JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2011, **95,** 237–262 NUMBER 2 (MARCH)

THE ROLE OF COMMON MOTOR RESPONSES IN STIMULUS CATEGORIZATION BY PRESCHOOL CHILDREN

AMANDA M. MAHONEY[1], CAIO F. MIGUEL[2], WILLIAM H. AHEARN[1], AND JULIANNE BELL[3]

[1]NEW ENGLAND CENTER FOR CHILDREN AND NORTHEASTERN UNIVERSITY
[2]CALIFORNIA STATE UNIVERSITY, SACRAMENTO
[3]UNIVERSITY OF ULSTER, COLERAINE

The purpose of this study was to assess the role of common motor responses as the ''speaker'' behavior on stimulus class formation, and the emergence of functional classes. Experiment 1 examined whether training one motor response to a set of three stimuli and a second motor response to another set of three stimuli would result in correct category-sort responses for 5 typically developing preschool children. Three of the children passed the categorization tests. Experiment 2 examined whether the classes formed in Experiment 1 were functional classes, and whether participants who did not pass categorization tests in Experiment 1 would do so following common vocal tact training. The 2 participants who failed categorization tests in Experiment 1 passed these tests in Experiment 2, although none of the participants passed the tests for functional classes. The results of the current study did not unequivocally support the naming hypothesis. Future research should therefore evaluate other possible sources of control that aid in stimulus categorization.

Key words: naming, equivalence, class formation, functional classes, categorization, humans

---

According to Horne and Lowe (1996), *naming* an object includes not only emitting the appropriate tact in its presence (speaker behavior), but also emitting a previously conditioned response in the presence of its spoken, written, or signed name (listener behavior). These authors suggest that three separately acquired repertoires (listening, echoing, and tacting) interact to produce more complex forms of verbal behavior. Listener behavior encompasses all responses that occur in the presence of verbal stimuli and, according to Skinner (1957, p. 225), are specifically conditioned to reinforce the behavior of the speaker[1]. In Horne and Lowe's example of how listener behavior can be taught, selection responses are occasioned by instructions (e.g., ''find the ball'') and are modeled (e.g., pointing to the ball) and reinforced by the caregiver (e.g., ''good girl!''). Through this process the child learns to differentially respond to the auditory stimuli produced by the caregiver. Later the child is taught to emit a variety of listener responses that encompass even more specific instructions (pick up, throw away, put, bounce, roll, etc.).

The echoic repertoire is established when the child repeats utterances emitted by caregivers. Over time they come to emit vocalizations that sound similar to those produced by their caregivers. In these instances the reinforcement may be either mediated or occur automatically (Miguel, Carr, & Michael, 2002; Vaughan & Michael, 1982) when the child's vocal production matches that of the model (see Horne and Lowe, 1996, for a more in-depth description of these processes). The echoic and listener repertoires are then used in teaching the tact. Now, in the presence of the nonverbal stimulus (e.g., a ball), the caregiver provides auditory stimulation (e.g., ''ball'') and the model (e.g., points to the ball) that occasions a verbal response on the part of the child (e.g., ''ball''). After repeated trials, the child may now emit the same verbal response in the absence of the model (e.g., will say ''ball'' in the presence of a ball without the echoic prompt). Once a child has been taught to respond as a speaker and a listener in

---

This study is based on a thesis submitted by the first author to the Bouvé College of Health Sciences at Northeastern University in partial fulfillment of the requirements for an M.S. degree in Applied Behavior Analysis. We thank Ellen Graham, Bethany McNamara, and the staff at The New England Center for Children in Southborough, MA for their invaluable on-site support. We also thank Dan Knight, Matthew Jameson, Angela Irwin, and Catherine Yates for their assistance with data collection.

Address correspondence to Caio Miguel, Department of Psychology, California State University, Sacramento, CA 95819-6007 (e-mail: miguelc@csus.edu).

doi: 10.1901/jeab.2011.95-237

[1] Schlinger (2008) modified this definition and suggested that listening is actually subvocal verbal behavior.

the presence of some stimuli, these repertoires become part of an interconnected relation that allows for the emergence of one after training only the other. Additionally, novel stimuli may become related by training common names.

There are two important outcomes of training common names: stimulus substitutability and transfer of function (Lowe, Horne, Harris, & Randle, 2002). Stimulus substitutability occurs when any stimulus in an arbitrary class occasions selection of other members. Transfer-of-function occurs when responses are occasioned by other stimuli in the class without direct training. Transfer-of-function is of interest because, once a response has been learned in the presence of one stimulus, stimuli that share that same name will occasion this same response. Goldiamond (1962) stated that this phenomenon is a necessary requirement to define a functional stimulus class. In other words, a functional stimulus class is said to exist when, following training of a novel response to only one or a few members of the class, the remaining members of the class will evoke the same response without explicit training. Horne and Lowe (1996) used the example of a child picking up their toys. After a child learns to tact several toys, the parent may hold up a single toy and state, ''pick up the others'' and the child will pick up the remaining toys. Additionally, after learning that one toy goes in the toy box, the child will also put the other toys in the toy box without being directly trained to do so. Functional classes have been defined as a set of stimuli that share a common function (Dougher & Markham, 1996). Though still used, this definition may not accurately describe the property of interest (function transfer). To clarify, multiple stimuli may all come to exert stimulus control over a single response, but if just one stimulus in the class becomes part of a novel relation, the other stimuli may do so as well without direct training. How these relations among stimuli and behavior are established in the absence of direct reinforcement is the cause of much theoretical dispute (Hineline, 1997). Horne and Lowe's account predicts that, once the naming relation is acquired, learning common names is sufficient to produce this emergent behavior. This has been supported by several recent studies (e.g., Horne, Hughes, & Lowe, 2006; Horne, Lowe, & Harris, 2007; Horne, Lowe, & Randle, 2004; Lowe et al., 2002; Lowe, Horne, & Hughes, 2005; Miguel, Pettursdottir, Carr, & Michael, 2008).

The procedure most commonly used to assess stimulus class formation as a function of naming consists of a categorization task, not too dissimilar from the typical matching-to-sample task used in equivalence research (Sidman, 1994). This procedure was first used by Lowe et al. (2002) and includes presenting two or more sets (classes) of arbitrary stimuli that cannot be related to one another based on physical similarity. A sample stimulus is then withdrawn from the array and the participants are asked to select the remaining stimuli that belong to the same set (i.e., the stimuli that occasion the same tact). A correct category-sort response is recorded when the participants select all of the stimuli from the array that belong in the same set as the sample. If the participants do not sort accurately, a second test is conducted. This second test is identical to the first except that the experimenter instructs the participants to tact the sample prior to selecting the remaining stimuli. These tests have been used to assess whether common tact or listener behavior facilitates categorization; in other words, whether correct sorting behavior is more likely after the participants have learned to tact and/or respond as a listener to the target stimuli.

Recently, Miguel et al. (2008) used this categorization procedure to further evaluate the role of speaker and listener repertoires in the development of stimulus classes. Six pictures of outlines of U.S. state maps were used in this study; three stimuli were trained as ''North,'' and three stimuli were trained as ''South.'' The training procedures and categorization tests were similar to those used by Lowe et al. (2002) and Horne et al. (2004). Half of the participants were taught a common tact (either ''North'' or ''South'') to each stimulus belonging to one of the two defined classes (i.e., speaker behavior) and the other half were trained to select the stimuli when the experimenter stated either ''North'' or ''South'' (i.e., listener behavior). During tact training, each stimulus was presented individually and the experimenter asked, ''What is it?'' During listener training, three stimuli were placed before the participants (two trained stimuli and one untrained stimulus).

The experimenter stated, ''Give me North.'' Categorization tests were conducted as described above. Of the 4 participants who had undergone tact training, 2 categorized correctly following the initial tact training while 2 other participants only categorized after additional tact training. During the additional tact training, all stimuli were present on the table and the experimenter stated, ''This is North. What is it?'' Following listener training, participants were tested on categorization as well as on their ability to tact the maps. One participant categorized without first tacting the sample stimulus while the other 2 participants categorized when the tact of the sample was required prior to selection. When correct categorization did not occur, the participants also failed to tact correctly. For this reason, the authors suggested that participants' naming repertoire may have played an important role in the formation of stimulus classes, although the results could also be argued as equivocal.

These classes of topographically distinct stimuli may also become part of the same stimulus class due to a common function. Lowe et al. (2005), for example, trained a motor response to one stimulus in a previously established arbitrary stimulus class of three members. Then, tests were conducted in which the remaining stimuli in each set were presented to determine whether the participants would emit the motor response corresponding to that stimulus set. These tests were conducted to assess whether the formed stimulus classes would emerge as functional classes without direct training. Results showed that the motor responses did transfer to the remaining stimuli.

To summarize, past naming research has suggested that each component of the naming relation (i.e., listener and speaker) can be taught separately and that naming is correlated with stimulus class formation (e.g., Horne et al., 2004; Lowe et al., 2002, Miguel et al., 2008). Lowe et al. (2005) extended these results by demonstrating the emergence of a functional class following training of a common nonvocal response to members of stimulus classes that had previously been established. These results add to our understanding of class formation by demonstrating that, following the specific training procedures used, the remaining stimuli will now evoke these novel responses and explicit reinforcement is unnecessary.

One interpretation of these outcomes of tact training is that the initially-trained vocal response is mediating transfer of the motor response. Past research, however, suggests that nonverbal humans who do not demonstrate the naming relation have acquired similar emergent relations (e.g. Carr, Wilkinson, Blackman, & McIlvane, 2000). Further, it has been suggested that untrained differential responding to the samples may not be linguistic (Sidman, 1997, 2000) and thus it would be incorrect to assume that participants are tacting the samples. Sidman (2000) asserts that performances during conditional discrimination tests are simply an outcome of the reinforcement contingency. In other words, equivalence is of phylogenic origin. According to Sidman, a descriptive account of the observed performances is all that is necessary to explain this phenomenon. Naming research, on the other hand, has formulated a behavioral account of the learning processes that may give rise to the observed untrained relations. In this account, one learns to respond to his or her own verbal behavior as a listener. During categorization tests, the correct comparisons are selected in the presence of the sample because all of the stimuli are related by common names. Though these processes may not be observed (e.g. covert tacts that potentially mediate transfer of functional responses) this explanation will persist if proven useful. Hayes, Barnes-Holmes, and Roche (2001) have constructed a theory that is a bit of an amalgam of Skinner's verbal behavior and equivalence research. This theory, termed relational frame theory, maintains that the function of the stimulus does not simply transfer among stimuli, but instead is transformed in accordance with the relation between the stimuli. In other words, if the stimuli in a class have evoked responses that have contacted consequences under the frame, or context, of ''opposite,'' for example, then a response that is reinforced in the presence of one stimulus may actually become further suppressed in the presence of a second stimulus. These authors assert that the relations that emerge among stimuli without direct reinforcement is what is meant by the phrase ''verbal'' and humans only acquire a verbal repertoire as these relations are learned.

A better understanding of whether the performances observed during categorization tests are ''linguistic'' might be achieved by subjecting nonvocal responses to the same tests. The terms linguistic or verbal do not imply ''vocal'' because signing and other nonvocal modes of communication may fall under this term (Skinner, 1957). However, differential outcomes when a vocal response is learned versus a nonvocal response may provide further indication of whether the performance observed during categorization tests is due to covert mediation by vocal–verbal behavior. Subjecting nonverbal organisms to these tests would also provide similar information. Horne et al. (2007) explored this issue by teaching common manual signs first. In that study, 8 typically developing participants, ages 2–4 years, were taught to engage in a single response in the presence of each member of one set of stimuli and a different response in the presence of each member of a second set of stimuli. In the presence of each stimulus of one set, the participants were trained to hold both fists in front of their body with the right fist on top of the left fist. In the presence of each stimulus in the other set, the participants were trained to place the fingertips of each hand onto their shoulders with elbows extended sideways. Training was also conducted to establish listener behavior, which consisted of selecting the stimulus in the presence of the manual sign modeled by the experimenter. Categorization tests (as described above) were conducted and all participants categorized correctly by selecting the comparison stimuli that had been trained to the same motor response as the sample. Two participants were then taught to emit vocal responses to one stimulus in each set (e.g., ''vek'' to a stimulus in Set 1 and ''zag'' to a stimulus in Set 2). These tacts then transferred to the remaining stimuli in each set without direct training.

The current article attempts to address some aspects of the Horne et al. (2007) study. Their results suggest that training a manual sign to each stimulus may be as effective as training a vocal response. Three components of their methods may lessen the strength of this conclusion. First, because the motor responses used were so topographically distinct and easy to tact, it is unclear whether the participants were covertly tacting the stimulus based on some property of the response-produced stimuli (e.g., the participant may have self-tacted, ''fist'' in the presence of one set of stimuli and ''shoulder'' in the presence of the other set of stimuli since those are salient body parts in use during the response). The authors stated as one of their main conclusions that, ''In line with Horne and Lowe's (1996) naming account, manual sign naming was found to be as effective as vocal naming in establishing arbitrary stimulus categorization'' (Horne et al., 2007, p. 367). But if motor responses are occasioning covert vocal naming there is essentially no difference between these two training procedures. This study attempted to decrease the likelihood of covert verbal behavior by using two motor responses that are topographically similar and thus difficult to differentially tact.

Another possible limitation of Horne et al. (2007) is related to the use of hand signs as the motor response. Of interest in these studies is whether mediational responding is necessary and, if so, of what sort. In the case of hand signs, participants may emit the response at any time such that the response mediates between stimulus presentation and the selection response. When the response occurs under these conditions, it is functionally no different than a vocal response. Past research has in fact demonstrated that, like vocal responses, hand signs can be evoked by one stimulus (the sample) as well as by a stimulus occasioning the selection response (the comparison) during selection-based tests, thus serving as mediational responses (e.g., Lowenkron, 1988). The specific motor response selected for the current study should serve to decrease the possibility that participants would engage in mediational responding. This will allow for speculation regarding the necessity of a mediational response at all (Sidman, 1994).

Third, the stimuli were trained in pairs initially, but training phases were also conducted with all of the stimuli on the table (see Horne et al., 2007, for a more thorough description). After the first two pairs of stimuli were trained, training was conducted with all four stimuli present on the table. After a review and additional training of the third pair, all six stimuli were present on the desk during training. This procedure could have inadvertently reinforced either overt or covert tacting of the stimuli in Set 1 and those in Set 2 in succession. For example, participants could have emitted the motor response and simultaneously looked at all three stimuli that had previously evoked that response. Also, this sort of arrangement could facilitate generalization because of the similarity in presentation between training and testing conditions (Miguel et al, 2008). If this occurred, categorization responses may have been reinforced and would not constitute an emergent performance. For this reason, remedial tact phases were not included in the motor response or tact training phases of the current study.

In short, past research has found that vocal responses are efficient in developing complex stimulus classes in young children (e.g. Horne et al, 2006; Miguel et al., 2008) and that motor responses alone have generated these classes as well (Horne et al., 2007). Though it is impossible to determine what participants are doing covertly, training procedures that reduce the likelihood that participants can covertly tact the stimuli may provide evidence that categorization is not a function of these mediating responses. Improvements in performance following training of a vocal response may provide further information about what the participants are actually doing during tests when a vocal response is available (i.e., covert vocal responses may be mediating emergent relations during test trials).

The purpose of the current studies was to evaluate the effects of training difficult-to-tact motor responses as the ''speaker'' relation on the emergence of categorization, listener relation, and the formation of functional classes. In Experiment 1, 5 preschool-age participants were taught to discriminate between two 3-member classes by a common motor response. Categorization tests were conducted following mastery of all six stimuli. Specifically, the purpose of Experiment 1 was to assess whether training one motor response to a set of three stimuli and a second motor response to another set of three stimuli (in a pair-wise training) would result in categorization. The purpose of Experiment 2 was to investigate whether the classes formed in Experiment 1 were functional classes, and to assess whether participants who did not pass categorization tests in Experiment 1 would pass them following vocal tact training.

EXPERIMENT 1

Method

Participants, Setting, and Materials

Five typically-developing children participated in the study; Kim and Meghan were 4 years old while Stephen, Andrew, and Gavin were 5 years old. Sessions were conducted daily (at approximately the same time) in a room (198 cm by 304 cm) located across the hall from the preschool classroom. The room contained a table, two chairs, and the materials necessary for that session (see below). Duration of sessions varied because conditions required varying tasks, but did not exceed 15 min.

Materials included six pieces of paper (20 cm by 14 cm) with line drawings to which motor responses were trained and which were sorted during tests for categorization. Additionally, two distracter stimuli were used during Listener tests. Prior to selecting stimuli for Set A and Set B, the stimuli were presented to two adults who were naïve to the purpose of the study. The adults were asked to separate the stimuli into two groups based on physical similarity. Set A and Set B were then formed from six of the stimuli; stimuli that were categorized together by the adults were not put in the same set. Two stimuli containing similar properties to training stimuli (i.e., paper with line drawings) were chosen to act as distracters.

Stickers or preferred edibles were available to the children at the end of each session contingent on participation. These items were kept in a container out of their view until the end of the session. Data sheets and a pen were available to the experimenters but were not in view of the participants.

Dependent Variables and Data Collection

The main dependent variable was the percentage of correct category sorts during categorization test sessions. During test sessions, all six stimuli were placed in a line in front of the participants. The experimenter sat next to each participant to ensure that inadvertent cueing did not take place through tracking of the experimenter's eye gaze. Only two experimenters conducted sessions and both had been trained in the problem of inadvertent cueing and how to lessen the likelihood of its occurrence. The experimenter held up one stimulus as prescribed according to the data sheet. The data sheet was split into six-trial blocks for all training and testing conditions. A *correct category sort* was scored when participants selected the remaining two stimuli that belonged to the same set and did not select the stimuli that belonged to the other set. In Experiment 1, stimuli were part of the same class if they evoked the same motor response.

During motor response training, each stimulus was presented between one and three times within each block (depending upon training phase) and only one stimulus was ever presented consecutively within a block (e.g., stimuli may have been presented as ABBABA but never AABBAB). With the exception of these constraints, presentation order was otherwise randomized. During motor response, vocal response, and listener pretesting, each stimulus was presented once per block and stimulus presentation was randomized. Pretesting was conducted to test for common responses among three stimuli that might be emitted in the absence of training so that novel stimuli could be introduced (though this never occurred).

Other dependent variables included the percentage of correct selection responses during Listener tests and the number of trials to criterion during motor response training. A correct response during Listener tests was scored when participants selected the sample that had been trained to the same motor response as emitted by the experimenter. Data on spontaneous vocal tacts or motor responses were also collected throughout sessions.

Interobserver Agreement (IOA)

Interobserver agreement (IOA) was assessed for 34.6% of all motor response training sessions and during 30% of all testing sessions. Agreements and disagreements were scored for each trial. An agreement was scored when both observers marked the same data (correct, incorrect, prompted, or no response) for that trial. Point-by-point agreement was calculated by dividing the number of agreements by the sum of agreements and disagreements multiplied by 100. IOA for training sessions yielded 98.2% agreement (range, 95.1–100%) and IOA during testing sessions yielded 100% agreement across all sessions and participants.

Independent Variable Integrity (IVI).

A second observer collected data during 26.9% of pretraining sessions, 30% of motor response training sessions, 33.2% of Listener test sessions, and 42.9% of categorization test sessions to assess independent variable integrity (IVI), scoring trials as correct or incorrect. A *correct trial* was scored if: (a) the relevant discriminative stimulus was presented correctly, (b) feedback protocol for that condition was followed, and (c) feedback or reinforcement was given or withheld as prescribed. IVI was calculated by dividing the number of correctly implemented trials by the total number of trials conducted. IVI yielded 99.7% integrity across all sessions (range, 98.2–100%).

Experimental Design

All participants were exposed to vocal and motor Tact Pretests, Listener Pretests, and Categorization Pretests 1 and 2. Each of these tests were conducted a second time following Motor-response Training. Categorization Test 1 included an instruction for the participants to look at the sample prior to selecting the remaining comparison stimuli. Categorization Test 2 included an instruction to look at the sample, emit the motor response previously associated with that stimulus, and then select the remaining comparison stimuli. Correct comparisons were the two remaining stimuli that had been trained with the same motor response. Categorization Test 2 was only conducted if the participant failed to categorize during Categorization Test 1. Phase order for Experiment 1 is shown in Figure 1. Also, see Table 1 for a depiction of the number of stimuli trained or tested in each condition, the number of trials per block, and the training criterion.

Pretraining with Everyday Objects.

Initially, training and testing were conducted with everyday objects to establish instructional control (Miguel et al., 2008). Pretraining was conducted prior to introduction of the arbitrary stimuli. Social praise was provided during all pretraining sessions contingent upon correct responses.

Motor-response pretraining. During motor-response pretraining, each participant was taught to emit a nonvocal motor response in

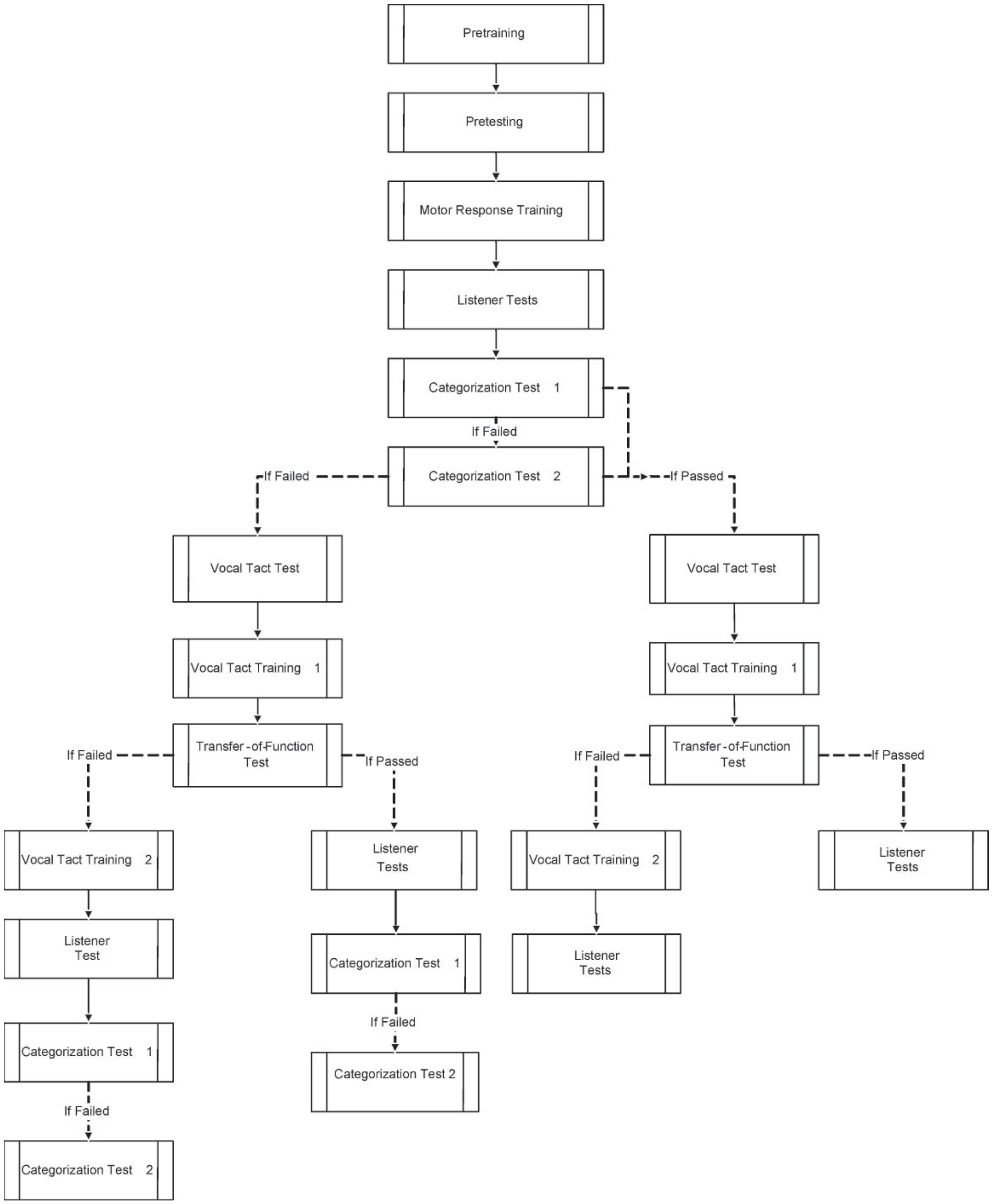

Fig. 1. Visual presentation of the order of training and testing conditions for Experiment 1 and Experiment 2.

the presence of everyday objects (a cup and a toy car). In the initial trial, the cup or the toy car were presented with the instruction, ''This one goes like this,'' while the experimenter modeled the conventional response, either pushing the car across the table or modeling drinking from the cup, and then said, ''Now you show me how it goes.'' The motor

Table 1

Experiment 1 Phases.

Experimental Phases- Experiment 1				
Phase	Task	Number of stimuli	Trials per block	Training criterion
Pre-training	Motor-response	6	6	1 block at 100%
Pre-training	Vocal-response	6	6	1 block at 100%
Pre-training	Listener	8	6	2 blocks at 100%
Pre-training	Categorization	6	6	2 blocks at 100%
Pre-test	Vocal-response	6	6	N/A
Pre-test	Motor-response	6	6	N/A
Pre-test	Categorization	6	6	N/A
1	Motor-response Training (A1/B1; A2/B2; A3/B3)	2 per set	6	2 blocks at 100%
2	Motor-response Training (All Sets)- Sr+ Fading	6	6	2 blocks at 100%
3	Listener tests	6	6	N/A
4	Categorization Test 1 (CT1)	6	6	N/A
5[a]	Categorization Test 2 (if CT1 not passed)	6	6	N/A

[a] Participants were only exposed to this test if they failed Categorization Test 1.

responses were then probed to determine if the participants would emit the response without the model. Three topographically distinct cars and three topographically distinct cups were presented singly in a randomized order determined by a data sheet with three presentations of each item type per block of trials. Training was conducted until the participants responded independently with 100% accuracy across one 6-trial block.

Vocal-response pretraining. Vocal-response pretraining was conducted prior to Experiment 1 (even though vocal responses were not trained until Experiment 2) to ensure that the participants could respond correctly to the instructions that would be given during vocal-response pretests. During these sessions, a stimulus was held up in front of the participant while the experimenter stated the directive, ''Tell me what it is.'' Correct responses were followed by praise, a brief intertrial interval of approximately 5 s, and the presentation of the next trial. Incorrect responses were followed by the statement, ''No, it is a (name). What is it?'' Training continued until 100% accuracy across one 6-trial block.

Listener pretraining. This training helped to ensure that participants would select the appropriate object when its function was modeled. During this training, a three-stimulus array was placed in front of the participant consisting of a car, a cup, and a distracter item (which was either a pen or a hat). During each trial a different type of car or cup was used as the sample and the comparison. At the beginning of each trial, the experimenter stated, ''Give me the one that goes like this,'' while modeling the conventional behavior. The experimenter then waited by extending out her hand directly in the middle (in relation to the participant) and behind the objects that were on the table. If the participant emitted an incorrect response (i.e., selected the wrong object, did not respond, or emitted the motor response without the object), the experimenter replaced the incorrect object, restated the directive-with-model, and pointed to the correct object while stating, ''This is the correct'' and again waited for the participant to select the object. If the participant did not respond or responded incorrectly for a second time, the experimenter repeated the procedure while guiding the participant's hand to the correct object. The same sequence was used if no responses occurred during a trial. Training was conducted in six-trial blocks and continued until a mastery criterion of 100% across two blocks of trials.

Categorization Pretraining

The purpose of this condition was to train the participants to respond appropriately to the instructions that were given during categorization tests. During Categorization Test 1, the experimenter held up the sample conventional item, stated ''Look, give me the others'' and waited for the participant to select from the remaining stimuli. Correct responses were followed by social praise (e.g. ''good job!'') and the removal of the stimuli from the table.

No intertrial activities were presented and the intertrial interval was approximately 5 s. If the participant did not respond within 5 s or responded incorrectly, the experimenter took out the correct comparisons and held them up while stating, ''These are the correct.'' The experimenter then replaced the items and repeated the trial within 5 s, physically guiding the participant to touch the correct comparisons, if necessary. If the participant selected only one of the correct comparisons, the experimenter stated, ''Are there any others?'' and waited 5 s for the participant to respond before scoring the trial as incorrect. If the participant selected all of the comparisons, the experimenter stated, ''I don't want all of the objects, just some of them,'' replaced the objects, and repeated the trial once. Prompted trials were noted. This training was conducted in preparation for Categorization Test 1. Training for Categorization Test 2 was conducted in the same manner, except the participant was asked to emit the appropriate motor response prior to selecting the comparison stimuli. The experimenter held up the sample stimulus, stated, ''Look, show me how this one goes,'' and handed the object to the participant. Immediately following the emission of the motor response or, if the participant did not emit the motor response within 5 s, the experimenter stated, ''Give me the others.'' For the selection response, the same correction procedure was used as for Categorization Test 1. This training was conducted in six-trial blocks until 100% accuracy across two blocks of trials.

Experimental Procedures

Once instructional control was demonstrated during pretraining, pretests were conducted, followed by motor response training and then posttests. In Experiment 1, pre- and posttests were conducted to assess the listener behavior and categorization responses. Participants were directly trained to emit a common motor response to each of the stimuli in Set A and a common motor response to each set of stimuli in Set B. Procedures for these conditions are described below (see Table 1).

Pretests. Pretests were conducted in order to (1) assess whether participants would assign common vocal responses or listener responses to the stimuli and (2) assess whether participants would categorize the stimuli according to physical features or assigned names prior to training.

During vocal-response pretests, the experimenter presented each stimulus singly and stated, ''tell me what it is.'' During listener pretests, two of the training stimuli and one distracter stimulus were placed on the table and the experimenter stated, ''Give me the one that goes like this'' while modeling the motor response. During categorization pretests, all stimuli were on the table and the experimenter removed one and then stated, ''Give me the others.'' During these tests, no programmed consequences were delivered for responses. Following all responses, the stimuli were removed, the experimenter recorded the response on the data sheet, and the next trial was presented. The intertrial interval was approximately 5 s. No intertrial activities were presented. Pretests of Categorization Test 2 were not conducted because if there were any common motor responses that might aid in categorization responses on Categorization Test 2, the stimuli would have been replaced with a set that did not occasion common motor responses with no explicit training. The purpose of motor-response pretests was to demonstrate that the participant did not emit common motor responses prior to training, and probing these responses again during a second categorization test may have resulted in formation of arbitrary common tacts by the participant. Thus, it was decided not to run pretests for Categorization Test 2. More detailed descriptions of listener pretests and categorization pretests are provided below.

Motor-response training. The purpose of this training was to teach common motor responses to two 3-member stimulus sets (Set A and Set B). Both sets of stimuli consist of pieces of paper printed with one of six arbitrary line drawings (see Figure 2). For Set A, motor responses consisted of folding the paper from one short side of the paper to the other. For Set B, motor responses consisted of folding the paper from one long side of the paper to the other (see Figure 3).

The stimuli were always placed in front of the participant facing the same direction, such that the short sides were always to the left and right of the participant and the long sides were the top and bottom. Training was conducted by placing one stimulus on the desk. The stimuli were presented according to the order

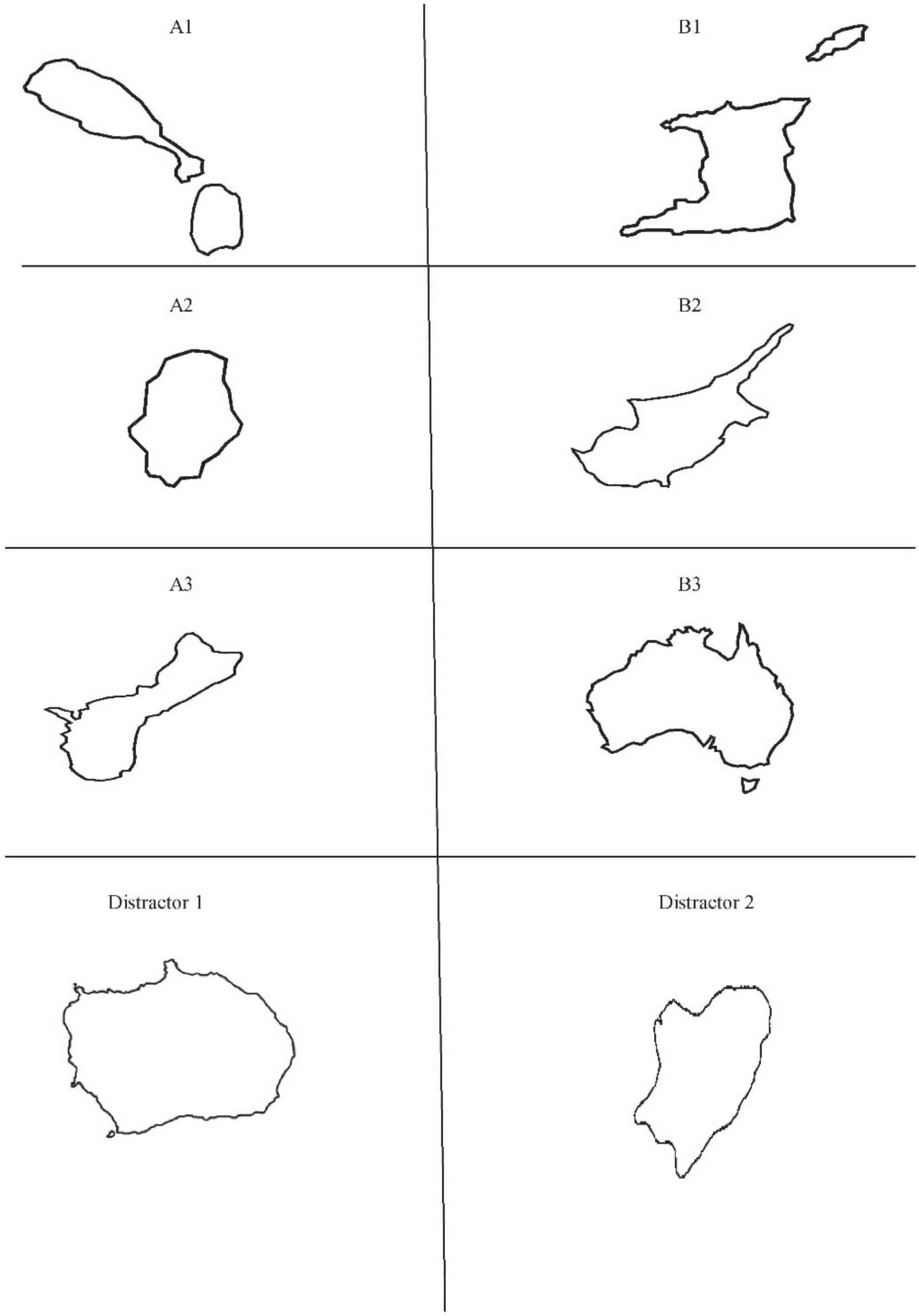

Fig. 2. Visual presentation of the stimuli used during all training and testing sessions with arbitrary stimuli.

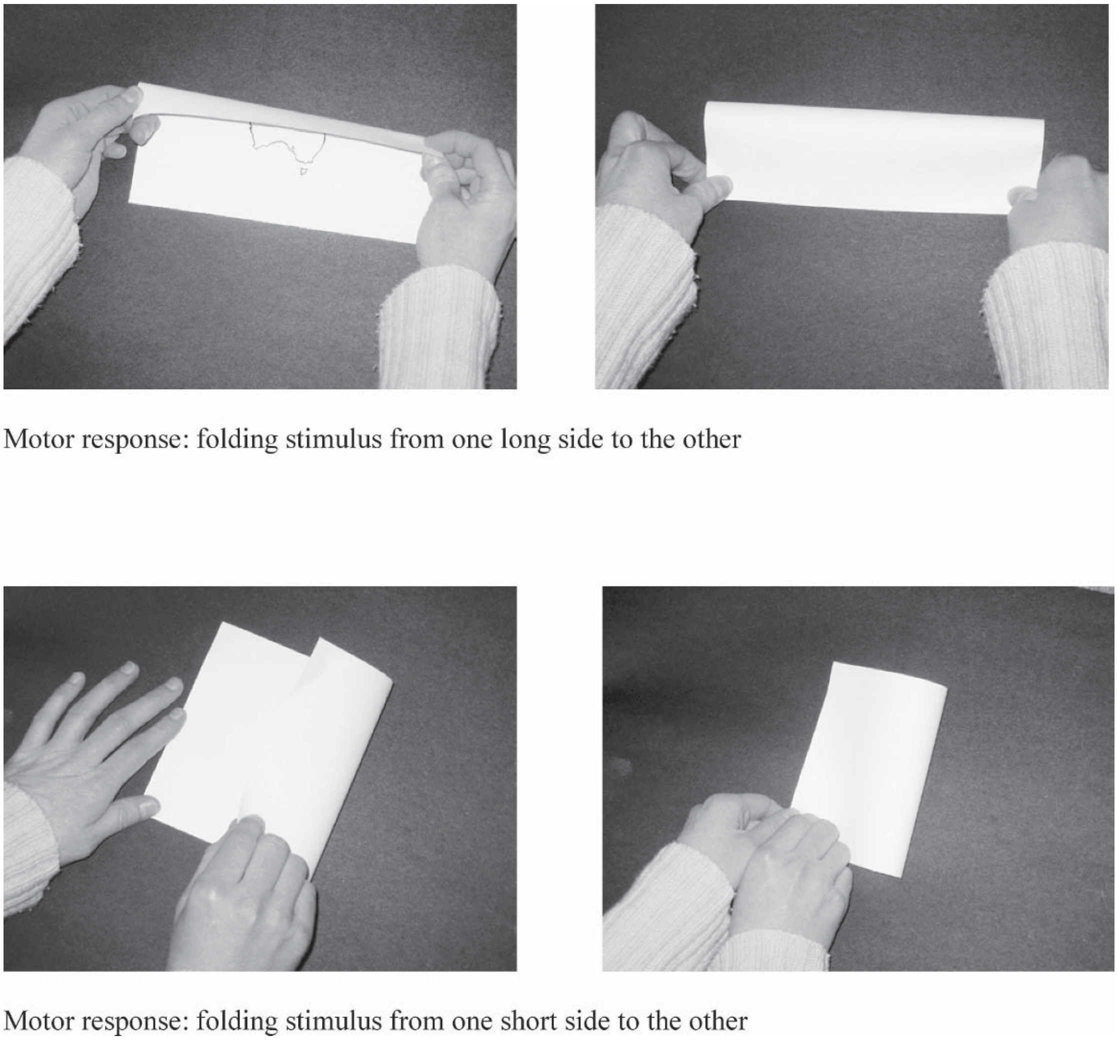

Fig. 3. Visual presentation of the motor responses used in Experiment 1.

prescribed on the data sheet. During all trials within the first block of training, the experimenter presented the stimulus and stated, ''This one goes like this,'' while modeling the correct response. The experimenter then instructed, ''Now you show me how it goes.'' If the participant responded accurately for one block of trials, the model was delayed by 3 s, the experimenter presented the stimulus and instructed, ''Show me how it goes.'' If the participant responded incorrectly or did not respond within 3 s, the experimenter modeled the response, stating, ''It goes like this.'' After an incorrect response, the trial was repeated once. Correct responses were followed by praise. At no time during training did the experimenter refer to the name of the object or its function. Because creases may have been made in the pieces of paper, following a single presentation of a stimulus, it was replaced with an exact copy to eliminate the possibility that the modeled product (i.e., the fold lines) might prompt responding.

Two motor responses were trained to six arbitrary stimuli to potentially form two 3-member classes. Training was conducted until mastery criterion (100% accuracy across two consecutive blocks) was achieved. Stimuli A1 and B1 were trained until the participant responded correctly across two blocks of trials and then stimuli A2 and B2 were introduced. These training blocks were interspersed with blocks of training with A1 and B1. Following mastery of stimuli A2 and B2, stimuli A3 and B3 were introduced. All three pairs of stimuli were then intermixed in three 6-trial blocks. After training criterion was met, the probability of reinforcement was reduced from 100% to 50% (i.e., every other correct response was reinforced) until the same criterion was again met. Final training sessions were conducted with no reinforcement. Prior to the reinforcement reduction phases, the participants were told, ''Now sometimes I will tell you when you are right and sometimes I won't.'' Reinforcement was reduced so that performance during subsequent tests could not be attributed to a sudden change in the rate of reinforcement.

Listener pre- and posttest. Listener tests were conducted to determine if the participants would select the appropriate stimulus when the experimenter modeled the relevant motor response. During each trial, the participant was presented with a three-stimulus array consisting of one stimulus from Set A, one from Set B, and one distracter stimulus. Stimulus presentation was always prescribed on a data sheet and was counterbalanced between the left, right, and center positions. The experimenter began each trial by instructing, ''Give me the one that goes like this,'' while modeling the appropriate motor response using blank paper. No consequence was provided during these trials but the responses were still recorded. Tests were conducted in six-trial blocks for a total of two blocks of trials.

Categorization Pre- and Posttest 1. Categorization tests were used to assess the effects of motor response training on categorization (i.e. stimulus class formation). During Categorization Test 1, the experimenter held up one sample stimulus from either Set A or Set B. The remaining stimuli from both sets were on the table in front of the participant. Comparison stimuli were presented in quasirandom order across trials. While holding up the stimulus the experimenter stated, ''Look,'' while waiting for the participant to hand over the remaining two stimuli. If the participant did not respond, the experimenter provided further direction by stating, ''Can you give me the others?'' If the participant selected only one of the remaining two stimuli, the experimenter stated, ''Are there any others?'' This additional prompt was only provided if the participant selected one stimulus, and it continued to be provided after each selection until the participant refrained from selecting for 10 s. Selection of the correct stimulus within 5 s of this additional prompt was scored as correct. If the participant selected just two stimuli the trial was scored as correct or incorrect and no further prompts were given, if the participant selected more than two stimuli but not all the stimuli the trial was scored as incorrect, and if the participant selected all the stimuli an additional instruction was given (see below). If the participant did not respond, the experimenter waited for 10 s and then presented the next trial. If the participant selected all of the pictures, the trial was scored as null. In this case, the experimenter provided further instruction by saying, ''I don't want all of the pictures, just some of them,'' and repeated the same trial. The null trial was not counted as one of the six trials and the participant had a second opportunity to respond correctly. After three consecutive null trials this condition was terminated and Categorization Test 2 was conducted. Trials in which these prompts were used were noted. Following correct, incorrect, and null trials, the response made by the participant was recorded on the data sheet, the stimuli were removed from the table, and the next trial was started. Tests continued for two 6-trial blocks during which each stimulus was presented once. No feedback or reinforcement was provided, and the exact response made by the participant was recorded.

Categorization Posttest 2. These tests were conducted in the same manner as Categorization Test 1, except that the participant was given the opportunity to emit the appropriate motor response prior to selecting the remaining comparisons. During these tests, the experimenter held up one sample while the remaining five stimuli were on the desk in front of the participant. The experimenter stated, ''Look, how does this one go?'' After the participant emitted the motor response the experimenter stated, ''Give me the others.'' If the participant did not emit the motor response, the experimenter waited 5 s, and then continued the trial by stating, ''Give me the others.'' Incorrect motor responses did not receive differential consequences. Howev-

Table 2
Experiment 1 Pretraining, Training, and Vocal Response Test Data.

Name	Age	Pre-training					Pre-Test	Motor-response Training	
		Motor-R	Vocal-R	List Test	Cat Test1	Cat Test2	Vocal-R	All Sets	Sr+ Fading
Stephen	5.1	1	1	1	1	1	0	43	4
Meghan	4.6	1	1	1	3	3	0	45	4
Kim	5.1	1	1	1	1	1	0	31	4
Andrew	4.9	2	1	1	1	1	0	29	9
Gavin	5.3	1	1	1	1	1	0	24	4

er, data were collected on both the motor response and the selection response.

Results and Discussion

Participant age and results of the pretraining with everyday objects are depicted in Table 2. During pre-training with everyday objects all but one participant emitted the correct motor response in the presence of each everyday object, all participants selected the correct object in the presence of the modeled function, and all but one participant selected the correct comparisons during categorization tests. Two training blocks were necessary for Meghan to categorize everyday objects in 100% of trials and one training block was necessary for Andrew to emit the correct motor response with everyday objects in 100% of trials.

During pretesting, none of the participants emitted a common motor or vocal response for either of the stimulus sets during motor-response pretests and vocal-response pretesting. In addition, none of the participants correctly sorted the stimuli during categorization pretests.

Training data for Experiment 1 are shown in Table 2. Testing data for Meghan and Stephen are displayed in Figure 4. During motor-response training (of arbitrary stimuli), all participants learned the motor response for each of the six stimuli in 45 trial blocks or less. For Stephen 43 blocks of trials were necessary. During listener tests, when the experimenter modeled one of two motor responses using a blank piece of paper, Stephen selected the correct comparison (i.e., the comparison that had occasioned that motor response during motor-response training) in 100% percent of trials. Stephen did not tact more than one stimulus in either Set A or Set B according to a common word or phrase or pass categorization tests during Experiment 1. During Categorization Test 1, he did not correctly select the remaining two stimuli in the set in any of the trials. On the first trial during the first block of six trials, Stephen selected all of the stimuli so the trial was scored as null and repeated. When repeated, he selected the three stimuli that were next to one another in the five-choice array. On subsequent trials during this block, Stephen continued to select either two or three stimuli that were at the right end of the five comparisons. On the first trial of the second block of Categorization Test 1, he selected one stimulus, which was in the same set as the sample (i.e. a correct comparison). After the experimenter prompt, ''Are there any more?'' he selected one additional stimulus that did not belong to the same set (i.e. an incorrect comparison). On all subsequent trials, he selected two stimuli that were next to one another but that were not both in the same set as the sample.

During Categorization Test 2, the motor response previously associated with the sample stimulus was emitted prior to selecting the remaining two stimuli in the set. Stephen emitted the correct motor response when occasioned by the sample stimulus and the experimenter instruction, ''show me how it goes'' in 83% and 100% of trials. When selecting the corresponding comparison stimuli, he followed the same pattern as in Categorization Test 1; he consistently selected two stimuli each time that were next to one another.

For Meghan, 45 blocks of trials were necessary to complete motor-response training. Following this training, she did not select the correct comparison during listener tests (she did select one stimulus on each trial, but it was always an incorrect comparison). During

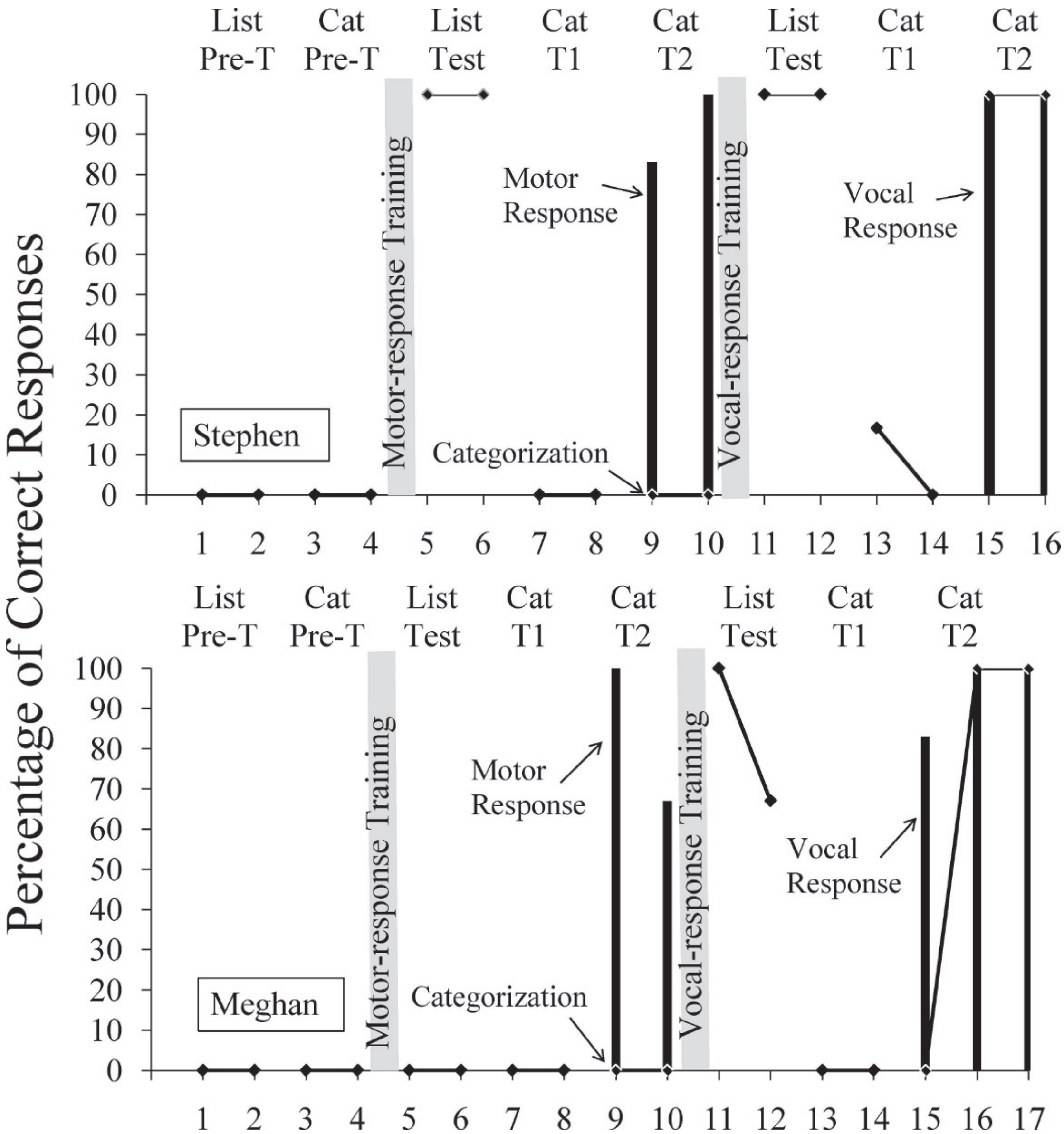

Fig. 4. Stephen and Meghan's listener and categorization test data for Experiment 1 and Experiment 2.

Categorization Test 1, she did not select the correct two comparisons in any of the trials. Similar to Stephen, she selected two or three stimuli on each trial and the stimuli were always next to each other. Due to a prolonged time period of one week between Categorization Test 1 and Categorization Test 2, probes of the motor responses were conducted prior to the second categorization test. Meghan emitted the correct motor responses during 83% and 100% of trials during these probes. During Categorization Test 2, she emitted the correct motor response to the sample in 100% and 63% of trials but did not categorize correctly (0%). The selection responses during these tests resembled that of Categorization Test 1.

Testing data for Kim, Andrew, and Gavin are shown in Figure 5. For Kim, 31 trial blocks were necessary to master the motor responses.

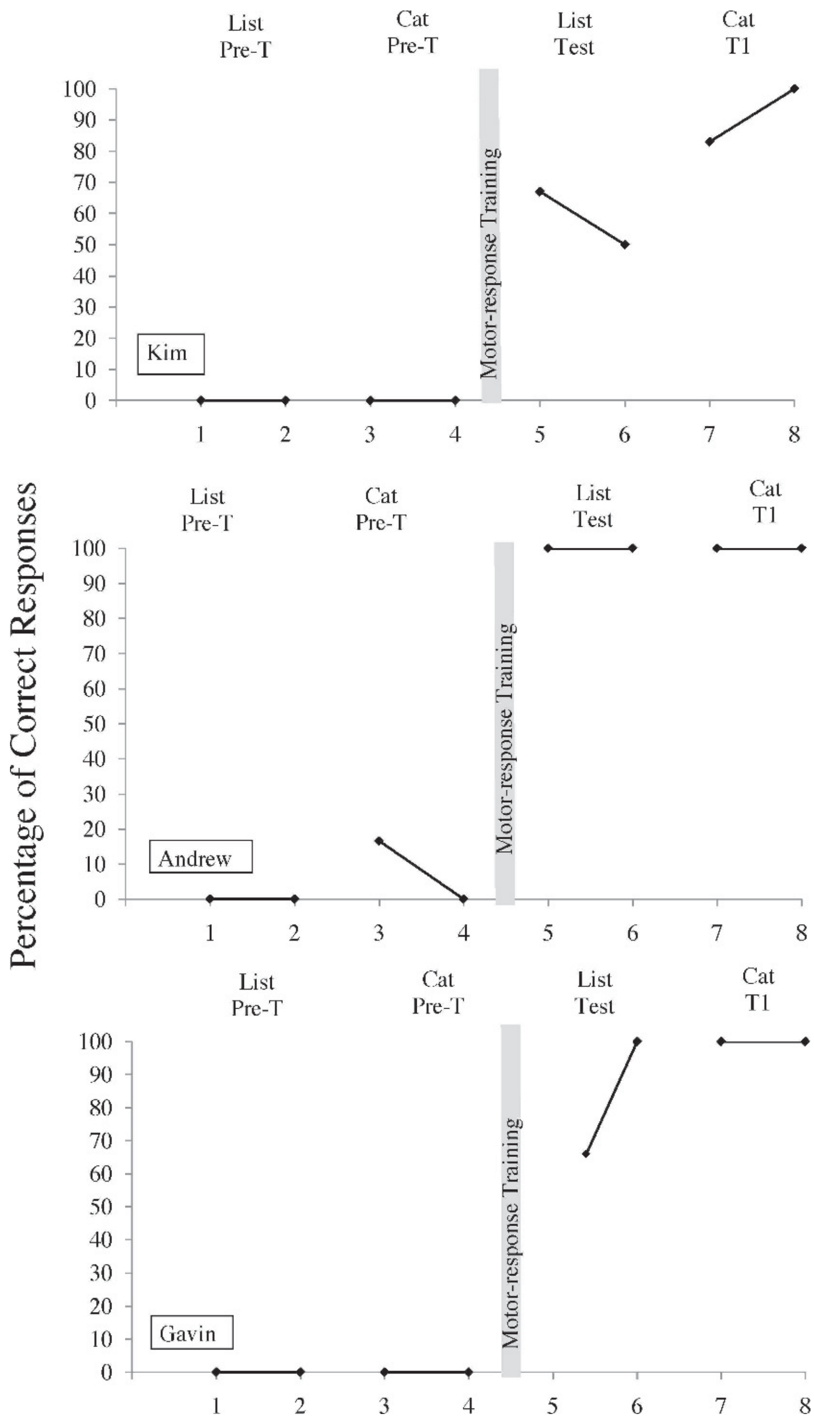

Fig. 5. Kim, Andrew, and Gavin's listener and categorization test data for Experiment 1.

During listener tests, in the presence of the experimenter's model of a motor response, Kim selected the correct comparison during 67% and 50% of trials. Kim did not tact more than one stimulus in either Set A or Set B with a common word or phrase. During Categorization Test 1, Kim selected the two stimuli that had been trained to the same motor response as the sample in 83% and 100% of trials. Because she passed this test, Kim was not exposed to Categorization Test 2.

For Andrew, 29 trials were necessary to complete motor-response training. Following this training phase, Andrew selected the correct comparison during listener tests in 100% and 100% of trials across two blocks of trials. Following listener tests, Andrew categorized correctly during Categorization Test 1 in 100% and 100% of trials. Andrew was not exposed to Categorization Test 2 in Experiment 1 because he passed Categorization Test 1.

Gavin reached mastery criterion during motor-response training in 24 trials. He selected the correct comparison stimuli in 67% and 100% of trials during listener tests. During Categorization Test 1, he selected the correct comparisons in 100% and 100% of trials and was not exposed to Categorization Test 2.

In summary, Kim, Andrew, and Gavin categorized following training of the motor response, while Stephen and Meghan did not categorize at all in either categorization test and therefore were tested again for categorization at the end of Experiment 2. All 5 participants moved on to Experiment 2. The first objective of data collection during Experiment 2 was to train the 2 participants who did not categorize following motor response training to emit vocal responses in the presence of the same two sets of stimuli and then test again for categorization. Past research has demonstrated that training common vocal responses is sufficient to establish categorization responses; however, to our knowledge there are no demonstrations of the effectiveness of common vocal responses after some other response topography (i.e. motor responses) was shown to be unsuccessful in generating categorization responses. A second objective of Experiment 2 was to evaluate whether the stimulus classes formed during Experiment 1 would be demonstrated as functional classes.

EXPERIMENT 2

Goldiamond (1966) asserted that the definition of a functional class should include the property that responses controlled by a few stimuli in the class will come to be evoked by the remaining stimuli in the class without explicit reinforcement. Though restricting the definition of a stimulus class somewhat, some ambiguity remains. Dougher and Markham (1996) called attention to the fact that this definition has not led to full agreement within the field regarding what constitutes a functional class. Though these authors raised several questions pertaining to this issue, the point of interest for this study concerns the relation between functional classes and other types of stimulus classes. Coherence of a stimulus class can be demonstrated through categorization tests, however it remains unclear whether the processes that give rise to stimulus class formation also give rise to functional class formation. In Experiment 2, all 5 participants were exposed to transfer-of-function tests; 3 participants had passed categorization tests in Experiment 1 and 2 participants had not. Naming theorists have never failed to establish functional classes, suggesting that their training and testing procedures that give rise to stimulus classes are also sufficient to establish functional classes with preschool-age children. Experiment 2 tested whether the common motor responses trained during Experiment 1 were sufficient to establish a functional class with this population. More specifically, Experiment 2 was conducted to assess (1) whether participants who did not categorize during Experiment 1 would do so after learning a common vocal responses for each stimulus that has been previously trained to a motor response and (2) whether the function of one stimulus in the set (i.e. the novel vocal response) would transfer to the remaining stimuli in the set with no further training. Participants who passed categorization tests in Experiment 1 were not exposed to categorization tests in Experiment 2. Following transfer-of-function tests, these participants still learned vocal tacts to the remaining two stimuli. The purpose of this was to demonstrate that participants would learn the remaining tacts if they were taught them directly, thus showing that failure to pass transfer-of-function tests was not due to conflicting stimulus control or property of the stimulus itself that might prevent tact transfer or acquisition. For example, failure to observe transfer of the tact to novel stimuli could be due to some property of a stimulus from Set A that is shared by stimuli from Set B. This shared physical property could occasion a common response (unobservable to the experimenter) that precludes acquisition of a novel response. However, learning common vocal responses readily (i.e. within one or two trials per stimulus) when they are directly trained would suggest that this is not the case.

METHOD

Participants, Setting, and Materials

Participants and materials were the same as in Experiment 1. After completion of Experiment 1, participants were immediately exposed to Experiment 2.

Dependent Variables and Data Collection

The main dependent variables were the percentage of correct tacts during transfer-of-function tests (described below) and the percentage of correct sorts during categorization tests. Other dependent variables included the percentage of correct vocal tacts during training and the number of correct selection responses during listener tests. Prior to transfer-of-function tests, tact training was conducted with one stimulus in each of the previously established motor response classes. Transfer-of-function tests assessed the transfer of these tacts to the other stimuli in each class. Categorization tests were conducted as in Experiment 1 and assessed the effects of training a vocal, rather than motor response on categorization.

Interobserver Agreement (IOA)

Interobserver agreement was conducted as in Experiment 1. IOA was assessed in 34.6% of training sessions and yielded 99.4% agreement (range, 97.9–100%) across all participants. IOA was assessed in 40% of testing sessions and yielded 100% agreement across all sessions.

Independent Variable Integrity (IVI)

A second observer collected data during 30% of pretraining sessions, 34.3% of Tact Training 1 sessions, 20% of transfer-of-function test sessions, 27.4% of Tact Training 2 sessions, and 44.4% of categorization test sessions to assess IVI by scoring trials as correct or incorrect. IVI was scored and calculated as in Experiment 1 and yielded 95.2% (range, 76.6–100%) agreement across all sessions. The low end of the range of IVI (76.6%) occurred during a tact training session in which a trial was skipped by the experimenter and was instead conducted at the end of that block of trials. All other sessions yielded IVI of 86% or better.

Experimental Design

For all participants, vocal tact tests were conducted followed by vocal tact training to one stimulus and transfer-of-function tests. If the correct tact did not transfer to the remaining stimuli in each set, these tacts were trained directly. After vocal tacts reached 100% accuracy (either through transfer among stimuli in a class or through direct tact training), Categorization Tests 1 and 2 were conducted with those participants who did not pass categorization tests in Experiment 1. Categorization Test 1 was conducted the same as in Experiment 1. In Categorization Test 2, the experimenter held up the sample stimulus and stated, ''What is it?'' to evoke the vocal tact (rather than the motor response) prior to selection of the comparison stimuli. Conditions were conducted in the following order: tact testing, tact training, transfer-of-function tests, Tact Training 2, Categorization Test 1, and Categorization Test 2. During all sessions, the experimenter was seated next to the participant. Reinforcement was provided as specified for each condition. Phase order for Experiment 2 is shown in Figure 1. Table 3 displays the phase labels, number of stimuli trained or tested number of trials per block, and the training or testing criteria for each condition.

Vocal-Response Pretraining

Pretraining was conducted to establish stimulus control of the instruction over tact responses. The same stimuli were used as in pretraining in Experiment 1. Though this condition was conducted prior to training in Experiment 1, it was repeated before Experiment 2 in part to ensure that the tact would be emitted in the presence of the instruction

Table 3

Experiment 2 Phases.

Experimental Phases–Experiment 2				
Phase	Task	Number of stimuli	Trials per block	Training/Testing criteria
Pretraining	Vocal response	6	6	1 block at 100%
Pretest	Vocal response	6	6	N/A
1	Vocal-response training (1 stimulus Set A + 1 stimulus Set B)	2	6	2 blocks at 100%
2	Vocal-Response Training 1–Sr+ Fading	2	6	2 blocks at 100%
3	Transfer-of-function tests	4	8	2 blocks at 75%
4	Vocal-Response Training 2	6	6	2 blocks at 100%
5	Vocal-Response Training 2–Sr+ Fading (if CT not passed in E1)	6	6	2 blocks at 100%
6	Listener tests	6	6	N/A
7	Categorization Test 1	6	6	N/A
8[a]	Categorization Test 2	6	6	N/A

[a] Participants were only exposed to this test if they failed Categorization Test 1

rather than a motor response or another incorrect response. During these sessions, the object was held up in front of the participant while the experimenter stated the directive, ''Tell me what it is.'' Correct responses were followed by praise and the presentation of the next trial. Incorrect responses were followed by the statement, ''No, it is a (name). What is it?'' Training continued until 100% accuracy across one six-trial block.

Procedures

Vocal-Response Training 1. The purpose of this condition was to train a novel response to one stimulus in each set so that transfer-of-function tests could be conducted following acquisition of these responses. Tact Training 1 was conducted with only one stimulus from Set A and one stimulus from Set B of the two sets trained in Experiment 1. The two stimuli that were selected varied across participants. Arbitrary labels that the participants could easily repeat were selected for these pictures (e.g., ''vug'' and ''ket''). Data sheets were constructed that prescribed stimulus-presentation order for the experimenter, such that the two stimuli used were presented three times in random order during each block of trials. During tact training, the experimenter held up either the stimulus from Set A or the stimulus from Set B while stating the directive, ''It is a [name]. Now you tell me what it is.'' The stimulus was not handed to the participants so that the motor response could be emitted. Once the participant answered correctly across all trials within one 6-trial block, the experimenter held up the stimulus and stated, ''Tell me what it is.'' Incorrect vocal responses were followed by the statement, ''No, it is a [name]. What is it?'' while the experimenter pointed at the object. If the participant did not respond, or stated, ''I don't know,'' the same correction was given, but the experimenter did not state, ''No.''

All tacts of the stimuli, whether correct or incorrect, were recorded. Correct responses were followed by social praise. Tact training continued until 100% accuracy across two consecutive 6-trial blocks. After this criterion was met, the probability of reinforcement was reduced as in Experiment 1 from 100% to 50% until criterion was again met, and then at 0% reinforcement until criterion was again met.

Transfer-of-function tests. The purpose of this condition was to test for the emergence of functional classes. Due to the previous training of one of two motor responses to each set of stimuli, past research (Lowe et al., 2005) suggests that a novel response trained in the presence of only one stimulus in the set might transfer without explicit reinforcement to the remaining stimuli in the set, which have been shown to be part of the same class through categorization tests. This condition tested whether the tacts trained to one stimulus in each set during Tact Training 1would transfer to the stimuli that had received no explicit training. These tests were conducted immediately after tact training had been completed (i.e., within the same session). The four remaining stimuli from Sets 1 and 2 in Experiment 1 were presented singly in front of the participant while the experimenter

Table 4

Experiment 2 Pretraining and Testing, Training, and Transfer-of-Tunction Test Data.

Participant	Age (yrs)	Vocal-R Pretrain[a]	Vocal-R Pretest[b]	Vocal Response Training: Number of training blocks until mastery				T-of-F Tests[b]
				Training 1	Sr+ Fading 1	Training 2	Sr+ Fading 2	
Stephen	5.3	1	0	7	4	3	6	0
Meghan	4.8	1	0	5	4	69	16	0
Kim	5.2	1	0	6	4	4	—	0
Andrew	5	1	0	5	4	5	—	0
Gavin	5.4	1	0	9	4	6	—	0

Note: Kim, Andrew, and Gavin were not exposed to Sr+ Fading 2 because they were not exposed to subsequent listener or categorization tests. ''T-of-F'' refers to transfer of function. T-of-F tests were conducted between Sr+ Fading 1 and Training 2.

[a] Data represent the number of training blocks until mastery.

[b] Data represent the percentage of correct trials.

stated the directive, ''Look, tell me what it is.'' No feedback or reinforcement was provided. Tests were conducted in eight-trial blocks (so that each of the remaining four stimuli was presented twice per block) across two consecutive blocks of trials. Passing criterion was set at 75% (or 6 out of 8 correct trials) across two consecutive blocks of trials.

Vocal-Response Training 2. Tact Training 2 was only conducted if the participant failed the transfer-of-function tests as described above. When this occurred, the same procedures described during the initial tact training were conducted with all three stimuli from each stimulus set. These stimuli were presented singly as prescribed randomly on the data sheet until 100% accuracy across two 6-trial blocks was achieved.

Listener tests. Listener tests were conducted in Experiment 2 to determine whether the participant would select the stimulus in the presence of its spoken name. Two trained stimuli and one distracter stimulus were placed on the desk and the experimenter stated, ''Give me the [name].'' No consequences were provided during these trials but the responses were still recorded. Tests were conducted in six-trial blocks across two blocks.

Categorization Test 1. Categorization Test 1 was conducted exactly as in Experiment 1.

Categorization Test 2. Categorization Test 2 was conducted as in Experiment 1 except that the instruction was changed to, ''Look, tell me what it is,'' to encourage a vocal rather than a motor response to the sample. After the participant emitted a response the experimenter stated, ''Give me the others.'' In order to keep the categorization tests as similar as possible to those in Experiment 1, the sample stimulus was still placed on the table in front of the participant. This provided the opportunity for the participant to emit either the motor response or the vocal response, however no participants emitted the motor response.

Results and Discussion

Pretraining results (Table 4) are shown as the number of pretraining blocks that were conducted. If the participant passed the first pretraining block of trials (which were unprompted to probe for independence before training) no further blocks were conducted. Training blocks were then conducted until the participant reached 100% accuracy across two blocks of trials. In Experiment 2, all participants emitted correct tacts for everyday objects with no explicit training (i.e., the first block was at 100% accuracy).

Table 4 also displays the training data during Experiment 2 and the transfer-of-function test data. Figure 4 contains Stephen and Meghan's listener test and categorization data (pre- and posttest) for Experiment 2. During Vocal-response pretests, none of the participants self-generated tacts that remained consistent across the two trained sets. They did, however, make up names for some, if not all, of the stimuli. Stephen named most of the stimuli, but not on every trial (on some trials he stated, ''I don't know''), and he did not assign one tact to more than one stimulus. Kim, Meghan, and Gavin did not tact every stimulus; they consistently stated, ''I don't know'' when asked to tact the stimuli. Kim did call one stimulus ''cow'' and Meghan called one stimulus ''unicorn''. Andrew stated, ''planets'' during some of the stimulus presentations but did not emit any other tacts.

During Tact Training 1, Stephen was taught to emit the vocal response, ''vug'' in the presence of A1 and the vocal response, ''ket'' in the presence of B2. It took seven sessions for Stephen to meet criterion during vocal-response training. During transfer-of-function tests, Stephen did not emit the tact that had been trained during Vocal-Response Training 1 in the presence of the other stimuli in each set. During Tact Training 2, three sessions were necessary for Stephen to meet criterion. Following a single training trial per stimulus, he tacted each of these four stimuli correctly on all subsequent trials. During Listener tests, when the experimenter stated, ''Give me the 'ket/vug,''' Stephen selected the correct stimulus from a three-choice array in 100% of the trials. During Categorization Test 1, in the presence of one stimulus from either Set A or Set B, Stephen selected the remaining two stimuli in that set in 17% and 0% of trials. On the first trial of the first block, Stephen selected the two correct comparisons. On the remaining five test trials, Stephen selected three comparison stimuli. On four of these five incorrect trials, he selected the three right-most stimuli in the comparison array. In the second block of trials, Stephen always selected three comparison stimuli. During Categorization Test 2, when the vocal response was required prior to selecting the remaining stimuli in the set, Stephen categorized correctly in 100% and 100% of trials. He also emitted the correct vocal response in 100% of the trials.

Meghan responded similarly during Tact Training 1 and 2 and during transfer-of-function tests. During Tact Training 1, five sessions were necessary for her to learn the tact relations to the stimuli A2 and B3. Following this training, the remaining two stimuli in each set failed to occasion these tacts. These tacts were then trained directly during 69 Tact Training 2 blocks. During Listener tests, when the experimenter stated one of the two tacts that had been taught during tact training, Meghan selected the correct comparisons in 100% and 63% of trials. During Categorization Test 1, Meghan selected one, two, or three stimuli on each trial, but they were never the correct comparisons. During Categorization Test 2, she selected the correct comparison in 0% of trials and tacted the sample correctly in 83% of trials. During these six trials, Meghan incorrectly tacted one stimulus as a ''ket'' and categorized accordingly, such that four stimuli evoked the tact ''ket'' and two stimuli evoked the tact ''vug''. When Meghan tacted the sample as ''ket'' she selected the remaining three stimuli that she also tacted ''ket'' when they acted as the sample, and when she tacted the sample ''vug'' she also selected the other stimulus that she tacted ''vug'' when it was the sample. This error was addressed by conducting a remedial tact training phase that included two additional blocks of trials. This phase was conducted exactly the same as previous tact training sessions. Meghan tacted the stimuli correctly in 83% and 100% of trials during these two blocks of trials. She was then exposed to Categorization Test 2 a second time. During these sessions, she tacted the sample correctly in 100% and 100% of trials and she categorized correctly in 100% of the trials.

During Tact Training 1, Kim was taught in six sessions to emit the vocal response ''ket'' in the presence of A3 and was taught to emit the vocal response, ''vug'' in the presence of B1. During Transfer-of-Function Tests, Kim did not emit the tacts that were trained during Vocal-response Training 1 in the presence of the untrained stimuli. In the presence of A1 and A2, she did not emit the vocal response, ''ket'' and in the presence of B2 and B3, she did not emit the vocal response, ''vug.'' During Tact Training 2, when these remaining tact relations were taught directly, four sessions were necessary before Kim emitted the correct vocal tact in 100% of trials across two consecutive blocks of trials. Kim was not exposed to categorization tests again in this experiment because she had passed these tests during Categorization Test 1 in Experiment 1.

Neither Andrew nor Gavin assigned common names to the arbitrary stimuli during vocal tact pretests. They each learned the vocal tact relations to one stimulus from each set in under six blocks of trials. Following this training, the remaining stimuli in the set did not evoke these tacts for either participant. Since both of these participants passed Categorization Test 1 in Experiment 1, neither of them was exposed to these tests again in Experiment 2.

In summary, the 2 participants who did not pass categorization tests in Experiment 1 passed listener tests and categorization tests after learning vocal tacts for the same stimuli in Experiment 2. None of the participants passed transfer-of-function tests after learning the tacts to one stimulus in each set. All participants were then exposed to Tact Training 2, during which these relations were trained directly.

GENERAL DISCUSSION

During Experiment 1, a motor response was taught to three stimuli in Set A and a second motor response was taught to three stimuli in Set B. Following this training, Stephen, Andrew, and Gavin engaged in the corresponding listener behavior by selecting the correct comparison in the presence of the experimenter's model of the motor response. Andrew and Gavin also passed subsequent categorization tests while Stephen did not. The remaining 2 participants (Kim and Meghan) did not pass listener tests. Meghan subsequently failed categorization tests while Kim categorized accurately. Stephen and Meghan were retested for categorization in Experiment 2 because they did not pass categorization tests in Experiment 1.

During Experiment 2, it was verified that after common motor response training, a vocal response that was trained in the presence of a single stimulus in the set was not emitted in the presence of the other stimuli until it was directly trained. In other words, following Tact Training 1, in which participants learned a vocal response to one exemplar of each set, no participants emitted these tacts in the presence of the remaining stimuli in each set during transfer-of-function training. All participants were subsequently taught to tact the remaining stimuli in each set. After learning the vocal tacts for each of the stimuli in the two sets, Stephen and Meghan passed Categorization Test 2 in Experiment 2.

There were three tests of emergent repertoires within this study, and the results of each will be discussed separately.

Listener Behavior

According to Horne and Lowe (1996), the speaker (either the motor or vocal response) and listener relations (demonstrated during Listener tests) are both necessary for categorization to occur. In this study, however, after speaker training, the emergence of corresponding listener behavior did not reliably predict performance on categorization tests. For instance, Stephen passed the tests of listener behavior but failed categorization tests in Experiment 1. Even though he emitted the motor response in the presence of the stimulus and selected the stimulus in the presence of the motor response, he did not categorize during Categorization Tests 1 or 2. In Experiment 2, after vocal tact training, Stephen passed listener tests (i.e., he selected the correct comparison in the presence of the auditory sample) and subsequently passed Categorization Test 2. Kim and Meghan failed listener tests when the sample was a motor response. Despite this, Kim passed Categorization Test 1 in Experiment 1. On the other hand, Andrew and Gavin passed both listener and categorization tests, consistent with what would be predicted by the naming hypothesis.

There is, however, an important limitation within the listener tests as they were conducted in this study. There is a distinction between the listener tests administered after training a motor response compared to the sort of listener behavior originally discussed by Horne and Lowe (1996). These authors describe that the relevant listener response may be covert vocal behavior that consists of saying the name of the sample stimulus, the hearing of which is discriminative for orienting and selecting the correct comparison stimulus. In their examples, the stimulus produced when tacting the sample, the auditory stimulus, is very similar whether emitted by the experimenter (when stating, ''give me the vek'' as in listener tests) or the participant (when tacting the sample stimulus ''vek'', or self-echoing, during categorization tests). Within the context of this study, these tests may be limited in that the relation between the sample stimulus (in the form of a motor response) and the selection response may be quite dissimilar to the listener relation that occurs during categorization tests. More specifically, during listener tests, the sample consisted of the experimenter modeling the motor response; during categorization tests, however, the sample consisted of the stimulus arising from the participant modeling the motor response herself. Horne et al. (2007) conducted tests of listener behavior as well and cited the same inconsistency between the listener behavior relevant during these tests and the listener behavior relevant during categorization tests. In their study, participants' failure to pass listener tests may also be explained, in part, by this difference.

Listener tests conducted following vocal response training did not present the same sort of logical problem as these tests with the motor response. When the sample is a vocal response, the auditory stimulation produced by the experimenter's vocal behavior is very similar to the type of stimulation arising from the participant's tact-to-sample. Unlike a motor response, this response (i.e., echo) shares point-to-point correspondence with the sample. This is a better test of the listener behavior and may help to explain why Stephen categorized in Experiment 2 and not in Experiment 1. In other words, though Stephen passed listener tests in Experiment 1, it is unclear whether these tests accurately reflected the listener repertoire that is important in mediating between the sample and the comparison during categorization tests. His performance during Categorization Test 2 in Experiment 2 might be accounted for by the availability of a common vocal tact for the stimuli in each set that may be emitted covertly at any time, while his failure to categorize during Experiment 1 might have been due to the difficulty in emitting a motor response while scanning comparison stimuli.

Kim's pattern of responding during listener tests deserves some additional speculation. Kim failed listener tests, but did so systematically in that she never selected the incorrect trained comparison. Kim's response pattern was interesting in that the distracter stimulus was almost always selected in the presence of the motor response associated with the stimuli in Set A while the correct comparison was always selected in the presence of the motor response associated with the stimuli in Set B. Future researchers should be aware that similar disruptions in performance are possible when distracters are introduced during testing conditions.

Categorization

The first distinction between this study and Horne et al. (2007) was the nature of the motor responses; specifically in the Horne at al. study they were topographically very different from one another. Horne et al. stated as one of their main conclusions that, ''manual sign naming was found to be as effective as vocal naming in establishing arbitrary stimulus categorization'' (Horne et al., 2007, p. 367). The current study did not achieve these results. Two participants did not pass categorization tests following manual sign training but did so following vocal response training. Importantly, it is possible that, for participants in both studies who passed categorization tests, manual signs evoked tacts (words) that were subvocally emitted during categorization tests. Future studies could explore whether manual signs without subvocal tacts are sufficient to produce categorization responses by replicating these procedures with participants who display no tact repertoire. The second issue with selection of a nonvocal response raised in the current study had to do with the use of hand signs. It is possible that participants may have used the hand sign as a meditational response between stimulus presentation and the selection response (either overtly or covertly, as one would emit covert vocal behavior). The motor response used in the present study necessitated the use of a piece of paper such that the likelihood of mediational signing would be decreased (only a close variation to the actual hand motion could be emitted without the tactile stimulation from the paper). Lessening the likelihood of a mediating motor response may explain the inconsistent results during listener tests, as well as categorization tests, again suggesting a second type of covert mediational response that may be used, when available, to improve performance during testing conditions (Lowenkron, 1988).

Past research has found that vocal responses are efficient in developing complex stimulus classes in young children and that motor responses alone have generated these classes as well (Horne et al., 2007). Though it is impossible to know what the participants were doing covertly, training procedures that reduce the likelihood that participants can covertly tact the stimuli (e.g., using a manual sign tact) may be reflected in less consistent performance during categorization tests. The poor performance during categorization tests following motor-response training with Stephen and Meghan suggests this may have been the case. Additionally, for both Meghan and Stephen, though there was no additional training between Categorization Tests 1 and 2 during Experiment 2, correct responding increased from 0% to 100% when participants were required to tact the sample stimuli. Responding during the second categorization test may have been accurate because emitting the vocal response prior to selecting the remaining stimuli provided more precise stimulus control over the selection response. According to the naming hypothesis, this stimulus control would involve both the speaker and listener repertoires: In the presence of the sample stimulus, the participant emits a vocal tact and this vocal tact then produces auditory stimulation to which the participant then responds as a listener by selecting the correct comparisons that had previously occasioned similar auditory stimulation.

The final methodological distinction between this study and the Horne et al. (2007) study has to do with stimulus presentation during tact training. Horne et al. first trained common manual signs in the presence of two pairs of stimuli and then conducted training with all four trained stimuli present on the table. After a review and additional training of the third pair, all six stimuli were presented on the desk during training. During tact training, the experimenter pointed to each stimulus when asking the participant to tact. Showing all of the stimuli to the participants together during a condition in which reinforcement was delivered may have inadvertently reinforced categorization responses. For example, Miguel et al. (2008) found that children who failed categorization tests after tacts were trained in a single-presentation format passed categorization tests after training with all of the pictures grouped together. In the current study, the stimuli within a set were never presented together prior to categorization tests and 2 participants did not pass these tests. These results suggest that learning to tact each stimulus in the class in a single-stimulus presentation may be insufficient to produce emergent categorization for some participants. Presenting the stimuli together may produce additional stimulus control (compared to that when stimuli are presented singly) that accounts for the improved responding during categorization tests for 3 participants (McIlvane & Dube, 2003).

Transfer-of-Function

Functional classes are defined as those classes in which all stimuli in the class occasion the same response (Sidman, 1994, p. 82). However, an additional test of a functional class is to train a novel response to one stimulus in each set and then test whether the remaining stimuli in the set evoke the response. The present study evaluated the transfer of a vocal response to untrained stimuli that were part of the same functional class (due to a common motor response).

All participants failed transfer-of-function tests in Experiment 2, though it is unclear why. In speculation, however, failure of participants to vocally tact during these tests may be attributed to the conditioning history of the participants, specifically that they had a history of emitting some tact during vocal-response tests. During vocal-response tests all participants emitted an untrained tact to at least one stimulus and often stated ''I don't know.'' This history may have decreased the likelihood that the vocal responses trained during Vocal-Response Training 1 would transfer to the remaining stimuli. For example, when asked ''What is it?'' a name that had been emitted prior (either overtly or covertly) may have been evoked and could have essentially blocked acquisition of the novel response, in a similar process to that reported on regarding blocking with compound stimuli (Pierce & Cheney, 2004).

It has been asserted that functional classes and equivalence classes are one in the same (Sidman, 1994). However, in the current study participants did not demonstrate cohesion of a functional class but did pass categorization tests. This would support the notion that the formation of an equivalence class does not necessarily imply presence of a functional class or vice versa (Sidman et al., 1989). In other words, participants who pass tests of equivalence may not pass transfer-of-function tests. The current results are analogous to this occurrence in that emergent categorization was observed with 3 participants but these participants did not pass transfer-of-function tests. Future research could investigate the role of the topography of the common response (e.g., vocal vs. motor or similar vs. topographically distinct), the role of training procedure (e.g., conditional discriminations vs. tact training), and the role of testing conditions (e.g., a single training and testing phase vs. multiple exemplar training).

Conclusions

The most parsimonious explanation for the observed performances is that, for some participants, training a vocal response may be more likely to exert stimulus control over categorization responses than training a motor response. However, it has been demonstrated that the training of an overt verbal response is unnecessary for the establishment of stimulus relations (Carr et al., 2000; Schusterman & Kastak, 1993; Sidman, 2000; Sidman, Cresson, & Willson-Morris, 1985; Sidman, Kirk, Willson-Morris, 1974; Sidman & Tailby, 1982). In the present study, stimulus relations existed for 3 participants when only a common motor response was trained, but for the other 2 this was insufficient. According to Sidman (1994), it is not necessary to assume that a verbal response occurs within relations between nonverbal stimuli; it is the relations among stimuli that are observable and therefore subject to analysis.

It remains unclear what the participants were doing covertly. For instance, Set A may have been labeled *fold to the side* and Set B may have been labeled *fold to the bottom*, or simply *side* and *bottom*, although the participants never overtly stated these tacts during the course of the study. It may be speculated that participants who passed categorization tests in Experiment 1 were either emitting similar vocal tacts covertly or were covertly emitting the motor response and that participants who failed these tests were not, but this analysis is not definitive. If participants emitted vocal tacts during categorization tests during Experiment 1, an explanation as to why no common tacts were emitted during tact pretesting is warranted. Tacts may not have been occasioned during testing due to the absence of stimulus control over that specific response, or the presence of a contextual cue that evoked a different response altogether. For instance, the experimenter instruction, ''What is it?'' may have indicated to the participant that the picture had a preselected name that they had not learned (apart from a tact of its function). This seems unlikely because all of the participants assigned arbitrary names to at least some of the stimuli, and stated, ''I don't know,'' in the presence of a few of the stimuli. None of the participants assigned common names to any stimuli within a set.

Further studies evaluating the relationship between naming and stimulus class formation are warranted. In Horne et al. (2007) children categorized after learning a common hand sign to each stimulus in the set. In the current study, only 3 of the 5 participants passed categorization tests after having learned only a motor response. However, there are several important differences between the methods utilized. First, the responses that were trained and tested were very different. The responses used in Horne et al. were hand signs that were emitted in the presence of the stimulus. These hand signs consisted of placing the hand on the shoulders and placing one fist on top of the other. These responses clearly corresponded to body parts for which a young child may emit the vocal tact. It is possible that the participants were tacting each stimulus using a vocal response that corresponded to the learned motor response. This, however, remains an inference because the authors reported that no participant emitted a vocal tact during any part of the study that may have been the result of this type of stimulus control. Second, there were some differences between the training procedures in the two studies. In the present study, the pictures were not seen together by the participant until categorization tests. In Horne et al., all stimuli were present on the desk during the final motor response training and listener response training phases. Training of the tact and selection response in this manner may have altered the stimulus control exerted by the categorization tests because the stimuli were seen together during both training and testing conditions. So, during training, the participant may have tacted the stimuli in Set A in succession and those in Set B in succession. It is possible that this response would then have contacted the reinforcement contingency in place during training phases. If this occurred, categorization was then learned in the presence of reinforcement prior to the tests conducted by the experimenters that are done under extinction.

These distinctions in the training and testing conditions between the present study and the Horne et al. (2007) study raise several experimental questions. For instance, future research could address whether common manual signs differ from vocal responses with respect to interaction among stimuli. Also, studies should evaluate under which training and testing conditions these classes will emerge.

The classes that formed following motor-response or vocal-response training (demonstrated through categorization tests) may be equivalence classes. Investigators conducting research in this area might explore this outcome by using methods similar to those introduced in Sidman, Wynne, Maguire, and Barnes (1989). These authors simply introduced novel stimuli to preexisting functional relations, trained a novel stimulus to one stimulus in the class, and then tested relations between the novel stimuli and the remaining stimuli in the class using conditional discriminations.

The present study used transfer-of-function tests to assess the formation of a functional class, unlike in the Sidman et al. (1989) study. However, the question of whether participants that have demonstrated class formation through categorization tests (or transfer-of-function tests) would then pass the tests of equivalence still remains. In other words, if a novel stimulus acts as the positive comparison in an array with a previously-trained stimulus as the sample, will the participant then select other stimuli from that class when they are among the comparisons and the novel stimulus acts as the sample? Such research would address questions about whether conditional discriminations would be sufficient for class expansion beyond those stimuli trained directly and could also test for transfer of vocal responses to the novel stimuli.

As mentioned above, the results of the current study did not unequivocally support the naming hypothesis. Future research should therefore evaluate other possible sources of control that aide in stimulus categorization. For instance, this responding could be evaluated as higher order operants that emerge due to repeated exposures to the training and testing conditions. Also, 3 participants categorized after learning only the motor response. This evidence provides further support that a verbal repertoire may be unnecessary for the establishment of these stimulus classes. Future research should evaluate what underlying mechanism is responsible for the formation and expansion of these classes.

REFERENCES

Carr, D., Wilkinson, K. M., Blackman, D., & McIlvane, W. J. (2000). Equivalence classes with individuals with minimal verbal repertoires. *Journal of the Experimental Analysis of Behavior, 74*, 101–114.

Dougher, M. J., & Markham, M. R. (1996). Stimulus classes and the untrained acquisition of stimulus functions. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 137–152). Amsterdam: Elsevier Science, B. V.

Goldiamond, I. (1962). Perception. In A. J. Bacharach (Ed.), *Experimental foundations of clinical psychology* (pp. 280–340). New York: Basic Books.

Goldiamond, I. (1966). Perceptions, language, and conceptualization rules. In B. Kleinmuntz (Ed.), *Problem solving* (pp. 280–340). New York: Wiley.

Hineline, P. N. (1997). How, then, shall we characterize this elephant? *Journal of the Experimental Analysis of Behavior, 68*, 297–300.

Hayes, S. D., Barnes-Holmes, D., & Roche, B. (2001). Relational frame theory: A précis. *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 141–154). New York: Kluwer Academic/ Plenum Publishers.

Horne, P. J., Hughes, J. C., & Lowe, C. F. (2006). Naming and categorization in young children IV: Listener behavior training and transfer of function. *Journal of the Experimental Analysis of Behavior, 85*, 247–273.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*, 185–241.

Horne, P. J., Lowe, C. F., & Harris, F. D. A. (2007). Naming and categorization in young children: V. Manual sign training. *Journal of the Experimental Analysis of Behavior, 87*, 367–381.

Horne, P. J., Lowe, C. F., & Randle, V. R. L. (2004). Naming and categorization in young children: II. Listener behavior training. *Journal of the Experimental Analysis of Behavior, 81*, 267–288.

Lowe, C. F., Horne, P. J., Harris, F. D. A., & Randle, V. R. L. (2002). Naming and categorization in young children: Vocal tact training. *Journal of the Experimental Analysis of Behavior, 78*, 527–549.

Lowe, C. F., Horne, P. J., & Hughes, J. C. (2005). Naming and categorization in young children: III. Vocal tact training and transfer of function. *Journal of the Experimental Analysis of Behavior, 83*, 47–65.

Lowenkron, B. (1988). Generalization of delayed identity matching in retarded children. *Journal of the Experimental Analysis of Behavior, 50*, 163–172.

McIlvane, W. J., & Dube, W. V. (2003). Stimulus control topography coherence theory: Foundations and extensions. *The Behavior Analyst, 26*, 195–213.

Miguel, C. F., Carr, J. E., & Michael, J. (2002). The effects of a stimulus–stimulus pairing procedure on the vocal behavior of children diagnosed with autism. *The Analysis of Verbal Behavior, 18*, 3–13.

Miguel, C. F., Petursdottir, A. I., Carr, J. E., & Michael, J. (2008). The role of naming in stimulus categorization by preschool children. *Journal of the Experimental Analysis of Behavior, 89*, 383–405.

Pierce, W. D., & Cheney, C. D. (2004). Behavior analysis and learning. In B. Webber (Ed.), *Reflexive behavior and respondent conditioning* (pp. 53–74). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schlinger, H. D. (2008). Listening is behaving verbally. *The Behavior Analyst*, *31*, 145–161.

Schusterman, R. J., & Kastak, D. (1993). A California sea lion (Zalophus californianus) is capable of forming equivalence relations. *Psychological Record*, *43*, 823–839.

Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.

Sidman, M. (1997). Equivalence relations. *Journal of the Experimental Analysis of Behavior*, *68*, 258–266.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, *74*, 127–146.

Sidman, M., Cresson, O., & Willson-Morris, M. (1974). Acquisition of matching to sample via mediated transfer. *Journal of the Experimental Analysis of Behavior*, *22*, 261–273.

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior*, *43*, 21–42.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37*, 5–22.

Sidman, M., Wynne, C. K., Maguire, R. W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the Experimental Analysis of Behavior*, *52*, 261–274.

Skinner, B. (1957). *Verbal behavior*. East Norwalk, CT: Appleton-Century-Crofts.

Vaughan, M. E., & Michael, J. L. (1982). Automatic reinforcement: An important but ignored concept. *Behaviorism*, *10*, 217–227.

Received: May 10, 2010
Final Acceptance: November 5, 2010