2021

# Expansion of Sidman's Theory: The Inclusion of Prompt Stimuli in Equivalence Classes

Simone K. Palmer
*Simmons University*

R. W. Maguire
*Simmons University*

Karen M. Lionello-DeNolf
*Assumption University*, k.lionellodenolf@assumption.edu

Paula Braga-Kenyon
*Northeastern University*



## Recommended Citation

Palmer, S.K., Maguire, R.W., Lionello-DeNolf, K.M. and Braga-Kenyon, P. (2021), Expansion of Sidman's theory: The inclusion of prompt stimuli in equivalence classes. Jrnl Exper Analysis Behavior, 115: 255-271. https://doi.org/10.1002/jeab.655.

**Expansion of Sidman's Theory: The Inclusion of Prompt Stimuli in Equivalence Classes**

Simone K. Palmer[1, 2], R. W. Maguire[1], Karen M. Lionello-DeNolf[3], and Paula Braga-Kenyon[4, 5]

[1]Simmons University

[2]Crossroads School

[3]Assumption University

[4]Kadiant

[5]Northeastern University

**Author Note**

Simone K. Palmer, Department of Behavior Analysis, Simmons University, Boston, MA.

This manuscript was completed in partial fulfillment of the requirements for the doctoral degree in Behavior Analysis at Simmons University. Thank you to Simmons University, Assumption University and Crossroads School for the support throughout this study. Simone Palmer is currently affiliated with Bierman ABA, Inc.

Correspondence concerning this document should be addressed to Simone KFVB Palmer, Department of Behavior Analysis, Simmons University, 300 The Fenway, Boston MA 02115. Email: si_vilasboas@yahoo.com.br

**Abstract**

Stimulus equivalence is defined as the ability to relate stimuli in novel ways after training in which not all of the stimuli had been directly linked to one another. Sidman (2000) suggested all elements of conditional discrimination training contingencies that result in equivalence potentially become class members. Research has demonstrated the inclusion of samples, comparisons, responses, and reinforcers in equivalence classes. Given the evidence that all elements of a conditional discrimination become part of the class, the purpose of this study was to determine if class-specific prompts would also enter into their relevant equivalence classes. Experiment 1 investigated the inclusion of prompts in an equivalence class using abstract stimuli with neurotypical students enrolled in higher education courses. Experiment 2 systematically replicated Experiment 1 using meaningful stimuli and individuals diagnosed with autism spectrum disorder. The results of both experiments demonstrated that class-specific prompts became part of equivalence classes with the other positive elements of the contingency. The results are discussed in terms of class expansion and the potential impact on equivalence-based instruction.

*Keywords:* stimulus equivalence, stimulus prompt, time-delay, equivalence-based instruction, conditional discrimination, college students, autism

**Expansion of Sidman's Theory: The Inclusion of Prompt Stimuli in Equivalence Classes**

Stimulus equivalence is defined as the ability to relate stimuli in novel ways after direct training of at least two relations in which not all of the stimuli had been linked to one another. Sidman (2000) stated that although equivalence relations originate in the reinforcement contingency, they also represent an outcome that extends beyond the establishment of the analytic units (i.e., two, three, four, five-term contingencies, etc.) to the emergence of new, untrained analytic units. Often, training is conducted in a conditional discrimination format (e.g., Sidman & Tailby, 1982), but evidence for equivalence has also been demonstrated after training on simple discriminations (e.g., Debert et al., 2007; Pilgrim, 2019). After discrimination training, for emergent relations to be considered equivalence classes, those relations must have the following three properties: reflexivity, symmetry, and transitivity (Sidman, 1990, 1994, 2000; Sidman & Tailby, 1982). In the context of conditional discrimination, reflexivity refers to matching a stimulus to itself (i.e., identity matching) without direct training or reinforcement. An example is matching stimulus A to stimulus A, stimulus B to stimulus B, and stimulus C to stimulus C. Symmetry refers to the reversibility of a previously trained conditional discrimination in the absence of direct training and reinforcement (e.g., sample-comparison reversibility or sample-comparison bi-directionality). For example, after direct training of the relation between A and B stimuli, the relation between B and A stimuli is demonstrated with no further training. Transitivity refers to the emergence of relations between stimuli that had not been directly related through training, but each had been related to a third stimulus; it is demonstrated by individuals relating samples from one trained conditional relation to comparisons from another (Sidman & Tailby, 1982). For example, after direct training of relations A to B and B to C, transitivity is shown by the emergence of the relation A to C.

Stimulus equivalence has been reliably demonstrated in both typically developing children (e.g., Sidman & Tailby, 1982) and adults (e.g., Dube et al., 1993; Johnson et al., 2014) and in individuals with developmental and intellectual disability (e.g., LeBlanc et al., 2003). The wide replicability of equivalence has generated tremendous excitement in the literature for both theoretical and applied reasons. In terms of theory, equivalence research has raised interesting questions related to verbal behavior (e.g., Guinther & Dougher, 2015; Sidman, 1986). In terms of application, designing teaching programs that use conditional discrimination procedures to promote emergent relations has the potential to decrease teaching time while simultaneously increasing what is learned (e.g., Albright et al., 2015; Fienup et al. 2015). In this paper, we report experiments relevant to both theory and application.

Sidman (1994, 2000) proposed that every element involved in the reinforcement contingency could become a member of an equivalence class as a result of direct training of the baseline relations depending on how teaching is structured. According to this idea, reinforcers and defined responses should become part of an equivalence class in addition to the sample and comparison stimuli when they are specific to that class. Empirical studies have confirmed Sidman's proposal (e.g., Dube et al., 1989; Johnson et al., 2014; Lionello-DeNolf & Braga-Kenyon, 2013; Minster et al., 2006; Shimizu, 2006).

Testing whether defined responses become part of an equivalence class is challenging because the test procedure needs to "present" the response as a sample stimulus without an additional antecedent stimulus (Sidman, 1994). Lionello-DeNolf and Urcuioli (2003) developed a procedure in which response patterns functioned as samples in a matching task without the need for differential occasioning stimuli. At the beginning of every trial, pigeons were presented a white key, and one of two response patterns (pecking fast or pecking slow) was required. There was no additional stimulus to cue which response pattern was correct on a given trial. If the pigeon made an incorrect response, the key went dark and the trial was repeated until the pigeon made the correct response pattern. Correct response patterns were then followed by a choice between two visual comparisons, and which comparison was correct depended on the response pattern that produced the comparisons. In other words, Comparison 1 was correct after pecking fast and Comparison 2 was correct after pecking slow. All of the pigeons learned to choose the correct comparison after response pattern samples. Subsequently, Shimizu (2006) and Lionello-DeNolf and Braga-Kenyon (2013) used this procedure with typically developing humans to demonstrate that that the different response patterns became members of equivalence classes with the samples and comparisons from the conditional discrimination tasks.

Testing Sidman's (2000) proposal that reinforcers become part of the equivalence class also requires a modification to the typical matching-to-sample (MTS) procedure. In the typical procedure, MTS is used to teach three or more relations (e.g., A1 matched to B1, A2 matched to B2, A3 matched to B3 where "A" refers to sample stimuli and "B" refers to comparison) stimuli, and the same reinforcers are used for correct matches across trials. After training the A–B relations, learners may then be taught B-C relations using the same reinforcer(s). In this situation, according to Sidman (2000), the reinforcer cannot become part of the equivalence class because the same reinforcers are used across three different relations and so they must "drop out" of the relation. To determine if reinforcers could become part of the class, a different reinforcer must be used for each relation taught. For example, reinforcer 1 (R1) would be used on A1-B1 and B1-C1 trials, R2 would be used on A2-B2 and B2-C2 trials, and R3 would be used on A3-B3 and B3-C3 trials. After training, the reinforcer would be presented as the sample stimulus, followed by the A, B, or C stimuli as comparison choices.

Dube and colleagues (1989) were the first to provide evidence that class-specific reinforcers can become members of equivalence classes. Two participants with intellectual disabilities were trained on identity MTS with class-specific reinforcers (i.e., A1-A1, B1-B1, C1-C1, and D1-D1 followed by R1 and A2-A2, B2-B2, C2-C2, and D2-D2 followed by R2). Next, the participants were trained on AB matching in which selections of B1 and B2 conditionally upon A1 and A2 were followed by R1 and R2, respectively. A novel stimulus (D) was then introduced through identity matching trials that were intermixed with AB and BC trials. A series of probe trials tested for the emergence of the two stimulus classes (i.e., A1-B1-C1-D1 and A2-B2-C2-D2). Both participants had high accuracies on the probe trials. These results suggest that stimulus-reinforcer relations served as the basis for stimulus class membership. Subsequent studies have also shown that class-specific reinforcers become part of the stimulus class (e.g., Johnson et al., 2014; Minster et al., 2006).

Confirmation of Sidman's proposal regarding the inclusion of reinforcers and responses in equivalence classes has applied importance because it suggests that other elements of the MTS procedure, such as prompts, may also become class members. Prompts are supplementary antecedent stimuli presented to increase the probability of a correct response (MacDuff et al., 2001; Touchette & Howard, 1984). This supplementary antecedent stimulus initially controls the target response, but is not functionally related to the task and may not be related to the discriminative stimulus that will eventually evoke the behavior (Touchette & Howard, 1984). People often learn to perform both simple and conditional discriminations without the need for prompting, but individuals with developmental or learning disabilities may require prompts o learn these relations (MacDuff et al., 2001). In this situation, stimulus control needs to be gradually transferred to the target antecedents by the systematic fading of the prompts (Green, 2001; MacDuff et al., 2001; Touchette, 1971; Touchette & Howard, 1984). Studies have reported issues related to the use of prompts during instruction, such as prompt dependency, faulty transfer of stimulus control, and restricted stimulus control (e.g., Green, 2001; MacDuff et al., 2001; Rincover, 1978). If prompts do become part of the equivalence class, then use of the same prompt across different classes may lead to class merger, potentially impeding the acquisition of discrimination across and within stimuli for some learners. For example, when teaching mathematics, a learner may be directly taught to choose comparison numerals 1, 2, and 3 given the corresponding quantities of dots as samples (A-B), and to choose the comparison printed words ONE, TWO, and THREE given the numerals as samples (B-C). After teaching, the learner may be able to match the printed words to the corresponding quantities (A-C and C-A matching) without further training. However, other outcomes are possible and frequently occur. For example, the learner may fail to acquire A-B or B-C baseline relations, or the learner may not show emergence of A-C/C-A matching. One reason for these unfavorable outcomes could be the use of a common prompt (e.g., a finger point or a gesture to the correct response) when teaching each baseline relation. In other words, if the prompt stimulus also becomes part of the equivalence class, use of a common prompt could prevent the separation of trained conditional relations into distinct equivalence classes.

The current study sought to test the hypothesis that class-specific prompts may enter equivalence classes. In Experiment 1, typically developing adults were trained on a series of arbitrary conditional relations with class-specific prompts using procedures similar to Dube et al. (1989) who investigated the inclusion of class-specific reinforcers in equivalence classes. Experiment 2 was a systematic replication of Experiment 1 in which children with autism were trained on conditional relations using class-specific prompts and meaningful (i.e., educationally relevant) stimuli.

# General Method

## Experimental Design

A pretest/posttest design (Sidman, 1971) was used in both experiments to compare the number of experimenter-defined classes prior to and following training as demonstrated by sorting and MTS test trial data.

## Participants

Experiment 1 employed typical adults. Experiment 2 employed children with autism spectrum disorders (ASD).

## Setting and Materials

Sessions were conducted in a quiet room within the location in which the participants were recruited. Experiment 1 used arbitrary stimuli and Experiment 2 used meaningful stimuli (Figure 1). The sorting task used 3.5 × 3.5-inch laminated cards, one copy of each stimulus across all three potential classes (Figure 1). An iPhone X© (https://www.apple.com) was used to photograph and video record permanent products during the sorting tasks.

Matching-to-sample trials were conducted on a computer with a touchscreen monitor (i.e., HP Laptop 15-bs015dx, Intel® Core™ i5-7200U CPU, http://www.hp.com) using customized software (*Matching to Sample Procedure Software*; Boldrin & Debert, 2017) that controlled presentation of all stimuli, consequences, intertrial intervals (1.5 s), and automatically recorded all data. The program also recorded the sample and comparison stimuli, comparison-stimulus locations for each trial, the timestamps for responses to the sample and comparison stimuli (i.e., latency to respond), and the date and time of each session.

**Dependent Variable and Measurement**

Two different types of procedures were used: sorting and MTS. Table 1 depicts the experimental conditions across both experiments. For both procedures, the experimenter sat to the side and slightly behind the participants so that the participants were not be able to observe the experimenter, but the experimenter was able to see the screen and the participants' responses. Sorting was included as (1) verification of the absence of pre-experimental relations between the stimuli used to establish equivalence classes and (2) an additional measure of class formation and expansion.

***Pretest: Sorting***

A sorting task served as the pretest, and the dependent variable was the number of stimulus cards sorted to form each of the experimenter-defined classes. A class-consistent response was defined by sorting stimuli according to the experimenter-defined classes. A stack of fifteen 3.5 × 3.5-inch laminated cards (i.e., one copy of each stimulus across all three potential classes) was randomly shuffled and placed in front of the participants. The participants were instructed to, "put the card into groups however you think they should go." Upon completion, the experimenter photographed the stimuli as the participants had grouped them. There were no programmed consequences or prompting delivered during sorting tasks.

***Training and Testing: Matching-to-sample (MTS)***

MTS training and testing established and assessed the acquisition and emergence of conditional discriminations. The dependent variables were the percent independent, class-consistent responses during training and testing for derived relations and the latency to respond to the comparison stimuli in s. A class-consistent response was defined as choosing a comparison stimulus that was related to the sample stimulus according to the experimenter-defined classes.

The total number of class-consistent responses was summed and the percentage calculated by the software. Latency to respond was defined as the time between the presentation of the comparison array and touching one of the comparison stimuli. Latency to both correct and incorrect responses was measured and data collected by the software. Each training and test session consisted of 18-trials. Figure 2 provides a schematic detailing the trained relations (e.g., A-B, A-C, and D-D) and tests for the emergence of untrained relations that would indicate equivalence class formation (e.g., B-A, C-A, B-C, C-B, D-A, A-D, D-B, B-D, D-C, C-D, A-A, B-B, C-C).

The MTS procedure involved a visual-visual simultaneous presentation of the stimuli with a non-differential observing response to the sample. At the start of each trial, a sample stimulus, a non-representative symbol in black on a white background measuring 2.5 × 2.5 inches, was presented at the top center of the screen. A touch to the sample produced the comparison stimuli. Three comparison stimuli, measuring 2.5 × 2.5 inches, appeared on the bottom of the screen side-by side, separated by one inch. The sample stimulus was present for the entire trial (i.e., simultaneous MTS). The participant selected a comparison either touching it on the screen or by using the touchpad and clicking on the right button.

During training, class-consistent responses were followed by a smiley face and a clapping/applause sound. Class-inconsistent responses resulted in the presentation of a brown "X" across the screen, accompanied by an auditory beeping sound. Test sessions were conducted without any programmed consequences.

***Prompting***

Stimulus prompts used during conditional discrimination training consisted of class-specific colors (i.e., blue, red, and yellow). These were displayed as the background of the sample and comparison stimuli. During training trials, the background color of the sample and the S+ comparison stimulus was the same, serving as a prompt, and the background colors for the S- comparisons were different from the S+ and each other. Although the color prompts can be considered distinct from the form used in MTS training, the presentation of the sample and comparison stimuli with the prompt could also be considered a compound stimulus. For example, from this perspective, A1 with red might be the sample, followed by a choice between three compound comparisons: B1 with red, B2 with blue, and B3 with yellow. During training, the color element of the compound (i.e., the prompt) was systematically faded out.

A delayed cue fading procedure was used to transfer control of responding from the color to the comparison stimulus. Experiment 1 included three fading steps. Initially, the colors were presented simultaneously with the onset of the comparison stimuli (e.g., 0-second delay). Contingent on class-consistent responding, the onset of the colors was delayed by 3-seconds after the presentation of the comparison stimuli. Finally, the color prompts were faded out entirely, again contingent on class-consistent responding at the 3-second delay. The delayed cue procedure in Experiment 2 included four fading steps: Presentation of the class-specific colors with the comparison stimuli occurred simultaneously, after a 2-second delay, after a 4-second delay, and finally with no presentation of class-specific color stimuli.

Criteria for advancement from one prompt level to the next were 89% or better class-consistent responding across three consecutive 18-trial sessions at a specific prompt level. Once the criteria were met, participants moved to the next fading step. If four or more errors occurred during two consecutive sessions, the previous cue level was reintroduced for all the relations. Relations were considered acquired once participants demonstrated 89% or better class-consistent responding across three consecutive 18-trial sessions without prompting.

**Phase 1 – Training.** During this phase, the first three baseline relations (i.e., A1-B1, A2-B2, A3-B3) were trained with class-specific prompts (i.e., E1, E2, E3, respectively).

**Phase 2 – Test.** Test trials assessed the emergence of symmetry (i.e., B1-A1; B2-A2; B3-A3) and if the prompt (E), presented as a sample and comparison, controlled class consistent-responding involving the A and B stimuli (i.e., E1-A1; E2-A2; E3-A3; E1-B1; E2-B2; E3-B3; A1-E1; A2-E2; A3-E3; B1-E1; B2-E2; B3-E3). During this phase, no programmed consequences occurred and testing occurred in one session.

**Phase 3 – Training.** During this phase, the second baseline relations (i.e., A1-C1, A2-C2, A3-C3) were trained with class-specific prompts (i.e., E1, E2, E3, respectively).

**Phase 4 – Test.** Test trials assessed the emergence of symmetry (i.e., C1-A1; C2-A2; C3-A3) and if the prompt (E), presented as a sample and comparison, controlled class consistent-responding involving the C and E stimuli (i.e., E1-C1; E2-C2; E3-C3; C1-E1; C2-E2; C3-E3). During this phase, no programmed consequences occurred and testing occurred in one session.

**Phase 5 – Test.** Untrained relations were tested to assess the emergence of transitivity (i.e., B1-C1; B2-C2; B3-C3; C1-B1; C2-B2; C3-B3). During this phase, no programmed consequences occurred and testing occurred in one session.

**Phase 6 – Training.** Participants were trained on identity matching-to-sample with novel stimuli (i.e., D1-D1; D2-D2; D3-D3) using class-specific prompts. Training procedures were identical to Phase 1. Mastery criteria were independent, class-consistent responding at 89% or above in one session.

**Phase 7 – Test.** Test trials determined if the prompts (E), presented as samples and comparisons, were related to the D stimuli, consistent with identity matching-to-sample training

(Phase 6). During this phase, no programmed consequences occurred and testing occurred in one session.

**Phase 8 – Test.** Participants were tested for all potential emergent relations between the D, A, B, and C stimuli (D1-A1; D2-A2; D3-A3; D1-B1; D2-B2; D3-B3; D1-C1; D2-C2; D3-C3; A1-D1; A2-D2; A3-D3; B1-D1; B2-D2; B3-D3; C1-D1; C2-D2; C3-D3). Class-consistent responding on these trials would indicate stimulus class expansion as a result of the prompt (i.e., color) becoming a member of the stimulus class.

**Phase 9 – Test.** This phase assessed reflexivity (i.e., A1-A1; A2-A2; A3-A3; B1-B1; B2-B2; B3-B3; C1-C1; C2-C2; C3-C3). During this phase, no programmed consequences occurred and testing occurred in one session.

***Posttest: Sorting***

After completing all MTS phases, participants were presented with the sorting task again. The procedures were identical to those in pretesting.

**Interobserver Agreement**

All sorting tasks were video recorded and photographs of each participant's pretest and posttest sorting were taken. Interobserver agreement (IOA) was assessed by two independent observers on the sorting tasks by tabulating the responses into a table as described by Fields and colleagues (2014). Agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying the result by 100. Reliability was assessed for 50% of the pretest and posttest data. Because all MTS data were recorded by the computer program, IOA data were not required.

**Procedural Integrity**

Data were collected on the experimenter's correct implementation of the procedures in each phase of this study using a procedural integrity checklist that included the steps required for each phase. To calculate procedural integrity, the number of accurately completed steps was divided by the total number of steps, multiplied by 100. Procedural integrity was assessed on 50% of sessions.

# Experiment 1

## Method

### *Participants*

The participants were 10 neurotypical adults enrolled in a higher education course; five participants were enrolled in an undergraduate course and five participants were enrolled in the first year (i.e., first or second semester) of a master's program in applied behavior analysis. Only students who were in the beginning stages of the coursework in applied behavior analysis were selected to participate to ensure that they had not been exposed to the concepts of stimulus control and equivalence classes before this study. The undergraduate students received course credit for participating, but the master's students did not. Participation was voluntary and the experimenter did not have any employment relationship with the participants. Table 2 depicts the participants' demographics at the beginning of the experiment.

Prior to the beginning of the study, each participant received a description of the study presented in written and spoken form and an informed consent form. The experimenter made clear that participation was voluntary and consent could be withdrawn at any time.

### *Procedure*

The procedures in Experiment 1 were as described in the General Method section.

**Results**

Across participants, session duration ranged from 1.02 to 1.3 hours and included all experimental tasks, excluding the time to review and sign the informed consent document. Results from the pre- and posttest are shown in Table 3 and results of the MTS sessions are shown in Table 4.

***Pretest: Sorting***

Data from the sorting tasks were examined to determine the number of stimuli participants placed into piles consistent with experiment-defined classes. Sorting data are represented by "strings" of three columns; each string represents a pile of cards. Each column in the string represents the number of stimuli from classes 1, 2, and 3, respectively, that the participant placed in that pile. For example, the string 5-0-0 indicates that the participant placed five cards from class 1, zero cards from class 2, and zero cards from class 3 in that pile. Thus, if participants sorted the cards according to the experimenter-defined classes, there would be three piles represented by the following strings: 5-0-0, 0-5-0, and 0-0-5 in any order. During the pretest (Table 3, upper portion), participants sorted cards into between two and five piles, none of which corresponded to experimenter-defined classes.

***Training and Testing: Matching-to-sample***

Accuracy in MTS sessions was calculated as the percent of independent (i.e., without prompting) class-consistent responding across all three classes and averaged across three consecutive sessions. All participants acquired the three trained baseline relations to criterion, and overall accuracy ranged from 90–100% for the A-B, A-C, and D-D relations, respectively.

Seven of the 10 participants passed tests assessing whether the prompt stimuli (i.e., colors) would function as samples and comparisons in the matching task. For these seven, class-consistent responding ranged from 94–100% across all tested relations.

For the three participants who did not pass the test, class-consistent responding ranged from 33–78% on the first prompt test (relation A-E). On subsequent tests, P10 displayed criterion-level class-consistent responding for all relations. For P5 and P7, class-consistent responding ranged from 28–83% for relations A-E, B-E, E-A, and E-B and was 100% for all subsequent tests. Thus, although class-consistent responding was not evident at the outset of testing for these three participants, it did emerge in further testing.

Criterion for passing tests for emergent relations was 89% correct. All participants passed both symmetry tests (i.e., B-A and C-A) and nine passed both transitivity tests (i.e., B-C and C-B). The exception was P10, who had an accuracy of 39% on transitivity tests. In addition, nine participants passed all equivalence tests (i.e., A-D, B-D, C-D, D-A, D-B, and D-C) with P10 being the one who did not. Notably, P10 was one of the participants who failed at least one prompt test. Due to a programming error, P1, P2, and P3 did not receive B-D, C-D, D-B, and D-C tests and P4 did not receive the D-A test. All participants passed equivalence (i.e., class expansion) tests except for P10, whose class-consistent responding was 100% for C-D and ranged from 28–83% for the remaining relations. All participants passed all the reflexivity tests with 100% class-consistent responding.

There were differences in response latencies to the comparison stimuli on correct and incorrect trials. On correct trials, latencies were longer on independent training trials than prompted trials, and the opposite was observed on incorrect training trials. In addition, for correct trials, latencies were shorter on symmetry test trials that transitivity and class expansion trials, but the opposite was observed on incorrect trials. Finally, for correct trials, latencies on prompt trials were shorter than those for incorrect trials. Detailed data on response latencies for training and testing trials is available upon request.

***Posttest: Sorting***

During the sorting posttest (Table 3, lower portion), nine of 10 participants sorted the cards into the three experimenter-defined classes, indicated by piles with the following strings: 5 0 0, 0 5 0, and 0 0 5 (in any order). The exception was P10 who placed the cards into three piles in a manner inconsistent with the experimenter-defined classes and different from how the cards were sorted in the pretest. This result can be contrasted with that of the pretest in which only four participants separated the cards into three piles, none of which were consistent with experimenter-defined relations.

**Discussion**

The results of Experiment 1 demonstrate the formation of three five-member equivalence classes and document the inclusion of prompts into the respective classes. Participants were trained on A-B and A-C relations with class-specific prompts, and were able to match the stimuli according to symmetry and transitivity. In addition, when the prompts were presented as samples and/or comparisons, participants generally matched them in a manner consistent with the established equivalence classes. Then, D-D MTS was taught using the class-consistent prompts, and this training was sufficient to support class expansion: Participants generally matched the D stimulus to the A, B, and C stimuli when presented as both sample and comparison stimuli with the only element linking them in training being the class-consistent prompt. Thus, Experiment 1 adds strong evidence in support of Sidman’s (2000) theory that all positive elements in a contingency enter an equivalence class. Prior to the current experiment, sample and comparison stimuli, responses, and reinforcers had been shown to be included in equivalence classes (e.g., Dube et al., 1993; Dube et al., 1989; Johnson et al., 2014; Lionello-DeNolf & Braga-Kenyon, 2013; Minster et al., 2006). Our study demonstrates that prompt stimuli, used to occasion the correct response during instruction, may also be included in equivalence classes.

The results also indicate that latency to respond to the comparison stimuli differed across trial types which is consistent with the stimulus equivalence literature (e.g., Tomanari et al., 2006). Mean latency to respond to a comparison increased from the directly trained stimulus-stimulus relations to the tests for emergent relations, especially on transitivity and class expansion tests for a majority of participants. Interestingly, latencies on trials in which the prompt was presented as a sample or comparison stimulus were similar to latencies on training trials, which provides further evidence that the prompt stimuli were part of the equivalence class.

The results of this study have potential implications for instruction in special education settings. Specifically, use of a common prompt in MTS instruction may lead to class merger, impacting the learner's ability to acquire discriminations. In applied settings, a common prompt (e.g., a finger point) is typically used in teaching. For equivalence classes to form under these conditions, the common prompt (as well as common responses and reinforcers) would have to "drop out" of the class. To the extent that an element common across classes does not drop out, it is possible that equivalence class formation may be prevented.

Another important difference between Experiment 1 and the application of equivalence-based instruction in special education settings is the type of stimulus used in teaching. In Experiment 1, non-representative stimuli were used to determine if prompts would enter equivalence classes. However, in educational settings in which prompts are typically used (e.g., with individuals with intellectual disability), the stimuli are educationally relevant and the learner may have had some prior exposure to them. Thus, Experiment 2 was conducted as a systematic replication of Experiment 1 using educationally relevant stimuli in a special education setting with individuals diagnosed with autism spectrum disorder (ASD).

# Experiment 2

## Method

### *Participants*

Participants were four individuals diagnosed with ASD who attended a private school for individuals with ASD or other developmental disabilities. Table 2 shows their demographics at the beginning of the experiment. All had prior exposure to sorting tasks, MTS tasks, and errorless instruction. During the sessions, their individualized behavior guidelines (i.e., responses to challenging behaviors, antecedent strategies prescribed) were implemented as needed.

The inclusion criteria included attending skills, independent use of a computer, reading skills, previous exposure to testing in extinction (e.g., participation in standardized testing), and demonstration of sorting and matching skills. Participants who demonstrated class formation during the pretest sorting task were excluded. If participants failed to acquire any of the baseline relations after three attempts to fade prompts, their participation ended.

Prior to the study, each participant (if applicable) and/or their guardian received a written and spoken description of the study in an informed assent and/or consent form, respectively. The experimenter made it clear that participation was voluntary, and consent could be withdrawn at any time.

### *Procedure*

**Preference Assessment.** A preference assessment (adapted from the Reinforcer Assessment for Individuals with Severe Disabilities [RAISD]; Fisher et al., 1996) was conducted to identify self-reported preferred items for each participant. The highest rated items were delivered to the participants at the end of each session regardless of performance.

**Pre-Experiment Sorting Tasks.** Prior to the sorting pretest, two sorting tasks were given to the participants using non-experimental stimuli to verify sorting skills. Failure to demonstrate sorting skills during at least one of the pre-experiment tests ended their participation. The first sorting task consisted of arranging colors into three categories based on physical identity. The stimuli were 15 2 × 2-inch index cards of three different colors (i.e., five pink, five blue, and five orange cards). They were presented to participants in a random pile followed by the instruction to “put them in groups.”

The second sorting task consisted of sorting conceptually related arbitrary stimuli. The stimuli were 15 3.5 × 5-inch index cards containing five pictures of different animals, five pictures of different edibles, and five pictures of different clothing items. The sorting task was conducted in the same way as the color-sorting task.

**Pretest, training, test session and posttest sessions.** Procedures in Experiment 2 were described in the General Method Section.

## Results

### *Pretest: Sorting*

Table 5 depicts the results from the pre-experiment sorting tasks with familiar stimuli based on identity (upper potion) and categories (lower portion). All four participants correctly sorted all stimuli. Table 6 (upper portion) depicts the results from the pretest sorting task with the experimental stimuli. The participants sorted the experimental stimuli into three to six piles, and none sorted according to experimenter-defined classes.

### *Training and Testing: Matching-to-sample*

MTS sessions took place over multiple days in which one or more sessions were conducted per day. Participation ranged from 24–52 days. Results from the MTS training and test sessions are depicted in Table 7. P11 failed to acquire the first trained relation (A-B); after three attempts to fade out the prompt, his participation ended, and his data are not reported. The remaining participants acquired the baseline relations in an average of 7.25 (range, 7–12), 8.25 (range, 8–17), and 6 (range, 4–13) sessions, for the A-B, A-C, and D-D relations, respectively. Overall accuracy ranged from 98–100% across the three trained relations.

On the prompt tests, P14 passed all tests with 100% class-consistent responding. Mean class-consistent responding on prompt tests by P13 was 94.38% (range: 83–100%) and all but one test was passed; the only test below criterion was C-E. By contrast, P12 showed class-consistent responding on prompt tests involving stimulus sets A, B, and C, but not D. On the former, class-consistent responding averaged 98% (range: 94–10%), but on prompt tests involving class expansion (D-E and E-D), class-consistent responding was 22%.

All participants passed both symmetry tests (i.e., B-A and C-A) and both transitivity tests (i.e., B-C and C-B). Class-consistent responding on symmetry tests was 100% with two exceptions: 94% on B-A and C-A by P12. On transitivity tests, class consistent responding was 100% with two exceptions: 94% on B-C and C-B by P12. However, only P13 and P14 passed equivalence tests (i.e., those involving the D stimuli that tested for class expansion). For these two participants, class-consistent responding on equivalence tests was 100% with two exceptions: 94% on B-D and D-C by P13. By contrast, P12 failed all equivalence/class-expansion tests: Class-consistent responding ranged from 17–33%. Finally, all participants passed all reflexivity tests with 100% class-consistent responding.

In general, response latencies during training were shorter on prompted than independent trials. In test sessions, latencies were shortest on reflexivity trials, followed by prompt tests, and symmetry. Latencies were the longest on transitivity and class expansion tests. However, there was greater variation in response latencies as compared to the adult participants in Experiment 1. Detailed analysis of latency data is available upon request.

***Posttest: Sorting***

During the sorting posttest (Table 6, lower portion), only P13 and P14 sorted the cards into the three experimenter-defined classes. On the pretest, P13 had sorted the cards into two piles, and P14 had sorted them into three piles that were not consistent with experimenter-defined classes. P12 only sorted the cards into one of the three experimenter-defined classes (12 (1) in Table 6). For the other two piles, P12 placed D1 with the class 3 stimuli and D3 with the class 1 stimuli. A second posttest was then conducted (12 (2) in Table 6) on a separate day to ensure the stimulus-stimulus relations were not acquired outside of the experimental context. The results were identical to those of the first posttest.

Because P12 failed tests for equivalence (A-D, B-D, C-D, D-A, D-B, and D-C) and the posttest sorting, retraining on the baseline relations was conducted, followed by all tests for emergent relations except reflexivity (on which class-consistent responding had been 100%). The results of the retraining and additional tests for emergent relations are depicted in the right-most column of Table 7. P12 acquired all baseline relations during the first training, thus two sessions per baseline relation, one prompted and one unprompted, were conducted during the retraining. There was only one session per tested relation for both original testing and retesting. Accuracy on all trained relations was 100%. On prompt tests, class-consistent responding was 95.8% (range: 83–100%); P12 passed all prompt tests except E-D, on which class-consistent responding was 83%. Relative to the first set of prompt tests, retraining seems to have facilitated emergence of the stimulus-prompt relations for the D stimuli. As in the previous test, P12 passed all tests of symmetry and transitivity (range: 94–100%). Finally, unlike in the first test, P12 passed all tests for equivalence/class-expansion: Mean class consistent responding was 96.2% (range: 89–100%). On the final posttest sort (12 (3) in Table 6), P12 sorted stimuli into three piles, each consistent with experimenter-defined relations, thus passing this test.

**Discussion**

The results of Experiment 2 replicate those of Experiment 1 and demonstrate the formation of three five-member equivalence classes that include four trained stimuli and class-specific prompts, thus expanding the literature on stimulus equivalence., These results add support for Sidman's (2000) theory that all positive elements in a contingency enter the equivalence class (e.g., Dube et al., 1993; Dube et al., 1989; Johnson et al., 2014; Lionello-DeNolf & Braga-Kenyon, 2013; Minster et al., 2006; Shimizu, 2006). Moreover, prompt inclusion was verified using both non-representative and meaningful (i.e., educationally relevant) stimuli across different populations (i.e., typically developing adults and individuals diagnosed with autism).

The analysis of the comparison responses latencies showed that the mean latencies were generally slower on emergent-relation test trials than training trials across a majority of participants in both Experiments 1 and 2. However, the participants diagnosed with autism displayed slower latencies compared to the typically developing adults. The mean difference between populations was smaller for prompted training, independent training, reflexivity, and symmetry trials (mean differences of 0.3, 0.8, 0.8, and 0.4 s, respectively) than for transitivity, class expansion, and prompt test trials (mean differences of 1.1, 2.5, and 1.0 s, respectively).

The results of Experiment 2 have implications for educational practices in applied settings because they demonstrate that prompts can enter equivalence classes when individuals diagnosed with ASD are trained on conditional discriminations with meaningful stimuli. The class expansion conducted via class-specific prompts lead to novel relations emerging without any direct training. Instructional strategies used across schools and organizations offering applied behavior analysis (ABA) services often include use of prompts, and mostly not class-specific prompts. We hypothesize that if those prompts are common across classes, stimulus classes may merge, which could hinder acquisition of the relations being taught. In other words, if one is teaching the difference between apples and bananas using a common prompt, and the prompt facilitates class merger, this may inadvertently teach the learner that apples and bananas are equivalent. This may be one underlying factor in prompt dependency. Moreover, even if use of a common prompt does not result in class merger, use of class-specific prompts may facilitate the acquisition of discriminations, similar to the way use of class-specific outcomes does (cf. Urcuioli, 2005), and may possibly reduce prompt dependency. For example, Braga-Kenyon et al. (2017) demonstrated that their participants acquired the trained conditional discrimination with fewer trials and fewer errors during training trials with class-specific prompts relative to trials in which a common prompt was used.

One factor that limits the conclusions that can be drawn from Experiment 2 is that the stimulus-stimulus relations were not pretested using the MTS procedure to ensure they were not established prior to training. Rather, pretesting consisted only of the sorting task. Because meaningful stimuli were used, one cannot exclude the possibility the participants might have had prior exposure to those relations  An MTS pretest, therefore, would have been a stronger method of establishing a lack of a pre-experimentally established relations. Nonetheless, the sorting pretest data indicated no evidence of experimenter-defined relations. Moreover, the lack of class-consistent sorting in the pretest was not a result of lack of sorting skills per se. All participants demonstrated the ability to sort based on identity and category using non-experimental stimuli during the pre-experimental phases. The extant literature supports correspondence between sorting and MTS tests (Arntzen et al., 2015; 2017; Fields et al., 2014). For this population, the advantage of sorting as the pre- and posttest measures was that it significantly reduced exposure to tests without reinforcement and reduced the amount of time away from their regularly scheduled school activities.

Unlike in Experiment 1, sessions in Experiment 2 were conducted across days to limit the time the participants spent outside their regular instructional activities. In addition, sessions were scheduled based on the participants' availability and individual behavior guidelines were followed (e.g., only engaging in task demands for 25 consecutive minutes, or honoring mands for break from demands). It is also possible that the participants could have gained access to the trained relations outside of the training and testing sessions because meaningful stimuli were used. Although this resulted in reduced experimental control, it does mimic the typical situation in special education settings. Importantly, even with reduced experimental control, the results largely confirmed the inclusion of prompt stimuli in the equivalence class.

## General Discussion

Experiments 1 and 2 add evidence supporting Sidman's (2000) theory that all positive elements present in a contingency enter equivalence classes, including samples, comparisons, reinforcers, responses, and prompts. The comparison of the pretest/posttest results, conducted via a sorting task, demonstrated that class formation included class-specific prompts with arbitrary

(Experiment 1) and meaningful (Experiment 2) stimuli with neurotypical adults (Experiment 1) and children diagnosed with ASD (Experiment 2).

Inclusion of prompts in equivalence classes could also be analyzed in terms of multiple element stimuli or compound stimuli. The prompt, when presented with the sample and comparison stimuli, could have become part of a 2-element compound stimulus (e.g., AE→BE) that, when elements were separated (i.e., during the fading procedure), demonstrated the same control as the compound stimulus (cf. Stromer et al., 1993). A potential outcome of using compound stimuli in MTS tasks is the development of inappropriate stimulus control, known as stimulus overselectivity (Lovaas et al., 1979) or restricted stimulus control (Dube et al., 2010). Restricted stimulus control may be defined as control by either a few relevant elements of a compound stimulus or by stimulus-irrelevant elements (e.g., position preference). During training in the current set of studies, the prompt could have controlled responding to the exclusion of the other visual stimulus element (i.e., responding consistent with identity MTS performance rather arbitrary MTS performance). Under these conditions, once the color prompt was faded completely, the other visual element may have failed to control responding consistent with training, documenting restricted stimulus control. Fortunately, the results of these two studies indicated that, for the majority of participants, all elements of the compound stimuli acquired discriminative control consistent with the compound stimuli training histories. These data lend support for the contiguous presentation of stimuli (e.g., elements of a compound stimulus) as a critical variable in the formation of equivalence classes (Schenk, 1993; Stromer et al., 1993; Maguire et al., 1994). Additionally, the requirement of a non-differential observing response during training may have promoted attention to the individual components during training (Dube et al., 2010).

The current results may have implications for enhancing errorless learning technology for individuals who display restricted stimulus control and faulty transfer of stimulus control from prompts used in teaching to the intended discriminative stimulus. Because errorless teaching procedures often employ common prompts (e.g., pointing), future research should investigate ≈ conditions under which prompts enter a class or do not. For example, research directly comparing use of class-consistent prompts to class-inconsistent and/or common prompts during conditional discrimination teaching and testing of emergent relations is needed. In addition, future research should investigate use of class-specific prompts as a remediation strategy for reducing prompt dependency.

Notably, not all participants in this study showed evidence of equivalence class formation. The reasons are not clear, and may be related to issues during training (e.g., selective attention) or to the testing sequence used (i.e., the simple-to-complex training/testing protocol). Future research could investigate use of other training/testing sequences and potential interactions with the use of class-specific prompts. In addition, use of different (1) prompting procedures (e.g., the symbols presented in class-consistent colors rather than superimposed), (2) methods of fading the prompts, and (3) differential observing responses should be explored. Collectively, the results of such studies have the potential to enhance errorless teaching methods in applied settings.

## References

Albright, L., Reeve, K. F., Reeve, S. A., & Kisamore, A. N. (2015). Teaching statistical variability with equivalence-based instruction. *Journal of Applied Behavior Analysis, 48*(4), 883–894. https://doi.org/10.1002/jaba.249

Arntzen, E., Granmo, S., & Fields, L. (2017). The relation between sorting tests and matching-to-sample tests in the formation of equivalence classes. *The Psychological Record*, *67*(1), 81–96. https://doi.org/10.1007/s40732-016-0209-9

Arntzen, E., Norbom, A., & Fields, L. (2015). Sorting: An alternative measure of class formation? *The Psychological Record*, *65*(4), 615–625. https://doi.org/10.1007/s40732-015-0132-5

Boldrin, L. S., & Debert, P. (2017). *Matching to Sample Procedure Software* (Version 1.0) [Computer Software]. São Paulo, SP: Instituto de Psicologia da Universidade de São Paulo.

Braga-Kenyon, P., Guilhardi, P., Lionello-DeNolf, K. M., & Dube, W. V. (2017). Teaching visual conditional discriminations using errorless learning: The role of prompts requiring simple and conditional discriminative control. *European Journal of Behavior Analysis*, *18*(2), 1–15. https://doi.org/10.1080/15021149.2017.1309624

Debert, P., Matos, M. A., & McIlvane, W. (2007). Conditional relations with compound abstract stimuli using a go/no-go procedure. *Journal of the Experimental Analysis of Behavior*, *87*(1), 89-96. https://doi.org/10.1901/jeab.2007.46-05

Dube, W. V., Green, G., & Serna, R. W. (1993). Auditory successive conditional discrimination and auditory stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior*, *59*, 103–114. https://doi.org/10.1901/jeab.1993.59-103

Dube, W. V., McIlvane, W. J., Maguire, R. W., Mackay, H. A., and Stoddard, L. T.

(1989). Stimulus class formation and stimulus-reinforcer relations. *Journal of Experimental Analysis of Behavior*, *51*, 65–76. https://doi.org/10.1901/jeab.1989.51-65

Fields, L., Arntzen, E., & Moksness, M. (2014). Stimulus sorting: A quick and sensitive index of equivalence class formation. *The Psychological Record*, *64*(3), 487–498. https://doi.org/10.1007/s40732-014-0034-y

Fienup, D. M., Wright, N. A., & Fields, L. (2015). Optimizing equivalence-based instruction: Effects of training protocols on equivalence class formation. *Journal of Applied Behavior Analysis*, *48*(3), 613–631. https://doi.org/10.1002/jaba.234

Fisher, W. W., Piazza, C. C., Bowman, L. G., & Amari, A. (1996). Integrating caregiver report with a systematic choice assessment. *American Journal on Mental Retardation, 101,* 15–25. Retrieved from http://ttac.odu.edu/_public/file/ReinfAssessIndSevereDisab.pdf

Green, G. (2001). Behavior analytic instruction for learners with autism: Advances in stimulus control technology. *Focus on Autism and Other Developmental Disabilities*, *16*, 72–85. https://doi.org/10.1177/108835760101600203

Guinther, P. M., & Dougher, M. J. (2015). The clinical relevance of stimulus equivalence and relational frame theory in influencing the behavior of verbally competent adults. *Current Opinion in Psychology, 2*, 21–25.

Johnson, C., Meleshkevich, O., and Dube, W. V. (2014). Merging separately established stimulus classes with outcome-specific reinforcement. *Journal of Experimental Analysis of Behavior*, *101*, 38–50. https://doi.org/10.1002/jeab.61

LeBlanc, L. A., Miguel, C. F., Cummings, A. R., Goldsmith, T. R., & Carr, J. E. (2003). The effects of three stimulus-equivalence testing conditions on emergent US geography relations of children diagnosed with autism. *Behavioral Interventions: Theory & Practice in Residential & Community-Based Clinical Programs*, *18*(4), 279–289. https://doi.org/10.1002/bin.144

Lionello-DeNolf, K. M., & Braga-Kenyon, P. (2013). Membership of defined responses in stimulus classes. *The Psychological Record, 63*, 769–784. https://doi.org/10.11133/j.tpr.2013.63.4.005

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. https://doi.org/10.1901/jeab.2003.79-21

Lovaas, O. I., Koegel, R. L., & Schreibman, L. (1979). Stimulus overselectivity in autism: A review of research. *Psychological Bulletin, 86*(6), 1236–1254. https://doi.org/10.1037/0033-2909.86.6.1236

MacDuff, G. S., Krantz, P. J., & McClannahan, L.E. (2001). Prompts and prompt-fading strategies for people with autism. In C. Maurice, G. Green, & R. M. Foxx (Eds.). *Making a Difference: Behavior Intervention for Autism*. Austin, TX: PRO-ED, 37–50. ISBN-13:978-0890798713

Maguire, R. W, Stromer, R., MacKay, H. A., & Demis, C. A. (1994) Matching to complex samples and stimulus class formation in adults with autism and young children. *Journal of Autism and Developmental Disabilities, 24*, 753–772. https://doi.org/10.1007/BF02172284

Minster, S. T., Jones, M., Elliffe, D., & Muthukumaraswamy, S. D. (2006). Stimulus equivalence: Testing Sidman's (2000) theory. *Journal of the Experimental Analysis of Behavior*, *85*, 371–391. https://doi.org/10.1901/jeab.2006.15-05

Pilgrim, C. (2019). Equivalence-based Instruction. In J. O. Cooper, T. E. Heron, & W. L.

Heward (Eds.), *Applied Behavior Analysis* (3rd ed., pp. 452–496). Hoboken, NJ: Pearson Education.

Rincover, A. (1978). Variables affecting stimulus fading and discriminative responding in psychotic children. *Journal of Abnormal Psychology, 87*(5), 541–553. https://doi.org/10.1037/0021-843X.87.5.541

Schenk, J. J. (1993) Emergent conditional discrimination in children: Matching to compound Stimuli. *The Quarterly Journal of Experimental Psychology Section B, 46*,(4), 345–365. https://doi.org/10.1080/14640749308401076

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

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, *14*, 5–13. https://doi.org/10.1044/jshr.1401.05

Sidman, M. (1986). Functional analysis of emergent verbal classes. In Zeiler, M. D. & Thompson T. (Eds.), *Analysis and Integration of Behavioral Units* (pp. 213–245). Hillsdale, NJ: Erlbaum. ISBN-13: 978-1138659612

Sidman, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman, & H. Lejeune (Eds.), *Behaviour Analysis in Theory and Practice: Contributions and Controversies* (pp. 93–114). Hillsdale, NJ: Erlbaum. ISBN-13 : 978-0863771446

Sidman, M. (1994). *Equivalence Relations and Behavior: A Research Story.* Boston: Authors Cooperative. ISBN-13:978-0962331169.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, *74*, 127–146. https://doi.org/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. https://doi.org/10.1901/jeab.1982.37-5

Stromer, R., McIlvane, W. J., & Serna, R. W. (1993). Complex stimulus control and equivalence. *The Psychological Record, 43*, 585–598. https://doi.org/10.1007/BF03395901

Tomanari, G. Y., Sidman, M., Rubio, A. R., & Dube, W. V. (2006). Equivalence classes with requirements for short response latencies. *Journal of the Experimental Analysis of Behavior*, *85*(3), 349–369. https://doi.org/10.1901/jeab.2006.107-04

Touchette, P. E. (1971). Transfer of stimulus control: Measuring the moment of transfer. *Journal of the Experimental Analysis of Behavior*, *15*(3), 347–354. https://doi.org/10.1901/jeab.1971.15-347

Touchette, P. E., & Howard, J. S. (1984). Errorless learning: Reinforcement contingencies and stimulus control transfer in delayed prompting. *Journal of Applied Behavior Analysis*, *17*, 175–188. https://doi.org/10.1901/jaba.1984.17-175

Urcuioli, P. J. (2005). Behavioral and associative effects of differential outcomes in discrimination learning. *Learning & Behavior, 33*(1), 1–21. https://doi.org/10.3758/BF03196047

**Table 1**

*Experimental Conditions from Experiments 1 and 2*

| Condition | Task Type | Session Type | Relation |
|---|---|---|---|
| Pretest | Sorting | Test | All |
| Phase 1 | MTS | Training | A-B+E |
| Phase 2 | MTS | Test | B-A; E-A; E-B; A-E; B-E |
| Phase 3 | MTS | Training | A-C+E |
| Phase 4 | MTS | Test | C-A; E-C; C-E |
| Phase 5 | MTS | Test | B-C; C-B |
| Phase 6 | MTS | Training | D-D+E |
| Phase 7 | MTS | Test | E-D; D-E |
| Phase 8 | MTS | Test | D-A; D-B; D-C; A-D; B-D; C-D |
| Phase 9 | MTS | Test | A-A; B-B; C-C |
| Posttest | Sorting | Test | All |

*Note*. Phases 1, 3, and 6 trained baseline, conditional discrimination relations using a class-consistent prompt. Phases 2 and 5 tested for symmetry and transitivity, respectively. Phases 4 and 7 tested whether the prompt would enter the class when presented as a sample or comparison stimulus. Phase 8 tested for class expansion (i.e., equivalence between the A, B, C, and D stimuli). Phase 9 tested for reflexivity. “MTS” refers to the matching-to-sample procedure.

**Table 2**

*Demographic Information for Participants at the Beginning of Experiments 1 and 2*

| **Experiment 1** | | | |
|---|---|---|---|
| **Participant** | **Gender** | **Age** | **Education Level** |
| 1 | Female | 24 | Graduate student |
| 2 | Female | 40 | Graduate Student |
| 3 | Male | 20 | Undergraduate Student |
| 4 | Female | 18 | Undergraduate Student |
| 5 | Female | 19 | Undergraduate Student |
| 6 | Female | 21 | Undergraduate Student |
| 7 | Female | 23 | Graduate Student |
| 8 | Female | 21 | Undergraduate Student |
| 9 | Female | 22 | Graduate Student |
| 10 | Female | 27 | Graduate Student |
| **Experiment 2** | | | |
| **Participant** | **Gender** | **Age** | **Diagnosis** |
| 11 | Male | 16 | autism spectrum disorder |
| 12 | Male | 12 | autism spectrum disorder |
| 13 | Female | 21 | autism spectrum disorder, attention deficit disorder, mood disorder, seizure disorder, and Dandy Walker syndrome |
| 14 | Male | 18 | autism spectrum disorder and Landau Kleffner syndrome |

*Note*. Participants' age is in years.

**Table 3**

*Individual Participant Data from the Sorting Pre- and Posttests in Experiment 1*

**Pre-class formation sorting**

| | **Pile 1** | | | **Pile 2** | | | **Pile 3** | | | **Pile 4** | | | **Pile 5** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Part** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** |
| **1** | 1 | 1 | 0 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 3 | | | |
| **2** | 4 | 4 | 4 | 1 | 1 | 1 | | | | | | | | | |
| **3** | 4 | 4 | 4 | 1 | 1 | 1 | | | | | | | | | |
| **4** | 3 | 2 | 0 | 1 | 2 | 5 | 1 | 1 | 1 | | | | | | |
| **5** | 1 | 0 | 4 | 3 | 4 | 0 | 1 | 1 | 1 | | | | | | |
| **6** | 4 | 4 | 4 | 1 | 1 | 1 | | | | | | | | | |
| **7** | 0 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 3 | 1 | 0 | | | |
| **8** | 3 | 1 | 0 | 1 | 2 | 0 | 1 | 1 | 1 | 0 | 0 | 2 | 0 | 1 | 2 |
| **9** | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 2 | 2 | 2 | 2 | 0 | | | |
| **10** | 1 | 1 | 1 | 3 | 3 | 1 | 1 | 1 | 3 | | | | | | |

**Post-class formation sorting**

| | **Pile 1** | | | **Pile 2** | | | **Pile 3** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Part** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** |
| **1** | 0 | 0 | 5 | 0 | 5 | 0 | 5 | 0 | 0 |
| **2** | 0 | 5 | 0 | 5 | 0 | 0 | 0 | 0 | 5 |
| **3** | 0 | 5 | 0 | 5 | 0 | 0 | 0 | 0 | 5 |
| **4** | 0 | 0 | 5 | 5 | 0 | 0 | 0 | 5 | 0 |
| **5** | 0 | 5 | 0 | 0 | 0 | 5 | 5 | 0 | 0 |
| **6** | 5 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 5 |
| **7** | 0 | 5 | 0 | 5 | 0 | 0 | 0 | 0 | 5 |
| **8** | 0 | 5 | 0 | 0 | 0 | 5 | 5 | 0 | 0 |
| **9** | 5 | 0 | 0 | 0 | 0 | 5 | 0 | 5 | 0 |
| **10** | 1 | 1 | 1 | 4 | 3 | 0 | 0 | 1 | 4 |

Note. "Part" refers to participant. "C1," "C2", and "C3" refer to experimenter-defined stimulus classes 1, 2, and 3, respectively. Sorting consistent with experimenter-defined relations would result in three piles represented by the strings 5-0-0, 0-5-0, and 0-0-5.

**Table 4**

*Percent Correct in Training and Derived Relations Testing for Participants in Experiment 1*

| Relation | Type | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A-B | Training | 100 | 100 | 99 | 100 | 100 | 100 | 100 | 98 | 100 | 90 |
| A-C | Training | 98 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 98 | 100 |
| D-D | Training | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| A-E | Prompt test | 100 | 100 | 100 | 100 | 33 | 100 | 67 | 100 | 100 | 78 |
| B-E | Prompt test | 100 | 100 | 100 | 100 | 28 | 100 | 72 | 100 | 100 | 94 |
| E-A | Prompt test | 100 | 100 | 100 | 100 | 33 | 100 | 72 | 94 | 100 | 94 |
| E-B | Prompt test | 100 | 100 | 100 | 94 | 33 | 100 | 83 | 100 | 100 | 89 |
| C-E | Prompt test | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| E-C | Prompt test | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| D-E | Prompt test | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| E-D | Prompt test | 100 | 100 | 100 | 94 | 100 | 100 | 100 | 100 | 100 | 100 |
| B-A | Symmetry | 100 | 100 | 100 | 100 | 100 | 100 | 94 | 100 | 100 | 100 |
| C-A | Symmetry | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 89 | 100 | 100 |
| B-C | Transitivity | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 39 |
| C-B | Transitivity | 100 | 100 | 100 | 94 | 100 | 100 | 100 | 100 | 100 | 39 |
| A-D | Equivalence | 100 | 100 | 100 | 100 | 94 | 100 | 100 | 100 | 94 | 83 |
| B-D | Equivalence | -- | -- | -- | 100 | 100 | 100 | 100 | 100 | 94 | 33 |
| C-D | Equivalence | -- | -- | -- | 89 | 100 | 100 | 100 | 94 | 100 | 100 |
| D-A | Equivalence | 100 | 94 | 100 | -- | 100 | 100 | 100 | 94 | 100 | 50 |
| D-B | Equivalence | -- | -- | -- | 100 | 100 | 100 | 100 | 100 | 100 | 67 |
| D-C | Equivalence | -- | -- | -- | 100 | 100 | 100 | 100 | 100 | 100 | 28 |
| A-A | Reflexivity | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| B-B | Reflexivity | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| C-C | Reflexivity | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

*Note*. The training scores indicate the percent of class-consistent responding in the absence of prompting. The "--" indicates that the test was not presented due to a computer error.

**Table 5**

*Individual Participant Data from the Pre-experimental Sorting Tasks in Experiment 2*

| **Identity sorting** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Pile 1** | | | **Pile 2** | | | **Pile 3** | | |
| **Part** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** |
| **11** | 0 | 5 | 0 | 5 | 0 | 0 | 0 | 0 | 5 |
| **12** | 5 | 0 | 0 | 0 | 0 | 5 | 0 | 5 | 0 |
| **13** | 0 | 5 | 0 | 5 | 0 | 0 | 0 | 0 | 5 |
| **14** | 0 | 5 | 0 | 0 | 0 | 5 | 5 | 0 | 0 |
| **Category sorting** | | | | | | | | | |
| | **Pile 1** | | | **Pile 2** | | | **Pile 3** | | |
| **Part** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** | **C1** | **C2** | **C3** |
| **11** | 5 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 5 |
| **12** | 0 | 5 | 0 | 0 | 0 | 5 | 5 | 0 | 0 |
| **13** | 5 | 0 | 0 | 0 | 0 | 5 | 0 | 5 | 0 |
| **14** | 0 | 5 | 0 | 5 | 0 | 0 | 0 | 0 | 5 |

*Note*. "C1," "C2," and "C3" refer to experimenter-defined classes of pink, light blue, and orange, respectively, for identity sorting and of animals, edibles, and clothing, respectively, for category sorting.

**Table 6**

*Individual Participant Data from the Sorting Pre- and Posttests in Experiment 2*

| **Pre-class formation sorting** | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Pile 1** | | | **Pile 2** | | | **Pile 3** | | | **Pile 4** | | | **Pile 5** | | | **Pile 6** | | |
| **Part** | **1** | **2** | **3** | **1** | **2** | **3** | **1** | **2** | **3** | **1** | **2** | **3** | **1** | **2** | **3** | **1** | **2** | **3** |
| **11** | 1 | 0 | 0 | 2 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 2 | 0 | 2 | 1 | 1 | 0 | 1 |
| **12** | 3 | 2 | 0 | 1 | 2 | 2 | 1 | 1 | 3 | | | | | | | | | |
| **13** | 1 | 2 | 1 | 4 | 3 | 4 | | | | | | | | | | | | |
| **14** | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | | | | | | | | | |

| **Post-class formation sorting** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Pile 1** | | | **Pile 2** | | | **Pile 3** | | |
| **Part** | **1** | **2** | **3** | **1** | **2** | **3** | **1** | **2** | **3** |
| **12 (1)** | 1 | 0 | 4 | 4 | 0 | 1 | 0 | 5 | 0 |
| **12 (2)** | 1 | 0 | 4 | 4 | 0 | 1 | 0 | 5 | 0 |
| **12 (3)** | 5 | 0 | 0 | 0 | 0 | 5 | 0 | 5 | 0 |
| **13** | 0 | 0 | 5 | 5 | 0 | 0 | 0 | 5 | 0 |
| **14** | 5 | 0 | 0 | 0 | 0 | 5 | 0 | 5 | 0 |

*Note.* “Part” refers to participant. “1,” “2,” and “3” refer to experimenter-defined classes 1, 2, and 3, respectively.

**Table 7**

*Percent Correct on Matching-to-Sample Sessions for Individual Participants in Experiment 2*

| **Relation** | **Type** | **P12** | **P13** | **P14** | **P12 (second)** |
|---|---|---|---|---|---|
| A-B | Training | 99 | 98 | 100 | 100 |
| A-C | Training | 99 | 100 | 100 | 100 |
| D-D | Training | 100 | 100 | 100 | 100 |
| A-E | Prompt Test | 100 | 100 | 100 | 100 |
| B-E | Prompt Test | 100 | 100 | 100 | 100 |
| E-A | Prompt Test | 100 | 89 | 100 | 100 |
| E-B | Prompt Test | 100 | 94 | 100 | 100 |
| C-E | Prompt Test | 94 | 83 | 100 | 100 |
| E-C | Prompt Test | 94 | 100 | 100 | 94 |
| D-E | Prompt Test | 22 | 100 | 100 | 89 |
| E-D | Prompt Test | 22 | 89 | 100 | 83 |
| B-A | Symmetry | 94 | 100 | 100 | 100 |
| C-A | Symmetry | 94 | 100 | 100 | 94 |
| B-C | Transitivity | 94 | 100 | 100 | 100 |
| C-B | Transitivity | 94 | 100 | 100 | 100 |
| A-D | Equivalence | 22 | 100 | 100 | 100 |
| B-D | Equivalence | 33 | 94 | 100 | 100 |
| C-D | Equivalence | 33 | 100 | 100 | 89 |
| D-A | Equivalence | 17 | 100 | 100 | 94 |
| D-B | Equivalence | 17 | 100 | 100 | 100 |
| D-C | Equivalence | 22 | 94 | 100 | 94 |
| A-A | Reflexivity | 100 | 100 | 100 | -- |
| B-B | Reflexivity | 100 | 100 | 100 | -- |
| C-C | Reflexivity | 100 | 100 | 100 | -- |

*Note*. The training scores indicate the percent of class-consistent responding in the absence of prompting. The "--" indicates that the test was not presented during the retraining phase.

**Figure 1**

*Experimental stimuli for Experiment 1(left columns) and Experiment 2 (right columns)*

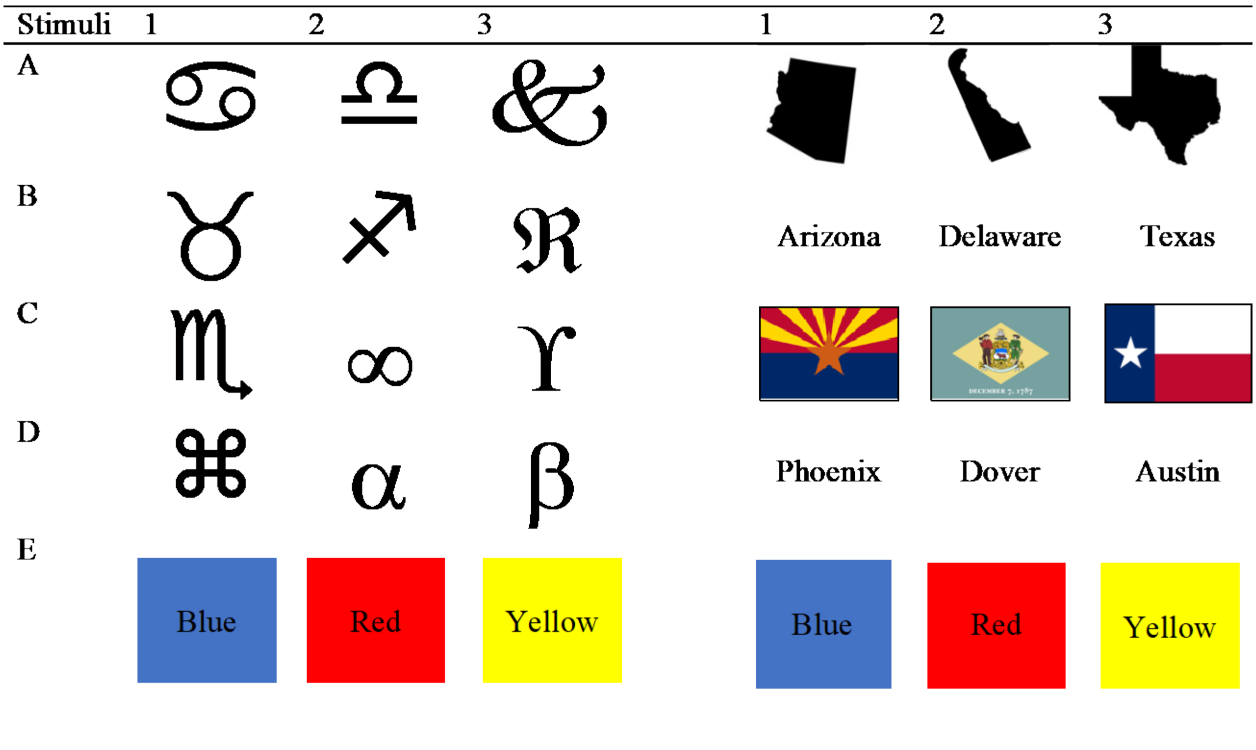
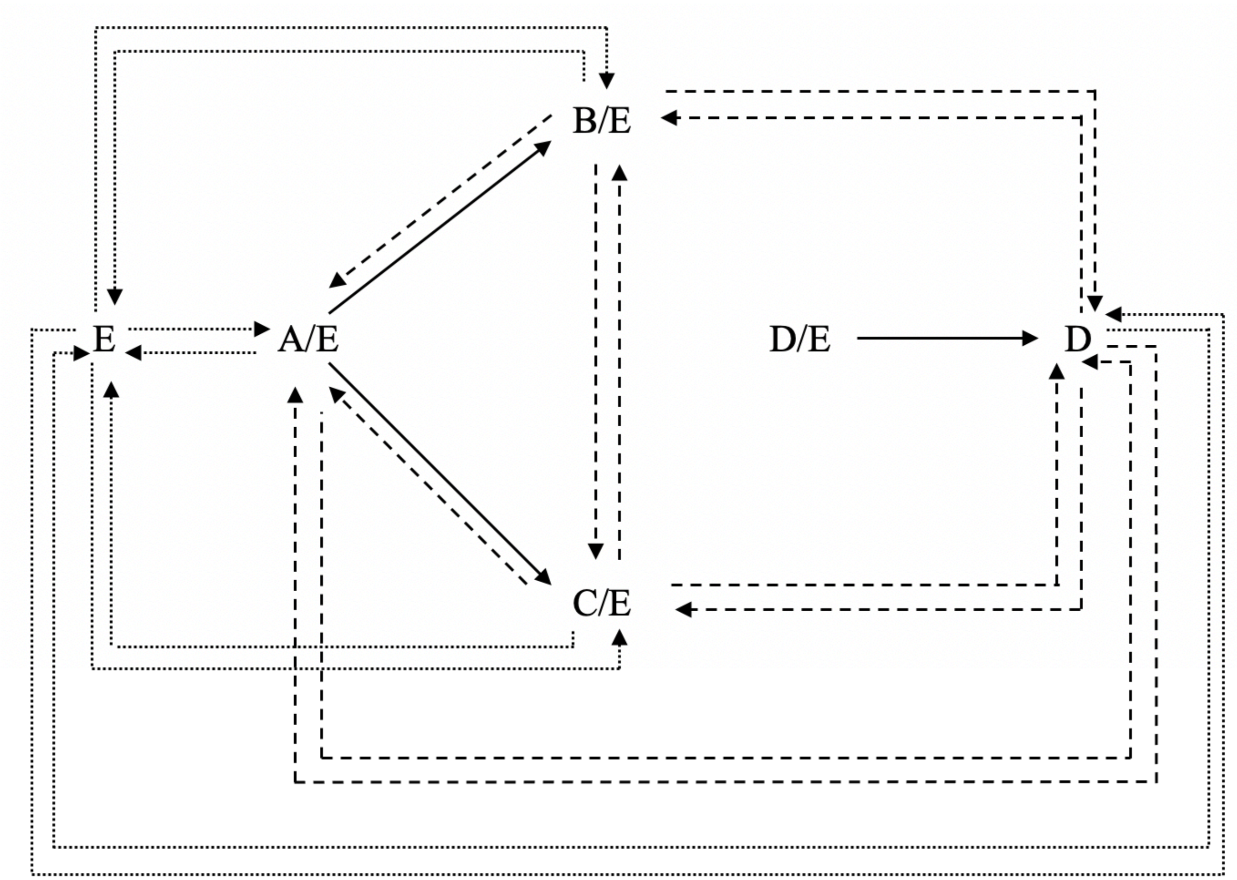


**Figure 2**

*Schematic representation of the trained and tested relations in Experiments 1 and 2*




*Note.* Stimuli in Experiment 1 were arbitrary symbols, whereas those in Experiment 2 were meaningful visual stimuli. Stimuli from Set A were related to stimuli from Sets B and C using the class-specific prompt (E). The solid arrows depict directly trained relations, the dashed arrows depict derived relations tested during probe sessions, and the dotted lines indicate relations between the arbitrary symbols and the color prompts, when each appeared as sample and comparisons.