the three questions randomly matched to the three cards. When conducting cross-category probes, the sequence of sessions was shape/number/color, shape/color, and then number/color.

Training

The instructor conducted two to four training sessions per day, three to four times per week. The average session duration was 1.5 min (range, 1-3 min), with at least a 5-min break between sessions. The 12-trial session arrangement was identical to the probe sessions. Training consisted of a progressive time delay based on Miguel and Kobari-Wright (2013; i.e., 0 s, 1 s, 2 s, 3 s, 4 s, no prompt). The delay was increased contingent on 92% (11/12) or higher correct independent or prompted responses at each delay for two consecutive sessions. A prompt regression was moving to a previous time delay (e.g., 2 s to 1 s) contingent on two consecutive sessions at 10/12 (83%) or lower.

In the beginning of the session, the instructor gained the participant's attention, placed a visual stimulus on the table, and asked a question. During the 0-s delay, the instructor delivered an immediate prompt. For example, in the presence of a yellow cat and "What color?" the instructor said, "color yellow." During the 1-s delay, the instructor modeled the response after 1 s of no responding or if an incorrect response occurred. A correct response during training, defined as echoing the keyword and emitting the correct tact (e.g., "color yellow"), resulted in praise on a fixed-ratio (FR) 1 schedule and a tangible/edible reinforcer on a participant-specific schedule (e.g., variable-ratio [VR] 5). The four keywords were "shape," "number," "it is a," and "color." If the participant responded with only the response and not the echoic (e.g., "yellow"), the correction procedure was repeating the question and immediately modeling the complete response; echoing the response resulted in praise but not a tangible/edible reinforcer.

The mastery criterion was 92% (11/12 trials) independent, correct intraverbal-tacts across three consecutive sessions with an overnight criterion. That is, a participant needed to respond at 92% or higher during two sessions on the same day and a third session on the next day, or 92% or higher during one session and during two consecutive sessions the next day. Reaching this mastery criterion resulted in post-training probes in the next session.

Massed Trial Training

Due to acquisition errors, we implemented several supplemental training conditions to strengthen component skills (see Table 2 for the data, errors, and modifications, and Table 3 for the number of sessions per condition). Massed trial training was initiated if a participant (a) experienced three prompt regressions (defined above), (b) emitted errors during retention sessions, or (c) emitted errors during probes as a result of mismatched echoic and tact responses (e.g., "shape two," "color hat") or as a result of incorrect question discrimination (e.g., responding with "it is a flower" to "What color?"). The purpose was to train intraverbal-tacts with prior errors with only one question to increase the stimulus control of the keyword over the tact response. For example, a session was composed of 12 trials of "What color?" with the Category 2 visual stimuli. The instructor used the same progressive time delay, mastery, and regression criteria as in the training condition. Contingent on mastery, the general training condition resumed at the prompt level implemented before massed trial training started.

Tact Training

Tact training occurred if a participant emitted two or more tacting errors for two consecutive sessions during training or post-training probes (e.g., responding "color green" or "green" in the presence of "What color?" and a picture of a yellow cat). For numbers, shapes, and colors, tact training occurred with the same stimuli as in the tact preassessment. For objects, tact training occurred with the nine experimental stimuli (i.e., three Trained, three Gen1, three Gen2). The instructor placed a horizontal array of the stimuli (e.g., six color cards) on the table in the same order each session. Like the tact preassessment, the instructor modeled tacting the first two stimuli (the same ones each session) and then pointed to the next stimulus, and so on. Contingent on correct, independent tacting, the instructor delivered participant-specific reinforcement. Each session consisted of 18 trials. That is, for colors, shapes, and numbers, six cards were presented with three opportunities to independently tact each one. For objects, nine cards corresponding with the Trained, Gen1, and Gen2 subsets were presented with two opportunities to independently tact each one. If a participant made an error, the instructor modeled the response

Table 2

Procedural Modifications

Part.	Condition, Session, and Data	Errors	Modification
Mark	C2 Gen1, Session 12, 75%	"What is it?" → "it is a yellow" "What color?" → "color cat"	Two booster sessions with C2 Trained subset
Mark	C2 Gen2, Session 13, 67%	"What color?" → "color spoon" "What color?" → "color car"	Two booster sessions with C2 Trained subset
Mark	C2 Gen2, Session 16, 58%	"What color?" → "color chicken" "What color?" → "color spoon" "What color?" → "color car"	Massed trial training of "What color?"
Mark	C2 Trained, retention session, 75%	"What color?" → "it is flower"	Massed trial training of "What color?"
Mark	C2 Gen2, Session 19, 50%	"What is it?" (blue hat) → "a chicken"	Tact training of objects
David	C1 Trained, 3 regressions	"What shape?" → "shape number" "What shape?" → "shape two"	Massed trial training of "What shape?"
David	C2 Gen2, Session 16, 50%	"What color?" (purple spoon) → "color blue" "What color?" (purple spoon) → "color green"	Tact training of colors
David	C2 Trained, 3 regressions	"What is it?" → "ih-ah" (unclear)	Articulation training of "it is a"
David	C1 Trained, booster sessions, errors	"What number?" (two) → "number one" "What shape?" (square) → "shape circle"	Tact training of numbers and shapes
Jack	C2 Gen2, Session 22, 75%	"What is it?" (blue hat) → "Santa"	Tact training of objects

Table 3

Number of Sessions per Category in Each Training Condition

	Training		Massed trial		Tact training		Retention	
Participant	C1	C2	C1	C2	C1	C2	C1	C2
Mark	12	15		7 (color)		7 (object)	32	16**
David	18	49	11 (shape)		5 (number) 11* (shape)	9* (color)	45	12
Jack	19				5 (object)		11	

Note. Words in parentheses indicate the focus of the supplemental training. One asterisk indicates the mastery criterion was not reached. Two asterisks indicate that retention sessions included two booster sessions.

and pointed to the next card contingent on a prompted response. The mastery criterion was 100% correct for one session. Subsequent posttraining probe sessions were preceded by a tact training session to ensure retention of the tact responses.

Retention Training

The first posttraining probe or tact training session of the day was preceded by a retention training session to ensure retention of the Trained responses at mastery level. The instructor also conducted retention sessions for the Category 1 and Category 2 Trained subsets prior to administering cross-category probes. Retention sessions were similar to training sessions except the instructor used least-to-most prompting. The instructor presented a question and a visual stimulus and allotted 5 s to respond. A prompt was contingent on an incorrect response or no response within 5 s.

Booster Training (Mark only)

This training was implemented with Mark when trial-by-trial error analyses for Category 2 Gen1 and Gen2 probes indicated correct echoics but incorrect intraverbal-tacts (e.g., responding "it is yellow" when asked, "What is it?" and "color spoon" when asked, "What color?"). The purpose was to increase control of the keywords "it is" and "color" over correct responses. Booster training was similar to retention training and consisted of two Category 2 training sessions. After booster training, the instructor administered the Category 2 and cross-category probes.

Articulation Training (David only)

When David experienced three prompt regressions during Category 2 training, we observed poor articulation of "it is a" (i.e., "ih-ah"). Before initiating massed trial training, we conducted articulation training sessions. These were identical to training sessions, starting at a 0-sec prompt delay; however, when the instructor asked, "What is it?" her model had more exaggerated pronunciation (i.e., elongated pronunciation of the three words, "it is a"). Prompting continued each trial until David matched this articulation. Progression to the next prompt delay was based on clear articulation for three consecutive sessions with the overnight criteria.

Posttraining Probe—New Sequence

We changed the posttraining probe sequence with all participants. The initial posttraining probe sequence consisted of three probes (Trained, Gen1, and Gen2) conducted consecutively with 5-min breaks between probes. In the new sequence, only one probe occurred per day. This decision was made when Mark's Gen1 and Gen2 responding decreased relative to responding in Trained probes on two consecutive days. We hypothesized that putting the newly learned responses on extinction resulted in variability that led to errors in each subsequent probe session. In an attempt to mitigate these extinction effects, the new probe sequence was a probe (Trained, Gen1, or Gen2) on a new day with retention training preceding each probe. In addition, all posttraining cross-category probes were conducted using the new sequence.

Social Validity

The first author presented a survey to parents and staff from the center where the study was conducted with a PowerPoint presentation of the study, as well as graphs and videos of baseline probes, training, and posttraining probes. The survey had 11 questions on a 5-point rating scale (a modified version of the *Behavior Intervention Rating Scale*; Elliott & Treuting, 1991).

Results

Figure 6 depicts the percentage of correct intraverbal-tacts (DV1; dots) and the percentage of correct intraverbal-tacts with correct and incorrect tacts (DV2; bars) for all participants during within-category probes. Figure 7 displays data for cross-category probes.

Baseline Probes

Baseline probes across participants for DV1 and DV2 averaged around 50% (i.e., 47%, 58%, and 52% for Category 1 and 49%, 55%, and 67% for Category 2 for Mark, David, and Jack, respectively). Exceptions were five (out of 19) baseline probes for David as they were above 70% (range, 75% - 83%). For Jack, before Category 1 training (up to Session 19), the mean percentage correct for DV1 in Category 2 was 58% (range, 33%-75%). Error analyses revealed tact errors with objects (e.g., saying "Santa" instead of "hat"). After training in Category 1 and object tact training in Category 2, responding in the last three baseline probes increased to 89% (range, 75%-100%) for DV1 and 95% (range, 92%-100%) for DV2.

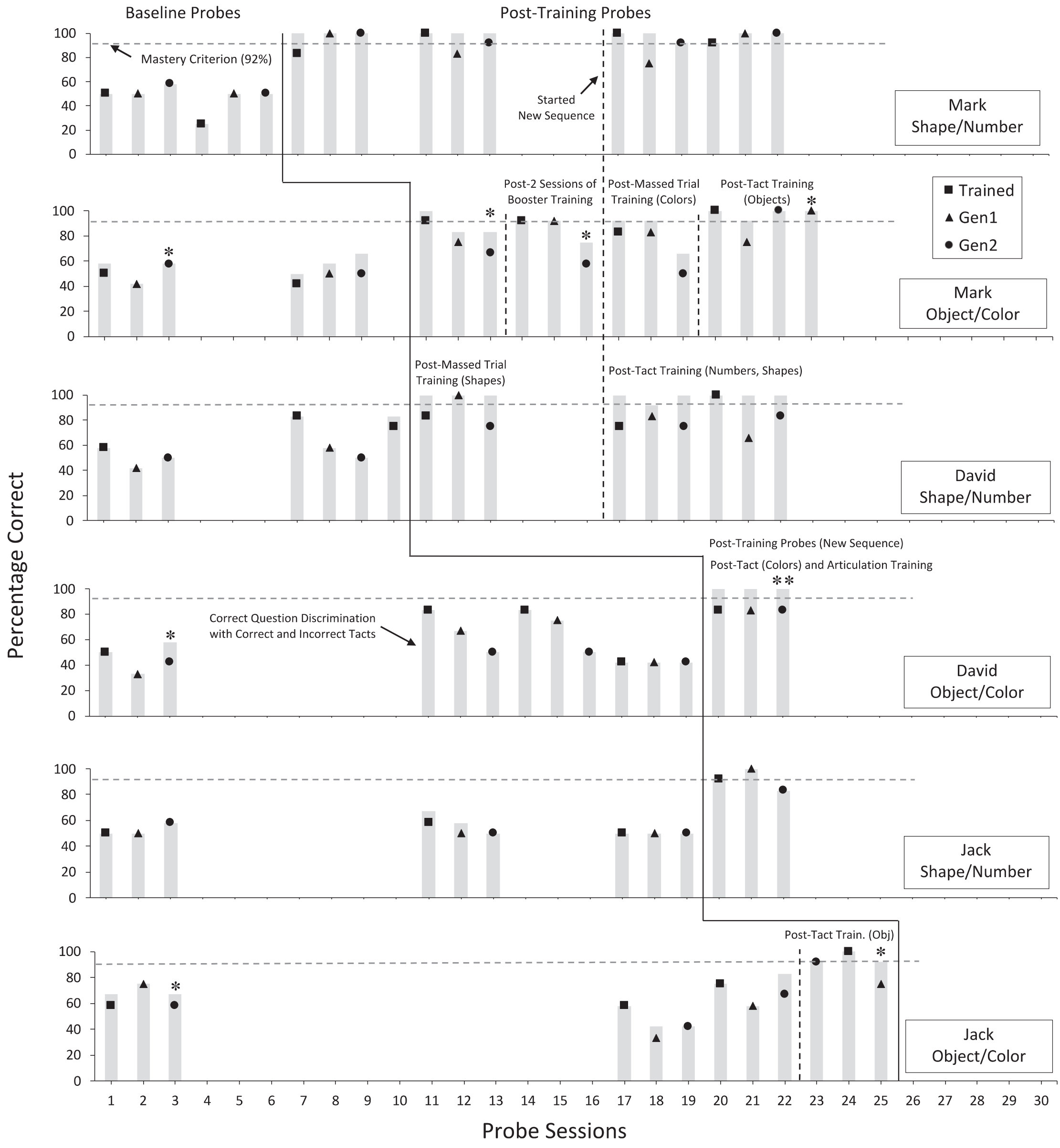
Posttraining Probes

For Category 1 for Mark, DV1 and DV2 were at or above mastery in 9/12 and 12/12 posttraining probes, respectively. For Category 2 for Mark, the posttraining probe average was 82% (range, 50%-100%), compared to a mean of 47% (range, 25%-58%) in baseline. Error analyses of the decreasing trend in the first set of posttraining probes (Sessions 11-13) revealed correct question discrimination but incorrect classes (e.g., responding "color cat" to "What color?" and "it is brown" to "What is it?"). After booster training, the Trained and Gen1 probes were at mastery, and the Gen2 probe was 58%. Error analyses indicated responding with correct question discrimination but incorrect classes mostly to "What color?" (e.g., "color cat"). Levels were slightly reduced in post-massed trial training probes and additional error analyses revealed tact errors (e.g., saying "chicken" when presented with "What is it?" and a card with a blue hat). Post-tact training probes showed increases to 100% across all subsets. Overall, levels of DV2 were similar to DV1 with the exceptions of higher levels of DV2 compared to DV1 in Sessions 7, 12, and 18 in Category 1 and Sessions 17, 18, and 21 in Category 2.

For David, across the nine Category 1 posttraining probes, the average was 82% (range, 67%-100%) for DV1 and 99% (range, 92%-100%) for DV2. During the initial training, David met the criterion for a training modification (i.e., three regressions in his training sequence). Error analyses indicated echoic responding with an incorrect class mostly to "What shape?" (e.g., "shape one"). Therefore, massed trial training of shapes occurred before the first set of posttraining probes. Due to

Figure 6

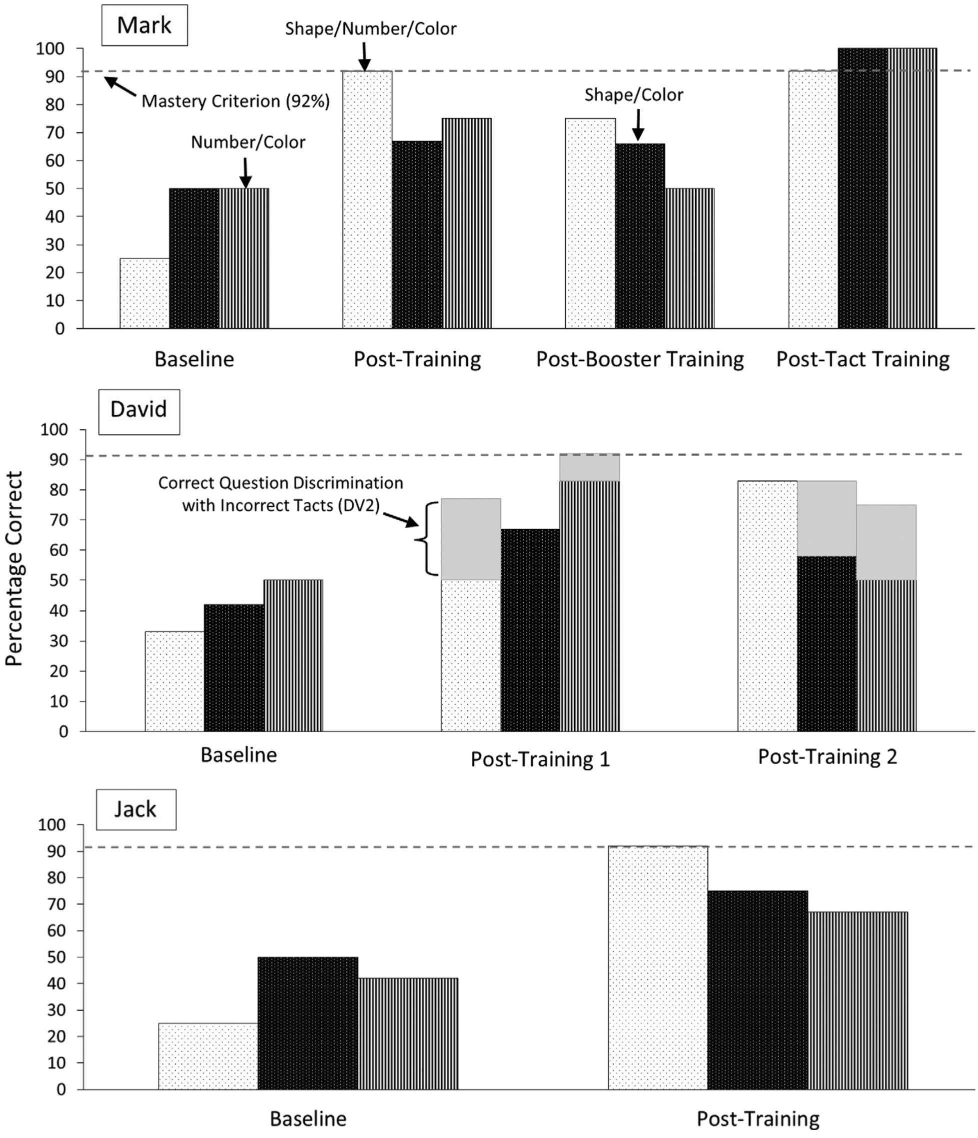
Percentage of Independent, Correct Responses During Probes



Note. Cross-category probes were administered after sessions denoted with an asterisk. Data points represent DV1; grey bars represent DV2.

tacting errors during these probes, we next conducted tact training of shapes and numbers. During tact training, David did not reach the mastery criterion for the trained tacts and demonstrated tacting errors during the last two sets of posttraining probes, represented by DV1

Figure 7
Percentage of Correct Responses During Cross-Category Probes



data below mastery and DV2 data at or above mastery. In Category 2 posttraining probes, David demonstrated 83% correct responding for DV1 and 100% for DV2. Most errors were incorrect tacts of colors.

For Jack, Category 1 posttraining probes averaged 92% (range, 83%-100%) for both DV1 and DV2. There were no posttraining probes for Jack in Category 2.

Cross-Category Probes

Baseline cross-category probes averaged around 40-50% (i.e., 52%, 42%, and 39% for

Mark, David, and Jack, respectively). Final posttraining cross-category probes for DV1, following the additional training for Mark and David, averaged 97%, 64%, and 78%, for Mark, David, and Jack, respectively. For David, DV2 averaged 80%.

Discussion

Progressive time delay with the requirement to echo keywords resulted in the acquisition and generalization of intraverbal-tacts with three children with autism. Specifically, the children learned to provide correct answers when asked, “What number?” and “What shape?” in the presence of different numbers inside shapes, as well as when asked, “What is it?” and “What color?” in the presence of different colored objects. The training procedure required the participants to echo the keyword from the question and emit the tact (e.g., answering “What color?” with “color blue”). Acquisition of these skills was demonstrated in a multiple-probe design across behaviors with two participants and a concurrent multiple-probe design across all three participants. In addition, within- and across-category generalization was demonstrated. These outcomes are important as answering questions about visual stimuli is a developmental milestone (Sundberg, 2008) and central to the expansion of comprehension (Pence Turnbull & Justice, 2012).

This study extends prior research on using echoic responding to facilitate multiply controlled listener responding (Causin et al., 2013), intraverbals (Kisamore et al., 2016), and intraverbal-tacts (degli Espinosa et al., 2020). This study also extends degli Espinosa et al.’s (2020) procedure in three ways. First, we reduced the number of training steps by omitting the initial echoic training phase and prompting and reinforcing echoics in the context of intraverbal-tact training. The reduced number of steps may allow for a more rapid adoption of these procedures in practice. Second, rather than the simple-conditional method of conditional discrimination training, we used the conditional-only method by interspersing trials with two questions (e.g., “What number” and “What shape?”) within each training session as the conditional-only method has been shown to be superior (Grow et al., 2011). Third, because we used a multiple-probe across behaviors design embedded in a concurrent multiple-probe design across participants rather than a multiple-baseline design across behaviors, we were able to detect generalization while retaining experimental control. The merits of these three methodological components across studies should be examined in future research.

There are several potential explanations for the effects of the training procedures, particularly the echoic requirement, on the acquisition and generalization of intraverbal-tact responses. First, echoing the keyword may have “enhanced stimulus control by all relevant controlling stimuli instead of just one of them” (Kisamore et al., 2016, p. 843). In baseline, responding was most often around 50% as the participants tacted one feature of the visual stimulus independent of the question (i.e., restricted stimulus control; Dube et al., 2010). For example, all three participants generally said the number (not shape) in Category 1 when asked both questions. In Category 2, when asked both questions, Mark and David generally said the object (not color), and Jack generally said the color (not object). In training, however, echoing “shape,” for example, may have enhanced the salience of the discriminative verbal stimulus and exerted stimulus control.

Second, the intraverbal control between the keyword (e.g., “color”) and the visual feature (e.g., blue) in an autoclitic frame facilitated stimulus control (degli Espinosa et al., 2020; Palmer, 2007). The participants had a learning history with the words/phrases “color,” “it is a,” “shape,” “number,” and other members of those classes. They would not have had a history, or would have had an aversive history, with word combinations such as "color cat," "it's a blue," "shape three," and "number circle." Emitting the tact in an autoclitic frame may have served to use the category word to "self-prompt" the correct tact (e.g., "color" self-prompted "blue" through an intraverbal relation). Anecdotally, Mark repeatedly stated the keyword before each tact.

Third, there may have been a phenomenon similar to joint control (Causin et al., 2013; Lowenkron, 2006). When asked, "What color?" in the presence of a yellow cat, the participant may have said, "color" as a form of rehearsal, and then scanned for the object and color components of the visual stimulus. When the component matched the product of the echoic (e.g., the yellowness of the cat matched the product of the echoic, "color"), the participant emitted that response (e.g., "yellow"). The relation in the current study is slightly different than the conceptualization by Lowenkron (2006) because the participant tacted an aspect of the visual stimulus with intraverbal control and the product of the echoic. However, the conceptualization is rather close to Lowenkron's original description of joint control, and our analysis may be better considered a specific type of convergent multiple control (Michael et al., 2011).

Finally, although a history of reinforcement can help explain the demonstration of trained conditional discriminations, mediation is required to explain generalized conditional discriminations (Causin et al., 2013; Lowenkron, 2006). For example, the participants had a history of reinforcement for responding with "color yellow" when asked "What color?" in the presence of a yellow cat, but no history of such responding in the presence of a red car (Gen2). Correct responding with "red" when asked "What color?" in the presence of a red car may have been facilitated by applying the strategy of echoing "color" and then emitting "red." Similarly, the participants did not have a history of responding with the cross-category probe stimuli but appeared to apply the strategy of echoing the keyword when presented with those untrained stimuli. Such echoing may be considered a problem-solving strategy that, once learned, may be applied to novel situations (Axe et al., 2019; Miguel, 2018; Palmer, 1991; Skinner, 1957).

Although we did not conduct a comparative or component analysis to determine if the echoic was necessary to facilitate the post-training responses, several findings support the facilitative role of the echoic. First, even though echoing the keyword was not required for correct responses in the probes, within-session analyses indicated that Mark and David emitted the echoic in 100% of the first and last sets of posttraining probes for Categories 1 and 2, as well as in the cross-category probes. Jack echoed in 100% of posttraining probes for Category 1, but in the last set of probes of Category 2, he echoed "color" in 33%-67% of responses and "it is a" in 0% of responses. In most cases, even when echoics were not prompted or reinforced, their presence supports the assertion that echoing the keyword facilitated correct intraverbal-tacts.

Second, after training, acquisition, and generalization in Category 1, Jack acquired Category 2 intraverbal-tacts without training. A potential explanation for this outcome is similar to the explanation provided above in which echoing the keyword may have generalized from Category 1 to Category 2 stimuli (Causin et al., 2013). Even though Jack echoed the keyword in few responses in Category 2 (see data above), this does not rule out covert echoing.

Third, Mark and David met criteria for the massed trial training condition. Mark's errors were mismatches of keywords and tacts (e.g., "color hat"). David met three regression criteria and therefore, did not proceed through the planned prompting hierarchy. The goal of massed trial training was to isolate a keyword to increase stimulus control of visual features (e.g., training "color X" in response to massed trials of "What color?"). This procedure may have also strengthened keyword-member-of-class (our own term) intraverbal responses (e.g., "color" evoking a color). After reaching mastery during massed trial training and restarting the general training, Mark and David reached mastery in three and four sessions, respectively. An alternative route would be starting with building a history of intraverbal relations through echoic training (i.e., degli Espinosa et al., 2020), and future research should explore relevant component analyses. Additionally, the error pattern for David in Category 2 was unclear pronunciation of "it is a"; he said "ih-ah." Horne and Lowe (1996) stressed the importance of clear pronunciation in the acquisition of naming and we hypothesized that clear pronunciation of keywords would strengthen keyword-member-of-class intraverbal control. The articulation training led to David mastering the trained set without additional modifications.

The current results may have implications for the sequence of training tacts, intraverbals, and intraverbal-tacts. Mark and David required, on average, twice as many training sessions with Category 2 (object/color) compared to Category 1 (shape/number). This may be due to the long history of reinforcement for tacting objects and colors many children with autism experience in early intensive behavioral intervention (Taylor & McDonough, 1996). Some curricula recommend teaching tacts of objects with "What is it?" and later, teaching tacts of colors by showing colored index cards and asking, "What color?" (e.g., Sundberg, 2008). Such curricular sequences appear to use simple-conditional discrimination training with respect to the question–visual stimulus conditional discriminations, which has been shown to render the question a neutral, as opposed to a conditional, stimulus (Grow et al., 2011). In other words, this sequence may produce overselectivity of the visual stimulus (Dube et al., 2016). Category 2 (shape/number) may have been acquired more quickly by Mark and David than Category 1 due to a shorter history of reinforcement of tacts of shapes and numbers under the control of the questions "What shape?" and "What number?", which appear in language curricula later than tacts of objects and colors (Sundberg, 2008). Sundberg suggested introducing "What is it?" and "What color?" as the first two questions for intraverbal-tacts (Intraverbal M-10). When introducing intraverbal-tacts, practitioners may avoid producing overselectivity by first targeting questions other than "What is it?" and "What color?" such as "What shape?" "What number?" "What does it say?" and "What does it have?"

Additionally, the results were socially valid. On the 5-point scale (1 = Strongly Disagree, 5 = Strongly Agree), the mean rating from parents and staff was 4.6 (range, 4.2-5.0). Some of the highest ratings were for the following items: "It is important to teach children with autism to answer questions about objects and pictures," "The intervention quickly improved the children's question answering," and "This procedure is a valuable contribution to behavior-analytic work." Two of the lowest ratings were to "Answering using a full sentence is an appropriate strategy" and "Most teachers would find this intervention suitable for the language challenge described." Although efficacious, future research should explore ways of ensuring the procedures are even more socially valid (e.g., inviting parents or teaching staff to collaborate on selecting the targets).

There were several limitations of the current study. First, some of the trained responses were grammatically uncommon. The four trained responses were (a) "What is it?" → "it is a X;" (b) "What color?" → "color X;" (c) "What shape?" → "shape X;" and (d) "What number?" → "number X." Responses (a) and (d) are grammatically common, whereas responses (b) and (c) are grammatically uncommon.

Social validity respondents expressed concerns about the children saying, "shape" and "number" as part of answers. Nevertheless, the procedure was effective and echoing the keywords may have facilitated correct responding. When selecting stimuli, it was important for physical features of the visual stimuli to relate to all questions. For example, it is more grammatically common to show animals and teach "What is it?" → "it is a cow" and "What does it say?" → "it says moo." However, "moo" does not have a physical form, whereas the color of the cow does. In practice, and as a child acquires more intraverbal-tacts, practitioners may teach questions with more common grammatical forms. Additionally, practitioners may extend responses, such as "the color is X" and "the shape is X," by using textual or audio prompts or scripts. Although future researchers may assess fading responses from "What color?" → "color blue" to just "blue," this fading may happen naturally.

Second, all participants emitted tacting errors despite preassessment results indicating 100% correct tact responding with all visual components of the experimental stimuli. Because David exhibited the most tacting errors (e.g., tacting yellow as green, 3 as 8, square as triangle), we arranged separate tact training sessions for numbers, shapes, and colors. During preassessments, we tested tact components using stimuli that were different than the stimuli used in the study (see Figure 2) to not overexpose the participants to the experimental stimuli. For example, we preassessed hat using a yellow hat (see Figure 2), and when Mark was asked, "What is it?" in the presence of the blue hat in the Gen 2 probe (see Figure 2), he said, "chicken" (see Table 2). Future researchers should consider testing tact components with the experimental stimuli to mitigate these tacting errors. Third, we did not counterbalance the order of the categories across participants. Counterbalancing would have demonstrated stronger experimental control if some participants experienced training and posttraining probes with object/color first and shape/color second.

Fourth, responses in four of David's baseline Trained probes were high (75%-83%) which threatens internal validity. In baseline (e.g., Category 1 [shape/number]), David answered "What number?" with the correct number, which was 50% of the trials. For "What shape?" he may have guessed (i.e., chance responding), which would be 50% of 50%, or 25%. Adding those together (50% + 25%) results in 75% correct responding. This limitation is tempered as baseline Trained responding in Category 2 dropped to near 50% in the last baseline probe. There is a substantial change in the level of correct responding between baseline and posttraining probes with DV2 (negating the tacting errors), and there is a substantial change in the level of correct responding between baseline and posttraining generalization probes for DV1. Finally, condition changes across participants were inconsistent due to individualized error analyses and modifications. More consistency would allow stronger conclusions regarding the conditions affecting the intraverbal-tact repertoire.

In addition to the suggestions above, there are numerous avenues for future research. First, future researchers should isolate requiring echoics as an independent variable by evaluating a prompt fading procedure with and without echoic training or requiring an echoic of an irrelevant word (Charlop, 1983). This may provide stronger support for the value of the echoic training and the mechanism involved. Second, researchers could target selection responses using the same antecedents as a prerequisite training to establish strict conditional discrimination. For example, when asked, "What number?" the child points to the number, and when asked, "What shape?" the child points to the shape. However, this may not be possible in all cases (e.g., a colored object).

Third, to assess the role of joint control in producing the responses, researchers may block the echoic of keywords by requiring a random word or phrase prior to responding (Clough et al., 2016). An example would be presenting a yellow cat, asking, "What color?" and requiring the response, "A, B, C, D, Yellow." Measuring latencies to responding in the echoic and blocking conditions could help elucidate mechanisms. Finally, future researchers should extend the categories beyond objects, colors, shapes, and numbers.

In conclusion, few studies have explored multiply controlled verbal behavior with children with autism. Echoics may be responsible for mediating the acquisition and generalization of intraverbal-tacts. Whereas the literature is replete with analyses of verbal behavior under simple stimulus control, future researchers should continue examining analyses and procedures for teaching generalized, multiply controlled verbal behavior.

REFERENCES

Axe, J. B. (2008). Conditional discrimination in the intraverbal relation: A review and recommendations for future research. *The Analysis of Verbal Behavior*, *24*(1), 159-174. https://doi.org/10.1007/bf03393064

Axe, J. B., Phelan, S. H., & Irwin, C. L. (2019). Empirical evaluations of Skinner's analysis of problem solving. *The Analysis of Verbal Behavior*, *35*(1), 39-56. https://doi.org/10.1007/s40616-018-0103-4

Bondy, A., Tincani, M., & Frost, L. (2004). Multiply controlled verbal operants: An analysis and extension to the Picture Exchange Communication System. *The Behavior Analyst*, *27*(2), 247-261. https://doi.org/10.1007/bf03393184

Causin, K. G., Albert, K. M., Carbone, V. J., & Sweeney-Kerwin, E. J. (2013). The role of joint control in teaching listener responding to children with autism and other developmental disabilities. *Research in Autism Spectrum Disorders*, *7*(9), 997-1011. https://doi.org/10.1016/j.rasd.2013.04.011

Charlop, M. H. (1983). The effects of echolalia on acquisition and generalization of receptive labeling in autistic children. *Journal of Applied Behavior Analysis*, *16* (1), 111-126. https://doi.org/10.1901/jaba.1983.16-111

Clough, C. W., Meyer, C. S., & Miguel, C. F. (2016). The effects of blocking and joint control training on sequencing visual stimuli. *The Analysis of Verbal Behavior*, *32*(2), 242-264. https://doi.org/10.1007/s40616-016-0067-1

degli Espinosa, F., Gerosa, F., & Brocchin-Swales, V. (2020). Teaching multiply controlled tacting to children with autism. *European Journal of Behavior Analysis*. Advance online publication. https://doi.org/10.1080/15021149.2020.1737407

DeSouza, A. A., Akers, J. S., & Fisher, W. W. (2017). Empirical application of Skinner's verbal behavior to interventions for children with autism: A review. *The Analysis of Verbal Behavior*, *33*(2), 229-259. https://doi.org/10.1007/s40616-017-0093-7

DeSouza, A. A., Fisher, W. W., & Rodriguez, N. M. (2019). Facilitating the emergence of convergent intraverbals in children with autism. *Journal of Applied Behavior Analysis*, *52*(1), 28-49. https://doi.org/10.1002/jaba.520

Devine, B., Carp, C. L., Hiett, K. A., & Petursdottir, A. I. (2016). Emergence of intraverbal responding following tact instruction with compound stimuli. *The Analysis of Verbal Behavior*, *32*(2), 154-170. https://doi.org/10.1007/s40616-016-0062-6

Dube, W. V., Dickson, C. A., Balsamo, L. M., O'Donnell, K. L., Tomanari, G. Y., Farren, K. M., Wheeler, E. E., & McIlvane, W. J. (2010). Observing behavior and atypically restricted stimulus control. *Journal of the Experimental Analysis of Behavior*, *94*(3), 297-313. https://doi.org/10.1901/jeab.2010.94-297

Dube, W. V., Farber, R. S., Mueller, M. R., Grant, E., Lorin, L., & Deutsch, C. K. (2016). Stimulus overselectivity in autism, Down syndrome, and typical development. *American Journal on Intellectual and Developmental Disabilities*, *121*(3), 219-235. https://doi.org/10.1352/1944-7558-121.3.219

Eikeseth, S., & Smith, D. P. (2013). An analysis of verbal stimulus control in intraverbal behavior: Implications for practice and applied research. *The Analysis of Verbal Behavior*, *29*(1), 125-135. https://doi.org/10.1007/bf03393130

Elliott, S. N., & Treuting, M. V. (1991). The behavior intervention rating scale: Development and validation of a pretreatment acceptability and effectiveness measure. *Journal of School Psychology*, *29*(1), 43-51. https://doi.org/10.1016/0022-4405(91)90014-I

Farber, R. S., Dickson, C. A., & Dube, W. V. (2017). Reducing overselective stimulus control with differential observing responses. *Journal of Applied Behavior Analysis*, *50*(1), 87-105. https://doi.org/10.1002/jaba.363

Fisher, W. W., Retzlaff, B. J., Akers, J. S., DeSouza, A. A., Kaminski, A. J., & Machado, M. A. (2019). Establishing initial auditory-visual conditional discriminations and emergence of initial tacts in young children with autism spectrum disorder.

Journal of Applied Behavior Analysis, *52*(4), 1089-1106. https://doi.org/10.1002/jaba.586

Floor, P., & Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy*, *9*(3), 327-339. https://doi.org/10.1207/s15327078in0903_4

Glodowski, K. R., & Rodriguez, N. M. (2019). The effects of scenic picture prompts on variability during the acquisition of intraverbal categorization for children with autism. *The Analysis of Verbal Behavior*, *35* (2), 134-148. https://doi.org/10.1007/s40616-019-00120-2

Grow, L. L., Carr, J. E., Kodak, T. M., Jostad, C. M., & Kisamore, A. N. (2011). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders. *Journal of Applied Behavior Analysis*, *44*(3), 475-498. https://doi.org/10.1901/jaba.2011.44-475

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*(1), 185-241. https://doi.org/10.1901/jeab.1996.65-185

Ingvarsson, E. T., Kramer, R. L., Carp, C. H., Pétursdóttir, A. I., & Macias, H. (2016). Evaluation of a blocked-trials procedure to establish complex stimulus control over intraverbal responses in children with autism. *The Analysis of Verbal Behavior*, *32* (2), 205-224. https://doi.org/10.1007/s40616-016-0071-5

Kisamore, A. N., Karsten, A. M., & Mann, C. C. (2016). Teaching multiply controlled intraverbals to children and adolescents with autism spectrum disorder. *Journal of Applied Behavior Analysis*, *49*(4), 826-847. https://doi.org/10.1002/jaba.344

Kisamore, A. N., Karsten, A. M., Mann, C. C., & Conde, K. A. (2013). Effects of a differential observing response on intraverbal performance of preschool children: A preliminary investigation. *The Analysis of Verbal Behavior*, *29*(1), 101-108. https://doi.org/10.1007/bf03393127

Lowenkron, B. (2006). An introduction to joint control. *The Analysis of Verbal Behavior*, *22*(1), 123-127. https://doi.org/10.1007/bf03393034

Michael, J., Palmer, D. C., & Sundberg, M. L. (2011). The multiple control of verbal behavior. *The Analysis of Verbal Behavior*, *27*(1), 3-22. https://doi.org/10.1007/bf03393089

Miguel, C. F. (2016). Common and intraverbal bidirectional naming. *The Analysis of Verbal Behavior*, *32*(2), 125-138. https://doi.org/10.1007/s40616-016-0066-2

Miguel, C. F. (2018). Problem-solving, bidirectional naming, and the development of verbal repertoires. *Behavior Analysis: Research and Practice*, *18*(4), 340-353. https://doi.org/10.1037/bar0000110

Miguel, C. F., & Kobari-Wright, V. V. (2013). The effects of tact training on the emergence of categorization and listener behavior in children with autism. *Journal of Applied Behavior Analysis*, *46*(3), 669-673. https://doi.org/10.1002/jaba.62

Miguel, C. F., Petursdóttir, A. I., & Carr, J. E. (2005). The effects of multiple-tact and receptive discrimination training on the acquisition of intraverbal behavior. *The Analysis of Verbal Behavior*, *21*(1), 27-41. https://doi.org/10.1007/BF03393008

Palmer, D. C. (1991). A behavioral interpretation of memory. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 261-279). Context Press.

Palmer, D. C. (2007). Verbal behavior: What is the function of structure? *European Journal of Behavior Analysis*, *8*(2), 161-175. https://doi.org/10.1080/15021149.2007.11434280

Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-based staff training: A guide for practitioners. *Behavior Analysis in Practice*, *5*(2), 2-11.

Pence Turnbull, K. L. & Justice, L. M. (2012). *Language development from theory to practice*. Pearson Education, Inc.

Petursdóttir, A., Lepper, T., & Peterson, S. (2014). Effects of collateral response requirements and exemplar training on listener training outcomes in children. *Psychological Record*, *64*(4), 703-717. https://doi.org/10.1007/s40732-014-0051-x

Skinner, B. F. (1957). *Verbal behavior*. Appleton-Century-Crofts.

Stauch, T., LaLonde, K., Plavnick, J. B., Savana Bak, M. Y., & Gatewood, K. (2017). Intraverbal training for individuals with autism: The current status of multiple control. *The Analysis of Verbal Behavior*, *33*(1), 98-116. https://doi.org/10.1007/s40616-017-0079-5

Sundberg, M. L. (2008). *VB-MAPP: Verbal behavior milestones assessment and placement program*. AVB Press.

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*(1), 23-43. https://doi.org/10.1007/BF03393090

Tager-Flusberg, H., Edelson, L., & Luyster, R. (2011). Language and communication in autism spectrum disorders. In D. Amaral, G. Dawson, & D. Geschwind (Eds.), *Autism spectrum disorders* (pp. 172-185). Oxford University.

Taylor, B. A., & McDonough, K. A. (1996). Selecting teaching programs. In C. Maurice, G. Green, & S. Luce (Eds.), *Behavioral intervention for young children with autism: A manual for parents and professionals* (pp. 63-177). Pro-Ed.

Received March 26, 2020
Final acceptance September 9, 2020
Action Editor, Tyra Sellers