The Psychological Record, 2013, 63, 769–784

MEMBERSHIP OF DEFINED RESPONSES IN STIMULUS CLASSES

Karen M. Lionello-DeNolf
University of Massachusetts Medical School, Shriver Center

Paula Braga-Kenyon
The New England Center for Children

Sidman (2000) has suggested that in addition to conditional and discriminative stimuli, class-consistent defined responses can also become part of an equivalence class. In the current study, this assertion was tested using a mixed-schedule procedure that allowed defined response patterns to be "presented" as samples in the absence of different occasioning stimuli. Four typically developing adults were first trained to make distinct response topographies to 2 visual color stimuli and then were taught to match those color stimuli to 2 different form-sample stimuli in a matching task. Three separate tests were given in order to determine whether training had established 2 classes, each composed of a response, a color, and a form: a form-response test in which the forms were presented to test if the participants would make differential responses to them and 2 response-matching tests to determine if the participants would match visual stimulus comparisons to response-pattern samples. Three of the 4 participants showed class-consistent responding in the tests, although some participants needed additional training prior to passing the tests. In general, the data indicated that the different response patterns had entered into a class with the visual stimuli. These results add to a growing literature on the role of class-consistent responding in stimulus class formation and provide support for the notion that differential responses themselves can become a part of an equivalence class.

Key words: equivalence, response, fixed-ratio (FR), differential reinforcement of low rates (DRL), humans

Stimulus equivalence has been traditionally defined as the ability to match (in a conditional discrimination task) certain stimuli together in novel ways after limited training. Specifically, after being trained to match stimulus A to B and stimulus B to C (where A, B, and C denote two or more stimuli), individuals can, without additional training, match in accordance with reflexivity (A–A, B–B, C–C), symmetry (B–A), and

This research was supported by NICHD Grants HD 39816 and HD 04147.

Paula Braga-Kenyon is now affiliated with Melmark New England.

Correspondence concerning this article should be addressed to Karen Lionello-DeNolf, University of Massachusetts Medical School, Shriver Center, 333 South St., Shrewsbury, MA 01545; E-mail: Karen.Lionello-DeNolf@umassmed.edu

DOI:10.11133.j.tpr.2013.63.4.005

transitivity (A–C, C–A; Sidman & Tailby, 1982). These effects have been demonstrated in typically developing human adults and children and also in adults and children with varying degrees of developmental disability (Adams, Fields, & Verhave, 1993; Barnes, McCullagh, & Keenan, 1990; Carr, Wilkinson, Blackman, & McIlvane, 2000). Sidman (2000) has suggested that when training includes class-specific reinforcers or responses, those too become part of the class of equivalent stimuli.

While there is considerable support for the inclusion of reinforcers in the defined class (Barros, Lionello-DeNolf, Dube, & McIlvane, 2006; Dube & McIlvane, 1995; Dube, McIlvane, Mackay, & Stoddard, 1987; Dube, McIlvane, Maguire, Mackay, & Stoddard, 1989; Joseph, Overmier, & Thompson, 1997; McIlvane, Dube, Kledaras, de Rose, & Stoddard, 1992; Schenk, 1994), fewer studies have considered the inclusion of the defined responses. One of the first of such studies was conducted by Manabe, Kawashima, and Staddon (1995). Three budgerigars were trained to make a high call to one visual stimulus (e.g., red) and a low call to another visual stimulus (e.g., green). Then, one bird (S4) was trained on identity matching to sample (MTS) in which red and green were the stimuli; making high and low calls to red and green (respectively) sample stimuli was required, but such behavior was not required for the comparison stimuli. All three birds were trained on a second matching task in which form stimuli were presented as samples and red and green were presented as the comparison stimuli. In order for the comparisons to appear, the birds had to make a call to the form sample stimuli, but the *type* of call was undefined. Either call would cause the comparisons to appear. In this latter task, S4 eventually began to make high calls to the sample associated with red and low calls to the sample associated with green, indicating that the high and low calls may have entered into a stimulus class with the visual stimuli. However, the other two birds did not begin to emit class-specific calls in a consistent manner. In a second experiment, these latter two birds were given training on two additional MTS tasks in a manner similar to S4 in Experiment 1. In this case, both birds began to make high calls to one stimulus and low calls to the other. These data indicate that for all three budgerigars, the high and low calls entered into a stimulus class with the visual stimuli (see Saunders & Williams, 1998, for an alternative explanation; see also Urcuioli, Pierce, Lionello-DeNolf, Friedrich, Fetterman, & Green, 2002, for a systematic replication with pigeons).

Whether defined responses can become class members has also been investigated with human participants. For example, Braga-Kenyon, Andrade, Ahearn, and Sidman (2006) trained seven typically developing adults to make distinct motor responses (such as a hand clap, finger snap, or finger tap) to either two or three stimuli (e.g., A1–R1, A2–R2, and A3–R3). Then, participants were trained on MTS with new stimuli as samples (B) followed by the familiar stimuli as comparisons (A). In order to produce the comparisons, participants had make the R1, R2, or R3 response, but they were not required to make those responses to particular sample stimuli. Five of the seven participants developed differential sample responding, suggesting that the visual stimuli and the responses had become members of a common class.

While the above studies provide suggestive evidence that responses can become part of an equivalence class, more direct evidence is needed. Sidman (1994, 2000) has outlined specifically what sort of training and test procedures need to be conducted in order to provide unequivocal evidence of membership of defined responses in equivalence classes. He suggested that training should consist of two symbolic matching tasks, with a different response required to the correct comparisons associated with each experimenter-defined class (e.g., A1–B1:R1, A2–B2:R2 and B1–C1:R1, B2–C2:R2). Testing should include all possible emergent relations—the typical tests for reflexivity, symmetry, transitivity, and equivalence as well as tests in which the defined responses themselves serve as samples (e.g., R1–A1, R1–B1, R1–C1 and R2–A2, R2–B2, R2–C2). If the responses were members of the class, then, for example, when “presented” with an R1 sample and a choice between A1 and A2 comparisons, participants should choose A1, but when “presented” with an R2 sample and the same two comparisons choices, participants should choose A2 (and likewise for the remaining sample-comparison combinations).

The difficulty with the testing sequence just outlined should be evident: How can different responses be “presented” as samples? One way to accomplish this is to present two different visual stimuli (i.e., observing stimuli) to which participants are trained to make the distinct response patterns. Then in the critical test, those observing stimuli would be presented as samples in a matching task, and making the required response pattern would result in presentation of the comparisons, thus allowing the participant to match a comparison to a response sample. The difficulty with this procedure, however, is that any test performance indicating the inclusion of the response in an equivalence class would not rule out the possibility that such performance was actually due to the inclusion of the visual stimuli rather than the responses per se (see Sidman, 1994, for a discussion of three-term contingencies and stimulus class formation). The issue is to find a way to “present” responses as samples in the absence of different occasioning stimuli.

Lionello-DeNolf and Urcuioli (2003) developed a mixed-schedule procedure in order to use response patterns as samples. Pigeons were trained to make two different responses to the same stimulus in order to avoid having to use differential observing stimuli to occasion the two responses. On each trial, a white square was presented, and the pigeons were required to make either a fixed-ratio (FR) 20 response or a differential reinforcement of low rates (DRL) 3-s response. The response pattern required was random, and there was no cue indicating to the pigeons which one was required. The variable of interest was what the pigeons did after making an error; all pigeons eventually learned to switch to the opposite response pattern after emitting an incorrect response. Then, the pigeons were trained on a matching task in which correct white-stimulus responses were followed by a choice between form comparisons: choosing one comparison after a DRL response pattern and the other comparison after an FR response pattern was reinforced. All pigeons learned this task to greater than 90% accuracy.

Lionello-DeNolf and Urcuioli (2003) did not test their pigeons for equivalence relations as defined by Sidman and Tailby (1982) and therefore were unable to confirm or disconfirm Sidman’s (1994, 2000) assertion regarding membership of defined responses in equivalence classes (but see Urcuioli, Lionello-DeNolf, Michalek, & Vasconcelos, 2006, and Urcuioli & Vasconcelos, 2008, for a test of response inclusion in an acquired equivalence class and related discussion). Using a similar procedure, Shimizu (2006) did include such a test. Eight adult participants were first trained to make three different computer mouse manipulations as defined responses; each manipulation consisted of four mouse movements that included changes in direction (e.g., right, down, right, and up). Next, participants were trained on two separate symbolic matching tasks in which the stimuli did not overlap (A–B, C–D) and in which different mouse manipulations were required when responding to comparisons (e.g., on A1–B1 and C1–D1 trials, response R1 was required when selecting B1 and D1, whereas on A2–B2 and C2–D2 trials, response R2 was required when selecting B2 and D2). Participants were subsequently tested on trials that assessed matching in accordance with reflexivity, symmetry, and transitivity among the visual stimuli and trials that assessed matching response samples to visual comparisons. On those latter tests, participants were presented with a white stimulus and were told to respond to it by making one of the previously learned mouse manipulations, and that if their manipulation was correct, the three comparison stimuli would be presented and they should respond as they did before. Results of this test showed that seven of the eight participants passed the response-stimulus test (and the eighth did on a retest).

The purpose of the current study was to extend the findings of Shimizu (2006) by testing for response membership using the mixed-schedule procedure developed by Lionello-DeNolf and Urcuioli (2003). The complex nature of the responses used by Shimizu were difficult to train (half the participants required remedial response training at some point) and required the use of written instructions as well as visual and auditory prompts (that were faded out prior to test), all of which may have facilitated stimulus class formation. If the visual stimuli associated with each response had become part of the class, those stimuli, instead of the responses, could have been responsible for participants passing tests for equivalence. In the current study, the defined responses were different patterns of “tapping” on a computer touchscreen and could be shaped without the use of verbal instructions or visual cues. Moreover, training on the mixed-schedule procedure prior to the test sessions meant that the participants did not need additional instructions at test because the testing format was already familiar to them.

Four adult participants were given training that should have resulted in the formation of two stimulus classes, each containing one response and two visual stimuli. Participants were then exposed to three separate tests: one in which a visual stimulus was presented in order to determine how the participants would respond to it and two in which responses were samples and the different visual stimuli were comparisons.

Method

Participants

The participants were four female graduate students (age range: 21–25 years) employed at The New England Center for Children, a school serving children diagnosed with special needs. Although all were enrolled in a part-time master’s program in behavior analysis, none had studied stimulus equivalence. Participation lasted for several hour-long sessions that took place over the course of 2 to 6 weeks. Participants were compensated for their time with movie passes, and this was not dependent on experimental performance.

Apparatus

All experimental sessions took place in the automated teaching lab, which has been extensively described elsewhere (Lionello-DeNolf & McIlvane, 2003); only relevant details will be described here. The participants were seated in front of a touchscreen computer monitor mounted on a side wall. Beneath the touchscreen was a small counter on which was placed a plastic bin for token accumulation. To the left of the touchscreen monitor, located on a different wall (at a 45-degree angle), there was a small compartment in which red poker-chip tokens could be dispensed.

The researchers observed participants from the other side of the wall via closed-circuit television. Experimental events were controlled and data were collected by a Macintosh G4 computer running LabView software (National Instruments) located on the researchers’ side of the wall. Tokens were dispensed via a Med Associates token dispenser (Model SG-501). The stimuli presented on the touchscreen were solid white, red, and green squares and two black Greek letters (omega and chi) presented on a white square background.

Procedure

Participants were individually brought into the lab and seated in front of the computer. They were told, “You have to touch the screen. If you touch it correctly, a token comes out here,” while the researcher pointed to the token compartment. One participant, MHA, was given the additional instructions, “When you touch, don’t drag your finger on the screen,” because that topography prevented the computer from recording a touch, and, “If you want to touch the square, touch in the middle,” because she frequently touched the edge of the square. This was problematic because the computer sometimes recorded such responses as hits to the stimulus and at other times recorded them as hits to the background. After delivering the instructions, the researcher left the teaching area, closed the door, and started the computer program.

Phase 1 color-response training. In Phase 1 training, red and green squares were presented on the center of the screen (on separate trials), and participants were required to make either a DRL 3-s response or an FR 10 response. For the DRL 3-s response, two responses were required in order to produce a token. The first press started a timer, and a second press at least 3 s later resulted in the delivery of a token in the token well. Presses that occurred less than 3 s after the timer started caused the timer to reset. For the FR 10 response, 10 consecutive presses were required, each less than 3 s from the previous response. Participants JLI and MHA were required to make the DRL response to red and the FR response to green; for participants GNR and LYK, these contingencies were reversed.

At the beginning of a trial, either the red or the green square was presented on the center of the screen. If the participant made the correct response pattern, a chime sounded, a token was dispensed, and the square disappeared from the screen. A new trial began after a 3-s intertrial interval (ITI). If the participant made an incorrect response pattern (e.g., pressing the square more than 3 s between each response on an FR trial), a "switch cue" sounded, the timer reset, and the stimulus remained on the screen until the correct response pattern was made. There were 40 trials per session. After completion of the 40 trials, participants' accuracies were evaluated. Between successive sessions on a given day, there was a break of 2 to 5 min in which the participant remained seated and the computer monitor was turned off. During this time, the researchers reviewed the data just collected and set up for the next session. Participants remained in Phase 1 color-response training until they reached an accuracy of 90% correct in a session of 40 trials or until they reached an accuracy of 87.5% on two consecutive sessions.

For GNR, after five sessions, accuracy remained at chance for both color stimuli. At that point, she was given further instructions to "touch the square as soon as it appears." After two additional sessions, accuracy remained at chance. At this point, the researcher demonstrated how to complete each response one time.

Phase 2 form-color matching. In Phase 2, participants were trained to match two Greek letter stimuli to the color stimuli from Phase 1. At the beginning of a trial, either the omega or the chi sample was presented in the center of the screen. A single touch resulted in its removal and the presentation of red and green squares on the left and right sides of the screen. A single touch to either comparison stimulus resulted in the end of the trial. If the correct comparison was chosen, the stimuli were removed, a token was dispensed, a "chime" sounded, and a 3-s ITI began. If the incorrect comparison was chosen, all stimuli were removed and a 3-s ITI began. Each Phase 2 session consisted of 20 trials, 10 with each sample stimulus. Red and green comparisons appeared in the left and right positions equally often across successive trials. No additional instructions were given to the participants. For JLI and GNR, red was the correct comparison after a chi sample and green was the correct comparison after an omega sample. For MHA and LYK, these contingencies were reversed. Phase 2 criterion was 90% correct for one session (20 trials). As in Phase 1, successive sessions were conducted after 2- to 5-min break periods in which the participant remained in the teaching area and the researcher examined the data and set up for the next session.

Phase 3 mixed-schedule training. In mixed-schedule training, the participants were trained to make two different responses to a single white square stimulus. The purpose of this training was so that in subsequent testing, the FR 10 and DRL 3-s response patterns could be used as samples in a matching task.

At the beginning of a trial, a white square stimulus was presented in the center of the screen. The participants were required to make one of two responses to it: either an FR 10 or a DRL 3 s (as defined in Phase 1). There was no cue to indicate which response pattern was required on a given trial. If the participant made the correct response (e.g., responded with 10 presses each less than 3 s apart when an FR 10 was required), the "chime" sounded, a token was delivered, and the white square disappeared from the screen. A new trial began after a 3-s ITI. If the participant made the incorrect response (e.g., responded with two presses equal to or greater than 3 s apart when an FR 10 response was required), a "switch cue" sounded, the timer was reset, and the white square remained on the screen. The trial then continued until the participant made the correct response pattern.

There were two variables of interest during this phase of training. Since there was no cue indicating which response pattern was scheduled for a given trial, participants were expected to randomly respond with one or the other response pattern. Thus, *hit rate* was defined as initial accuracy on each trial (i.e., percentage correct). This was expected to remain at chance throughout training. The main variable of interest was the *switch rate*. This is defined as a switch to the other response pattern after an initial error on a trial. For example, if a DRL response was scheduled for a trial and the participant initially responded with an FR response, would participants, after hearing the switch cue, switch response patterns and make a DRL response? The switch rate was expected to increase with training. Correct response patterns that occurred after two or more errors on a trial were not included in the switch rate.

Each session consisted of 60 trials in which FR and DRL responses were required an equal number of times according to a prearranged random sequence. Prior to the start of the first training session, participants were told, "Now things will look different, but you will still receive a token for correct responses." The criterion to complete this phase of training was one session with a switch rate of 90% or better or two consecutive sessions with a switch rate of 85% or better for both sessions. For GNR, switch rate varied between 63% and 82% over six 60-trial sessions. Despite this, GNR was moved to the next phase of the experiment.

After completing Phase 3, participants were given a series of three different tests in order to test membership of the FR and DRL response patterns in the stimulus class. However, if completion of Phase 3 occurred at the end of a 1-hr period (or if the participant indicated she wished to stop for the day), the test series was given on a different day. In those cases, participants were given refresher sessions on Phase 1 and Phase 2 training prior to being tested. The order in which the tests were administered was counterbalanced across participants.

Form-response test. In the form-response test, either the omega or the chi stimulus was presented on the center of the screen and remained there until the participant completed either a DRL 3-s response pattern or an FR 10 response pattern. Thus, any two presses spaced 3 s or more apart ended a trial, even if those two presses were preceded by some presses spaced less than 3 s apart. Once either pattern was complete, the stimulus disappeared from the screen and a 3-s ITI followed. No feedback or tokens were delivered in the test. Twenty test trials were presented, 10 with each stimulus.

Response-form matching test. In this test, each of the 20 trials began with presentation of the white square stimulus. Participants were required to make either a DRL or an FR response to it, and the contingencies were similar to that in the Phase 3 mixed-schedule training described above. However, once the correct response pattern was made, the trial did not end. Instead, the two Greek letter stimuli were presented as comparison stimuli. A single press to either comparison stimulus ended the trial, which was followed by the ITI. The chi and omega stimuli appeared in the left and right comparison positions equally often, and DRL and FR served as samples on 10 trials each. No feedback or tokens were delivered for any comparison selection.

Response-color matching test. Contingencies in this test were similar to that in the response-form test, except that the Greek letter comparison stimuli were replaced by the red and green stimuli.

After completing the first test series, data were examined. In some cases, participants were given further training in an attempt to facilitate class formation. Procedural details in those cases are described in the Results section.

Results

Training

Participants completed Phase 1 (color-response) training in an average of 4.8 sessions (range: 2–10), Phase 2 (form-color matching) training in an average of 2.3 sessions (range: 2–3), and Phase 3 (mixed-schedule) training in an average of 3.3 sessions (range: 1–8). Final accuracy averaged 92.5% (range: 87.5–100%) and 97.5 % (range: 95–100%) for Phases 1 and 2, respectively. For Phase 3, on the final session, the average hit rate was 52.5% (range: 48.3–55%) and the average switch rate was 91% (range: 75–100%).

Inter-response time (IRT) and latency to respond to each stimulus were also analyzed during Phases 1 and 3 to ensure that differential responding developed to the two stimuli (Phase 1) and to the white square (Phase 3). Table 1 shows the average IRT and latency data in the final (criterion) training session for each phase for individual participants. In Phase 1, IRTs averaged 297.0 ms for the color stimulus for which participants were required to make an FR response and 3,781.9 ms for the color stimulus for which participants were required to make a DRL response. Latency data, however, were similar: 703.0 ms and 781.1 ms for each stimulus, respectively. In Phase 3 mixed-schedule training, participants continued to emit two distinct response patterns in the absence of a differential observing stimulus. On trials in which an FR pattern was required, IRTs averaged 438.7 ms, and on trials in which a DRL pattern was required, IRTs averaged 2,990.8 ms. On average, participants waited longer to respond on trials in which an FR pattern was required (623.7 ms) than those on which a DRL pattern was required (504.9 ms).

Table 1
Average IRT (Latency) Data for FR- and DRL-Response Required Trials, in Milliseconds, for Phases 1 and 3 Training for Individual Participants

	Phase 1		Phase 2	
Participant	FR	DRL	FR	DRL
GNR	231.9 (834.8)	3,686.0 (619.4)	468.9 (668.6)	3,054.9 (404.8)
JLI	236.7 (595.3)	3,697.2 (741.2)	806.3 (461.3)	4,754.3 (434.2)
LYK	162.7 (584.9)	3,404.9 (664.3)	221.4 (792.4)	2,099.5 (684.3)
MHA	196.7 (786.9)	4,339.4 (1,099.6)	258.4 (572.5)	2,054.5 (496.2)
M	297.0 (703.0)	3,781.9 (781.1)	438.7 (623.7)	2,990.8 (504.9)

Testing

Figures 1 and 2 depict the percentage correct during each training session and the percentage of class-consistent responding on each of the initial tests for individual participants. Also shown is accuracy on any additional training as well as class-consistent responding on any repeated tests.

GNR. Data for GNR are shown in Figure 1 (top). GNR received one refresher session on each phase of training prior to test because training and test sessions took place on different days. Although she required more sessions to achieve criterion levels of performance in Phase 1 than any other participant, and although her switch rate never reached high (i.e., 90%) levels, GNR made class-consistent responding on at least 90% of the trials in each of the three tests. Thus, this participant showed strong evidence that the DRL and FR responses had become part of a class, along with the color and Greek letter stimuli.

JLI. Data for JLI are shown in Figure 1 (bottom). JLI achieved high accuracy in Phases 1 and 2 in two to three sessions and achieved a switch rate of better than 90% in her first Phase 3 session. However, on the form-response test, she made FR responses to both

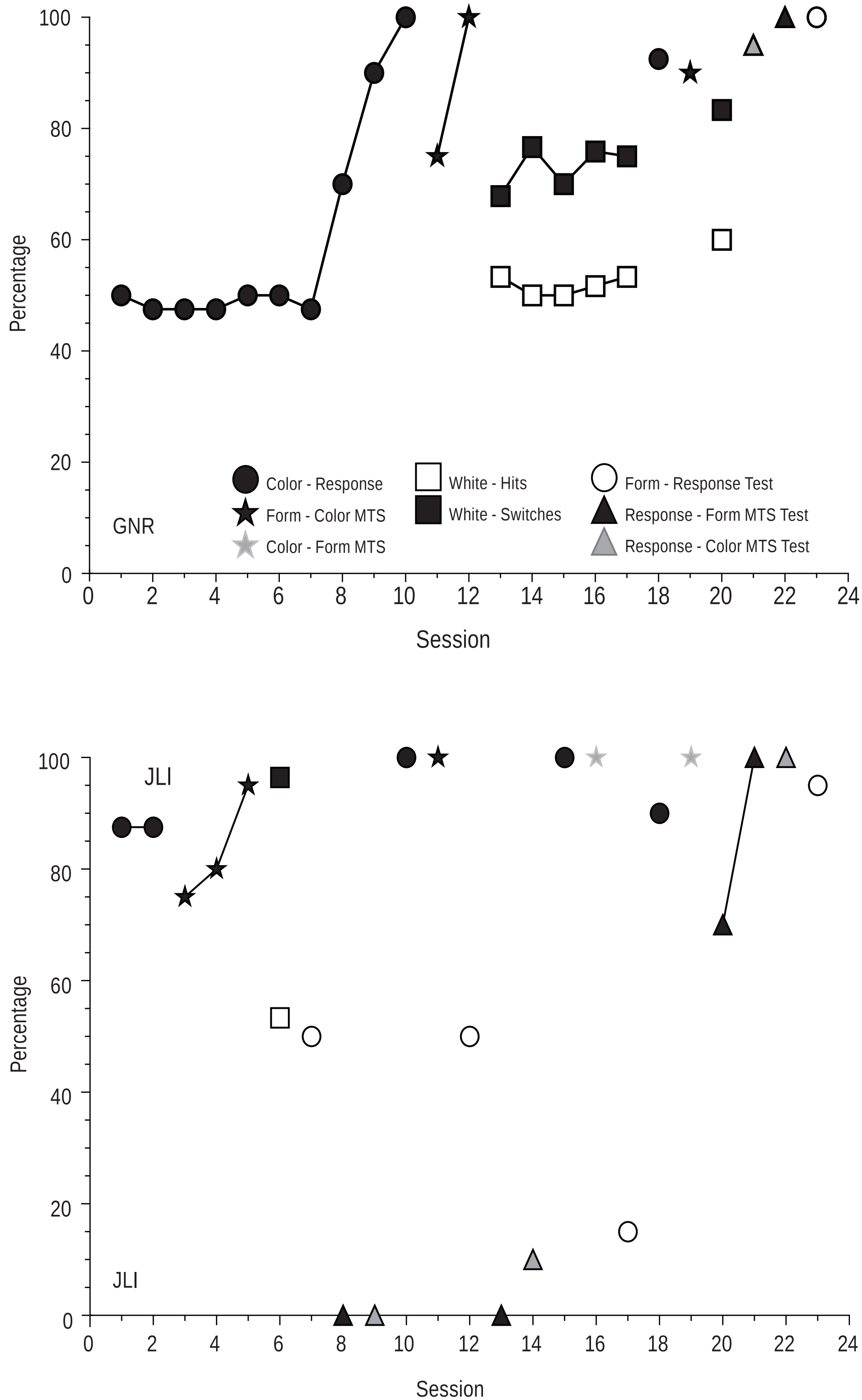

Figure 1. Percentage correct during training and refresher session and percentage of class-consistent responding during test sessions for GNR and JLI.

stimuli, regardless of potential class membership. In addition, on the response-matching tests, she chose the comparison belonging to the *opposite* class on every trial. At this point, JLI was given additional refresher sessions on Phases 1 and 2, and the entire test series was repeated. The results, however, were no different.

Next, JLI was given a refresher on color-response training (Phase 1). Then, she was trained on color-form matching—the symmetrical version of the trained Phase 2 task. This training was followed by a repeat of the form-response test. This time, class-consistent responding was below chance. JLI began to make some DRL responses to the Greek letter stimuli, but those responses were to the stimulus belonging to the opposite class. After two more refresher training sessions, the response-form matching test was given. Class-consistent responding began to emerge during this test. For the first several trials, JLI chose the comparison belonging to the opposite class. Then, that behavior pattern ceased and JLI made class-consistent responding until the end of the session. A second session was given, and all responses were class consistent. On the subsequent response-color matching test, class-consistent responding was 100%. Finally, the form-response test was given again, and this time, more than 90% of the responses were class consistent.

MHA. Data for MHA are shown in Figure 2 (top). MHA required several sessions to acquire the DRL and FR response patterns to the color stimuli in Phase 1. Phase 2 and Phase 3 training were each completed in two sessions. On the response-matching tests, responding was largely class consistent, with 80% class-consistent choices on the response-form test and 100% class-consistent choices on the response-color test. On the form-response test, however, class-consistent responding was at chance levels. Just as JLI did, MHA made all FR responses to both the Greek letter stimuli. She was then given training on color-form matching (the symmetrical version of Phase 2 training) in the same manner as JLI, and the entire test series was repeated. Unlike for JLI, this training did not facilitate class-consistent responding on the form-response test. MHA was then given experience with Phase 1 (color-response) training without tokens or feedback for correct responses prior to being given the form-response test a final time. In this final test, class-consistent responding was 90%, indicating that the DRL and FR responses had become members of a class that also included the color and form stimuli.

LYK. Data for LYK are shown in Figure 2 (bottom). Criterion was met in two sessions for each of the three training phases. Prior to the first test series, LYK was given refresher sessions on Phases 1 and 2. On the form-response test, LYK made all FR responses, just as JLI and MHA did on their initial form-response tests. On the response-form matching test, 100% of choices were class consistent, but on the response-color matching test, 100% of choices were to the opposite-class comparison. LYK was then given one session of training on color-form matching (the symmetrical version of the Phase 2 task), and given the tests again, in a different order. This time, on the response-form test, 100% of choices were to the opposite-class comparison, but on the response-color test, 100% of choices were to the class-consistent comparison. This was the opposite result of the first test series. On the form-response test, LYK again made all FR responses. Refresher sessions on Phase 1 and Phase 2 symmetry were repeated, and a session of Phase 1 training without feedback or tokens was conducted. On the subsequent tests, she made all FR responses to the forms, and made class-opposite responses on both response-matching tests.

LYK's behavior patterns seemed to indicate that she randomly responded in a consistent way on each test (i.e., arbitrary assignment; de Rose, 1996). We speculated that training had not produced *any* stimulus classes, even between the visual stimuli. Thus, we decided to train her on two different matching relations composed entirely of visual stimuli (non-representative black forms on a white background) and then give her standard equivalence tests of reflexivity, symmetry, and transitivity. The results of these tests were similar to those described above: On half the tests, she responded in accordance with equivalence, and on the other half, she responded to the opposite-class stimulus.

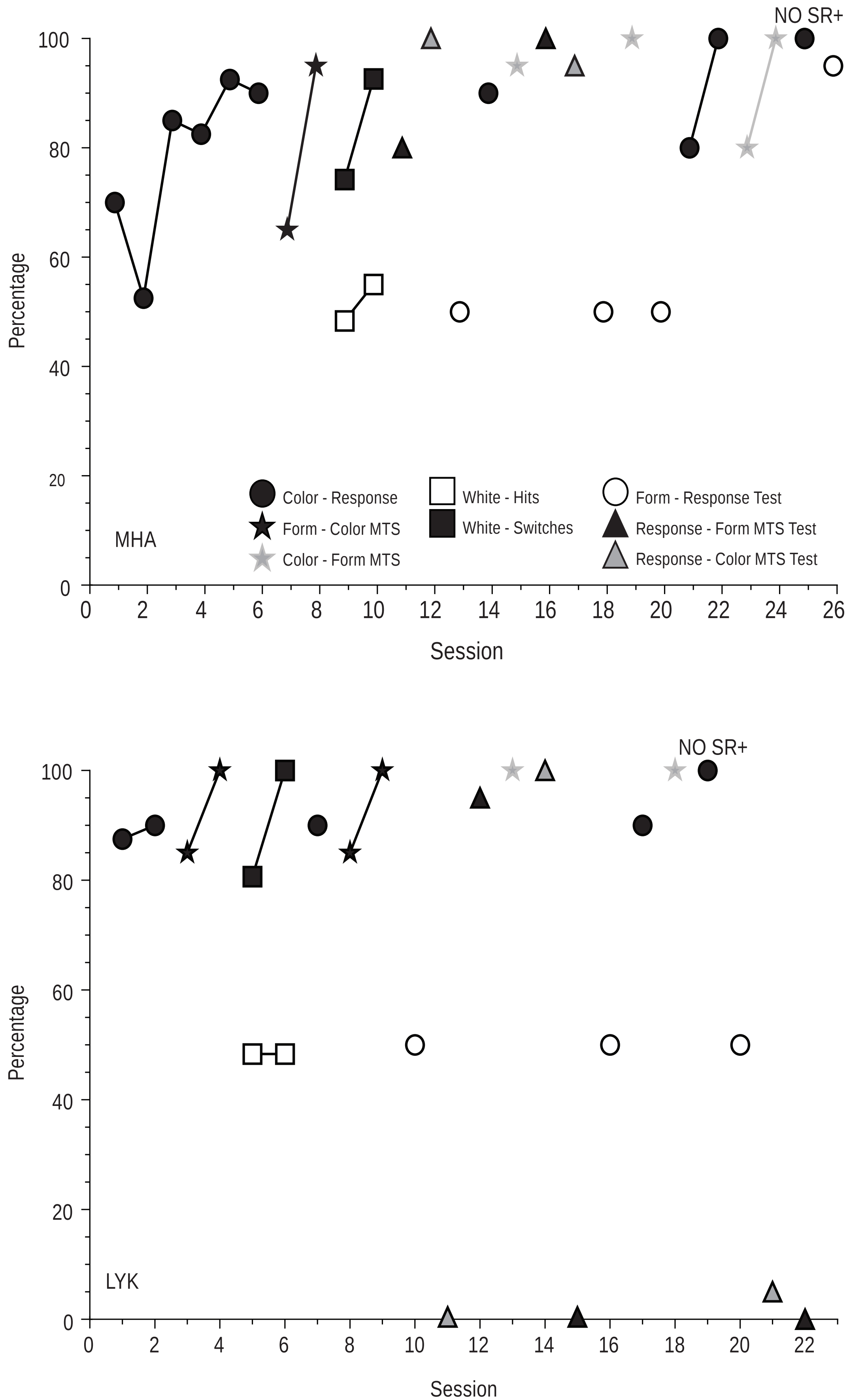

Figure 2. Percentage correct during training and refresher sessions and percentage of class-consistent responding during test sessions for MHA and LYK.

Response-time data. Table 2 shows IRT and latency data for all participants on each of the test sessions. For the form-response test, “FR” and “DRL” refer to the stimulus with which FR and DRL responses were expected to be associated on the basis of equivalence. For the response-form and response-color matching tests, FR and DRL indicate the required response pattern to the white stimulus. Data in bold indicate sessions in which class-consistent responding was obtained. IRT data from the matching tests indicate that all the participants continued to respond with distinct response patterns, although some participants tended to respond more quickly on DRL-required trials in test than they did on those trials in Phase 3 training (e.g., MHA).

Table 2
IRT (Latency) Data, in Milliseconds, From Each Test Session for Individual Participants

Participant	Test session	Form-response test		Response-form MTS test		Response-color MTS test	
		FR	DRL	FR	DRL	FR	DRL
GNR	1	**240.7 (609.8)**	**3,985.8 (796.1)**	**510.7 (780.3)**	**2,501.0 (430.2)**	**447.1 (1001.0)**	**2,511.1 (495.0)**
JLI	1	283.6 (432.9)	257.1 (406.0)	466.6 (453.3)	3,375.2 (425.4)	519.4 (357.2)	4,549.9 (314.7)
	2	250.7 (401.4)	248.8 (402.8)	341.23 (533.4)	3,061.8 (505.1)	350.9 (505.4)	2,584.3 (306.5)
	3	1,081.6 (373.3)	298.2 (438.9)	247.3 (395.3)	2,552.1 (515.7)	**283.2 (529.9)**	**2,106.5 (390.2)**
	4	**205.7 (562.5)**	**2,140.7 (539.7)**	**231.7 (480.6)**	**2307 (372.8)**		
LYK	1	197.4 (433.8)	347.8 (479.6)	**191.8 (515.6)**	**1,810.9 (386.3)**	197.0 (530.3)	1,854.7 (487.9)
	2	172.7 (299.0)	251.7 (478.3)	202.1 (513.5)	1,914.5 (394.2)	**193.4 (555.7)**	**1,872.4 (467.6)**
	3	198.7 (397.3)	197.5 (675.8)	225.5 (605.4)	2,175.8 (462.2)	206.8 (683.2)	2,206.3 (576.4)
MHA	1	258.6 (699.4)	301.5 (703.2)	**249.3 (811.6)**	**1,852.2 (774.4)**	**231.3 (686.4)**	**1,924.4 (611.1)**
	2	324.3 (832.6)	270.6 (821.6)	**279.6 (803.8)**	**2,008.5 (817.7)**	**236.2 (604.1)**	**1,859.8 (616.2)**
	3	169.2 (595.7)	180.2 (646.4)				
	4	**181.2 (744.3)**	**2,696.9 (1,788.1)**				

Discussion

Three of four participants showed evidence that different response patterns do become part of a class that also contains visual stimuli. All the participants were trained to make two distinct response patterns (FR and DRL) to two different color stimuli. Then, all were trained to match those same stimuli to different form samples in a matching task. All participants received training on a mixed-schedule procedure to establish a history of making both response patterns to a common stimulus. Testing consisted of presenting the form stimuli and requiring participants to make one of the two defined responses to them, and "presenting" the responses as samples followed by a choice between the color and the form stimuli (in separate test sessions). In the initial test series, one participant (GNR) passed all tests, indicating that the different response patterns had joined into the class with the visual stimuli. Results for the remaining participants, however, were more variable. All three initially made only FR responses to the form stimuli in the form-response test. After subsequent training, two (JLI an MHA) eventually made differential responses consistent with class membership. On the response-color/form matching tests, one participant (MHA) made class-consistent choices from the outset. Another participant (JLI) made all class-inconsistent responses initially, but then made class-consistent responses after subsequent exposure to a color-form matching task. This task was the symmetrical version of Phase 2 training, and her responses were symmetrical from the first trial.

One participant, LYK, did not show evidence of the FR and DRL response patterns becoming part of a stimulus class. Her choices on the response-matching tasks were either completely class consistent or completely class inconsistent, and this changed for each type of test with each test session. LYK was subsequently given training on MTS with only visual stimuli and then given standard tests for reflexivity, symmetry, and transitivity. Her behavior on these tests showed the same pattern as that on response-inclusion tests—sometimes it was class consistent, and other times it was class inconsistent. Thus, LYK did not form equivalence classes even when trained in the typical manner.

The results of this study add to the growing body of literature investigating the status of class-consistent–defined responses in equivalence classes (Dymond & Barnes, 1994, 1995; Manabe et al., 1995; Shimizu, 2006; Urcuioli et al., 2006; Urcuioli et al., 2002; Urcuioli & Vasconcelos, 2008a, 2008b). Several previous studies also were successful in presenting responses as samples (e.g., Dymond & Barnes, 1994, 1995; Shimizu, 2006) to human participants using different procedures, and the results of this study complement and extend those data. In Shimizu (2006), the responses were different four-step sequences of computer-mouse manipulations, and training involved the use of written instructions as well as visual and auditory prompts. The participants were then required to make those responses when choosing comparisons in subsequent matching tasks. In test, a white square was presented and the participants were told to make one of the previously learned responses (i.e., mouse manipulations), and if they were correct, comparison choices would appear and they should respond to them as they did before. Our results extend the results of Shimizu by showing that similar results can be obtained using a task that does not require written or other visual cues in order to train the responses, thereby eliminating the possibility that those cues were responsible for facilitating stimulus class formation. In addition, the mixed-schedule procedure does not require any additional instructions at the time of testing because the participants are already familiar with the testing format.

In Dymond and Barnes (1994), three 3-member equivalence classes (A1, B1, C1; A2, B2, C2; A3, B3, C3) were established prior to training participants to make class-consistent responses to one member from each class. Participants were presented with a cue (the words *spacebar task*) on the computer screen followed by the presentation of either B1 or B2; participants were taught to either press the spacebar (e.g., to stimulus B2) or to refrain from pressing the spacebar (e.g., stimulus B1) and were given written feedback (i.e., "correct" or "wrong" appeared on the screen). In a second phase of the trial, the stimulus was re-presented and followed by a choice between the two B stimuli (i.e., identity matching). In a series of steps, the visual stimuli that occasioned the spacebar press (or the absence of a spacebar press) and the sample stimuli were faded out, such that in a final training phase, only the words *spacebar task* were presented, the participants were required to either respond or refrain from responding, and there was no cue indicating which was correct. If the participant behaved correctly (e.g., did not press the spacebar when "no response" was required), the written feedback still appeared and was followed by comparison presentation. If the participant behaved incorrectly (e.g., did not press the spacebar when a response was required), the word *incorrect* appeared, and it was also followed by comparison presentation. In test, the comparisons were the C stimuli from the previously established equivalence classes and accuracy was assessed by determining if participants chose a comparison that "matched" or "went with" whatever response (spacebar press or refrain from spacebar press) they had just made (i.e., C1 after "no spacebar press" and C2 after "spacebar press"). All the participants responded in a class-consistent manner.

Like Dymond and Barnes (1994), our study involved a motor response to a single visual stimulus (pressing a key on a keyboard versus tapping a computer touchscreen), but unlike Dymond and Barnes, our study involved two discrete response patterns across classes and participants, whereas for Dymond and Barnes, one "response" was *not* pressing the spacebar. This may be an important difference between the studies, as it was not reported what behavior the Dymond and Barnes' participants engaged in when they refrained from pressing the spacebar. Nonetheless, the results of Shimizu (2006), Dymond and Barnes, and the current study all provide convergent evidence supporting Sidman's (1994) assertion that defined responses can become members of equivalence classes (see also Dymond and Barnes, 1995, for an extension of their general procedure to relations other than equivalence).

Although the current results provide suggestive evidence supporting Sidman's (1994, 2000) hypothesis, there remain some open issues. For example, only one participant (of the three who showed evidence of class formation) made differential responses to the form stimuli when they were presented in test. Similar variability was found in Shimizu (2006). In that study, on all test trials, *any* response to a class-consistent comparison was sufficient to earn reinforcement, but participants were given the opportunity to make one of the defined responses. Four of the eight participants failed to make a class-consistent response on at least some of the test trials, even when they passed tests that involved response samples and visual comparisons. The variability observed in the current study may be related to specific aspects of the training procedure. In both Phase 1 training to make FR and DRL responses to the color stimuli and in Phase 3 mixed-schedule training, participants heard a switch cue when making an incorrect response pattern. For example, on a majority of trials, participants pressed the stimulus once, and when nothing happened, they pressed again. If they did not hear the switch cue, then they rapidly continued pressing the stimulus until it disappeared. If they heard the cue, they paused before making another press. In the response-form test, however, there was no switch cue, as either an FR or a DRL was sufficient to end the trial. Thus, these participants did in test what they had been trained to do: In the absence of the switch cue, they continued responding until the stimulus disappeared.

What is unclear, however, is what aspects of the additional training (if any) some participants received caused them to begin to emit differential responses. For JLI, responses became differential after both color-form matching exposure *and* class-consistent performance on the response-matching tests. It is unclear why that should be the case, as the response-matching tests would serve to strengthen the switch-cue guided behavior (since each trial started with the white square stimulus). For MHA, the answer is a little clearer. She too was given training on color-form matching, but this had no effect on subsequent form-response behavior (and she matched in a class-consistent manner on the other tests from the outset). Prior to her final test, she was given a refresher training session on color-response training without any reinforcement, in which accuracy was 100%. Thus, just prior to test, she experienced a session in which she made differential responses to the color stimuli, did not hear the switch cue, and was not given any reinforcement: conditions that were identical to those in test with the exception of which visual stimuli were presented.

One way to test the switch-cue hypothesis is to simply rearrange the order in which training and testing is conducted. For instance, training could be conducted in Phase 1 without a switch cue (such that the stimulus just remains on the screen until a DRL response is made). Then, form-color matching could be conducted, followed by the form-response test. Once that test is passed, participants would then be given mixed-schedule training and tested for response membership in the class. Doing so, and also testing for the full complement of emergent relations, will provide strong evidence for response membership in equivalence classes.

References

ADAMS, B. J., FIELDS, L., & VERHAVE, T. (1993). Effects of test order on intersubject variability during equivalence class formation. *The Psychological Record, 43*, 133–152.

BARNES, D., MCCULLAGH, P. D., & KEENAN, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *The Analysis of Verbal Behavior, 8*, 19–30.

BARROS, R. S., LIONELLO-DENOLF, K. M., DUBE, W. V., & MCILVANE, W. J. (2006). Equivalence class formation via identity matching to sample and simple discrimination with class-specific consequences. *Brazilian Journal of Behavior Analysis, 2*, 79–92.

BRAGA-KENYON, P. R., ANDRADE, M., AHEARN, W. H., & SIDMAN, M. (2006). *Inclusion of defined responses in equivalence relations: A systematic replication of Manabe et al. (1995).* Paper presented at the 32nd annual meeting of the Association of Behavior Anaysis, Atlanta, GA.

CARR, D., WILKINSON, K. M., BLACKMAN, D., & MCILVANE, W. J. (2000). Equivalence classes in individuals with minimal verbal repertoires. *Journal of the Experimental Analysis of Behavior, 74*, 101–114. doi:10.1901/jeab.2000.74-101

DE ROSE, J. C. (1996). Controlling factors in conditional discriminations and tests of equivalence. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 253–277). Amsterdam, The Netherlands: Elsevier.

DUBE, W. V., & MCILVANE, W. J. (1995). Stimulus-reinforcer relations and emergent matching to sample. *The Psychological Record, 45*, 591–612.

DUBE, W. V., MCILVANE, W. J., MACKAY, H. A., & STODDARD, L. T. (1987). Stimulus class membership via stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior, 47*, 159–175. doi:10.1901/jeab.1987.47-159

DUBE, W. V., MCILVANE, W. J., MAGUIRE, R. A., MACKAY, H. A., & STODDARD, L. T. (1989). Stimulus class formation and stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior, 51*, 65–76. doi:10.1901/jeab.1989.51-65

DYMOND, S., & BARNES, D. (1994). A transfer of self-discrimination response functions through equivalence relations. *Journal of the Experimental Analysis of Behavior, 62*, 251–267. doi:10.1901/jeab.1994.62-251

DYMOND, S., & BARNES, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more than, and less than. *Journal of the Experimental Analysis of Behavior, 64*, 163–184. doi:10.1901/jeab.1995.64-163

JOSEPH, B., OVERMIER, J. B., & THOMPSON, T. (1997). Food and nonfood related differential outcomes in equivalence learning by adults with Prader-Willi syndrome. *American Journal on Mental Retardation, 101*, 374–386.

LIONELLO-DENOLF, K. M., & MCILVANE, W. J. (2003). Rebirth of the Shriver Automated Teaching Laboratory. *Experimental Analysis of Human Behavior Bulletin, 21*, 12–17.

LIONELLO-DENOLF, K. M., & URCUIOLI, P. J. (2003). A procedure for generating differential "sample" stimuli without different exteroceptive stimuli. *Journal of the Experimental Analysis of Behavior, 79*, 21–35. doi:10.1901/jeab.2003.79-21

MANABE, K., KAWASHIMA, T., & STADDON, J. E. R. (1995). Differential vocalization in budgerigars: Towards an experimental analysis of naming. *Journal of the Experimental Analysis of Behavior, 63*, 111–126. doi:10.1901/jeab.1995.63-111

MCILVANE, W. J., DUBE. W. V., KLEDARAS, J. B., DE ROSE, J. C., & STODDARD, L. T. (1992). Stimulus-reinforcer relations and conditional discrimination. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 43–67). Reno, NV: Context Press.

SAUNDERS, K. J., & WILLIAMS, D. C. (1998). Do parakeets exhibit derived stimulus control? Some thoughts on experimental control procedures. *Journal of the Experimental Analysis of Behavior, 70*, 321–324. doi:10.1901/jeab.1998.70-321

SCHENK, J. J. (1994). Emergent relations of equivalence generated by outcome-specific consequences in conditional discrimination. *The Psychological Record, 44*, 537–558.

SHIMIZU, H. (2006). Testing response-stimulus equivalence relations using differential responses as samples. *Journal of the Experimental Analysis of Behavior, 86*, 239–251. doi:10.1901/jeab.2006.04-03

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

SIDMAN, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*, 127–146. doi:10.1901/jeab.2000.74-127

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. doi:10.1901/jeab.1982.37-5

URCUIOLI, P. J., LIONELLO-DENOLF, K. M., MICHALEK, S., & VASCONCELOS, M. (2006). Some tests of response membership in acquired equivalence classes. *Journal of the Experimental Analysis of Behavior, 86*, 81–107. doi:10.1901/jeab.2006.52-05

URCUIOLI, P. J., PIERCE, J. N., LIONELLO-DENOLF, K. M., FRIEDRICH, A., FETTERMAN, J. G., & GREEN, C. (2002). The development of differential sample behavior in pigeons. *Journal of the Experimental Analysis of Behavior, 78*, 409–432. doi:10.1901/jeab.2002.78-409

URCUIOLI, P., & VASCONCELOS, M. (2008a). Effects of within-class differences in sample responding on acquired sample equivalence. *Journal of the Experimental Analysis of Behavior, 89*, 341–358. doi:10.1901/jeab.2008-89-341

URCUIOLI, P., & VASCONCELOS, M. (2008b). On the origins of emergent differential sample behavior. *Journal of the Experimental Analysis of Behavior, 90*, 61–80. doi:10.1901/jeab.2008.90-61