Published in final edited form as:

Psychol Rec. 2014 June 1; 64(2): 195–208. doi:10.1007/s40732-014-0021-3.

Controlling Relations in Stimulus Equivalence Classes of Preschool Children and Individuals with Down Syndrome

Priscila C. Grisante^{a,b}, Julio C. de Rose^{a,b}, and William J. McIlvane^{b,c}

^aUniversidade Federal de São Carlos

^bInstituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino

^cUniversity of Massachusetts Medical School

Abstract

We evaluated emergent stimulus-stimulus relations after two different training procedures. Participants were five typically developing preschool children and three individuals with Down Syndrome. Experiment 1 used two-comparison matching to sample (MTS) to establish AB and BC relations. Experiment 2 used two-comparison and blank-comparison MTS, each on 50% of training trials to establish AB and BC relations. In both experiments, tests for emergent relations (AC, CA) were conducted to assess equivalence class formation. In Experiment 2 subsequently, class expansion was assessed after CD training. All participants showed positive equivalence test outcomes. Seven showed class expansion. After class formation tests in both studies, probe tests were conducted for *select* and *reject* relations in baseline relations. Initial results were somewhat variable, but became more consistent after class expansion.

Keywords

select and *reject* relations; preschool children; Down Syndrome; matching to sample

In Set Theory, equivalence relations are defined as having the properties of reflexivity, symmetry, and transitivity. Sidman and colleagues (Sidman, 1986, 1994, 2000; Sidman & Tailby, 1982) applied this criterion to arbitrary stimulus-stimulus relations, arguing further that equivalence relations involving such stimuli are symbolic relations.

In typical stimulus equivalence studies, relations between stimuli are usually established by arbitrary matching-to-sample (MTS) procedures. Samples of a set designated A (A1, A2... An) are matched with those of a set B (B1, B2...Bn) displayed in an array of two or more comparison stimuli. Selection of a given comparison stimulus (e.g., B1) is reinforced when a related sample is present (e.g., A1) but not when unrelated samples are present. Such training establishes arbitrary matching relations involving each sample from set A and a corresponding comparison from set B (AB relations). When two or more sets of relations are

Correspondence concerning this manuscript should be addressed to: Julio C. de Rose, Universidade Federal de São Carlos/ Departamento de Psicologia, Rodovia Washington Luís, km 235 - SP-310, São Carlos - São Paulo - Brasil, CEP 13565-905, juliocderose@yahoo.com.br.
Authors' email addresses, respectively: prisgrisante@gmail.com; juliocderose@yahoo.com.br; william.mcilvane@umassmed.edu.

established with humans (e.g., AB, BC), the properties of reflexivity, symmetry, and transitivity are often demonstrated by emergence of untrained relations such as BA, CB, AC, CA. These emergent relations document the properties of equivalence relations. Moreover, several studies have provided evidence that members of stimulus equivalence classes are indeed related symbolically (e.g., Barnes-Holmes et al., 2005; Bortoloti & de Rose, 2009, 2012; Haimson, Wilkinson, Rosenquist, Ouimet, & McIlvane, 2009).

Most humans who learn matching relations between sets of stimuli typically show equivalence classes, either promptly or in repeated testing (referred as “delayed emergence;” see Sidman, 1994). However, some may demonstrate MTS relations but not equivalence classes. Such variability remains poorly understood. Devany, Hayes, and Nelson (1982) have proposed that stimulus equivalence is correlated with the development of language (see also Hayes, Barnes-Holmes, & Roche, 2001; Lipkens, Hayes, & Hayes, 1993; Horne & Lowe, 1996). Other authors have suggested that outcomes of tests for stimulus equivalence depend critically on the nature of stimulus control topographies (SCTs) (cf. McIlvane & Dube, 1992) established during MTS training (e.g., Carrigan & Sidman, 1992; de Rose, 1996; de Rose, Hidalgo, & Vasconcellos, 2013; Dube & McIlvane, 1996; McIlvane, 2012; McIlvane, Serna, Dube, & Stromer, 2000; Sidman, 1979, 1994).

When a typical MTS procedure is employed, the experimenter can record participants’ responses, but cannot know with certainty which aspects of the sample and comparison stimuli control responding (McIlvane & Dube, 1992). In a recent article in this journal, de Rose et al. (2013) illustrated in their Figure 1 how correct responding in two-comparison MTS may result from different SCTs. Both panels of the figure show the participant selecting the stimulus defined as correct in the presence of the sample. The left panel shows a selection resulting from a sample-S+, or *select*, SCT (represented by an arrow pointing from the sample to the correct comparison). In this case, the participant need not attend to defining features of the S-. By contrast, the right panel shows a selection resulting from a sample-S-, or *reject*, SCT (represented by an oval-tipped arrow pointing from sample to incorrect comparison). Although the participant touches the correct comparison, the controlling relation involves merely the sample and the incorrect comparison.

Carrigan and Sidman (1992) offered an analysis of outcomes that might be expected on tests of emergent relations when SCTs are *select only* or *reject only*. With the former, the expected equivalence classes should emerge. With the latter, expected classes should not. Their analysis offers one explanation for inter-subject variability in equivalence class test outcomes (e.g., Carr, Wilkinson, Blackman, & McIlvane, 2000; Kato, de Rose, & Faleiros, 2008; Saunders, Drake, & Spradlin, 1999; Smeets & Barnes-Holmes, 2004) and occasional inconsistency in individual performances (e.g., Saunders & McEntee, 2004).

Certain techniques may (1) clarify whether participants exhibit *select* and *reject* relations following MTS training and (2) increase the probability that such SCTs will be acquired. Two are pertinent to the present study – the novel stimulus substitution procedure (e.g., Cumming & Berryman, 1965; Stromer & Osborne, 1982) and the blank comparison (or mask) procedure (McIlvane, Withstandley, & Stoddard, 1984; McIlvane et al., 1987).

Novel stimulus substitution procedure

With this procedure, SCTs are inferred by systematic responding to arrays that include novel stimuli. Given AB training, for example, *select* SCTs are assessed with trials presenting (1) A1 or A2 as samples and (2) comparison displays that include the corresponding stimulus and a novel stimulus (B1/N1 and B2/N2, respectively). *Select* SCTs are shown if the participant selects B1 in relation to A1 and B2 in relation to A2 (and does not select novel stimuli). *Reject* SCTs may be demonstrated when each sample is displayed with a comparison stimulus defined as *incorrect* for that sample and a novel stimulus (e. g., A1 and B2/N3, and A2 and B1/N4). If the participant consistently selects the novel stimuli, this is evidence that *reject* SCTs were established in training.

In their Experiment 1, Stromer and Osborne used novel stimulus probes after training a two-comparison AB conditional discrimination and after a BA Symmetry test. They found that each of four adolescents with mild intellectual disabilities exhibited both *select* and *reject* relations on all probes.

Blank comparison procedure

A typical procedure is as follows: A two-comparison MTS baseline is established. Then, stimulus control shaping covers one of the comparison stimuli (S+ or S−) on each trial with a solid color mask that increases in size progressively. In the final performance, half of the trials with each sample present S+ stimuli along with a mask covering S-stimuli; the other half present S-stimuli along with the mask covering S+ stimuli. The final performance requires participants to attend to and to respond (or not) to non-masked comparison stimuli on the basis of *select* or *reject* relations with the samples.

The blank comparison procedure can be used either *after* conditional discrimination training (to evaluate established SCTs) or *during* training (to promote acquisition of specified SCTs). To illustrate the former use, suppose a participant is taught AB matching in the two-comparison format and also identity matching in the blank-comparison format. Thereafter, one could present the AB trials in the blank-comparison format to assess whether independent *select* and *reject* relations were demonstrable (McIlvane et al., 1984).

To illustrate the latter use, one might arrange trial distributions such that (1) each non-masked comparison stimuli is a valid match with its corresponding sample on 50% of trials and (2) the masked comparison is a valid match on the other 50% of trials (encouraging no bias towards *select* or *reject* relations). Alternately, one might bias the participant towards acquiring *select* or *reject* relations by decreasing or increasing, respectively, the proportion of trials on which the masked comparison stimulus is the valid match (de Rose et al., 2013).

Kato et al. (2008) found that normally capable adults who showed rapid emergence of stimulus equivalence relations also exhibited reliable *select* and *reject* relations in the blank-comparison format. These results suggested that establishing both *select* and *reject* relations during baseline conditional discriminations training might promote development of experimenter-specified equivalence relations. Supporting this conjecture, de Rose et al. (2013) have shown that acquisition of reliable *select* and *reject* control was accompanied by

reliable and rapid equivalence class formation by school-age children. These results are noteworthy because rapid equivalence class formation was found with a two-node linear design that trained relations AB, BC, and CD. Studies comparing different training designs have shown linear designs to be the least effective (e.g., Arntzen, Grondahl, & Eilifsen, 2010; Arntzen & Holth, 1997). Further, Fields, Adams, Verhave and Newman (1990) showed that delay in emergence of equivalence relations increases with the number of nodes. Thus, de Rose et al (2013) suggested that rapid class formation under the otherwise unfavorable training conditions of their study might be due to the fact that their procedures established both *select* and *reject* SCTs for all trained MTS relations.

Rationale for the present experimentation

Certain human participant populations (e.g., preschool children, older individuals with intellectual disabilities, etc.) tend to show more variability in acquiring MTS baselines and in displaying equivalence relations than those just mentioned (e.g., Augustson & Dougher, 1991; Lazar, Davis-Lang, & Sanches, 1984; O'Donnel & Saunders, 2003; Pilgrim, Jackson, & Galizio, 2000). Given the results of Stromer and Osborne (1982), Kato et al. (2008), and de Rose et al. (2013), one plausible hypothesis relating to such variability is that positive or negative outcomes on tests for equivalence class formation and on tests for *select* and *reject* relations might prove highly correlated. For example, positive outcomes on equivalence class tests would be accompanied by positive outcomes on tests for *select* and *reject* relations.

Following this reasoning, we conducted two experiments investigating acquisition of two-comparison MTS and equivalence class formation with preschool children and participants with Down Syndrome (DS). In both experiments, we sought to assess *select* and *reject* SCTs using the novel stimulus probe procedure. Experiment 1 trained AB and BC relations with a two-comparison MTS procedure. Experiment 2 trained AB and BC relations with 50% of trials using the blank comparison; we reasoned that the blank-comparison training might increase the probability of acquisition of both *select* and *reject* SCTs. Experiment 2 included an additional phase to assess class expansion and potentially associated SCT probe outcomes.

Experiment 1

Method

Participants—The study included five typically developing preschool children and one child and two adults with DS. The two adults could read simple passages of text, and the child with DS could read some words with simple syllables (consonant-vowel). Preschool children knew the names of some letters of the alphabet. No participant had a research history involving matching-to-sample tasks. Further information on the participants is presented in Table 1.

Participants were recruited from their daycare settings or schools. Parents received information about the research and signed a consent form. The Peabody Picture Vocabulary Test-R (Dunn & Dunn, 1981) was administered individually to each participant. This test

assesses receptive vocabulary and yields an equivalent mental age. The WISC III and WAIS III were also administered individually to the child and the adults with DS, respectively. These tests provide a measure of IQ as measured by both verbal and nonverbal tasks.

Setting and Apparatus—For participants with DS, sessions were conducted in a room at their school. Preschool children were studied in a room at the UFSCar Laboratory for the Study of Human Behavior. Stimuli were presented on a Macintosh computer (iMac using MTS version 11.3 [Dube, 1991]). The program also recorded responses and presented programmed consequences for correct and incorrect selections. Participants sat facing the computer, and the researcher sat to the right and slightly behind. Every MTS trial (simultaneous format) displayed five white windows on a gray background, one centered in the screen and the other four in its corners. Sample stimuli were presented on the center window and comparison stimuli on the outer windows. Participants responded to stimuli by clicking on them using the mouse button. Responses to empty windows had no programmed consequences.

Procedures—Sessions for preschool children lasted about 15 minute and were conducted 3-4 times per week; testing spanned 7-14 weeks. Sessions for participants with DS lasted about 30 minutes and occurred three times per week; testing spanned 2-5 weeks.

The procedure had pretraining, training, and testing phases. During all trial blocks of all conditions, the correct and incorrect comparison stimuli, their positions, and the order of sample presentation varied unsystematically across trials with restrictions (the same sample and the same correct comparison position occurred 2 times in succession, comparison stimuli appeared 3 times in succession). Correct MTS selections during pretraining and training conditions were followed by a display of colorful stars, a musical phrase, and the intertrial interval (1 s). Incorrect responses were followed by removal of all sample and comparison stimuli, a 3 s blackout, and the intertrial interval.

In the beginning of the first session, the experimenter sat next to the participant and provided verbal instructions in Portuguese during the initial trials: “Click here” (pointing to a stimulus presented alone in the first pretraining block); “Did you see the little stars? They appeared because your choice was correct!” When the task involved samples and comparisons, verbal instructions were: “click here (pointing to sample)”; “Did you see that two other pictures appeared? Now you have to choose one of them”. After the choice, the experimenter indicated the meaning of consequences. If the choice was not correct, for example, the experimenter said: “Did you see that the screen became dark and the little stars did not show up? It’s because your choice was not correct, let’s try again”.

Back-up reinforcement procedures involved individualized posters displayed nearby the data collection room. Each poster had a series of two-row tables. Upon reaching learning criterion for each session, the participant was permitted to affix a patch within a cell of a table. When two cells were filled, the participant could choose from an array of reinforcers (coloring materials, small toys, story books, school supplies, etc.).

Pretraining: During this phase of the procedures, we sought to develop an MTS baseline. Familiar stimuli were used initially and were followed by trials with abstract stimuli (Figure 1). Participants were taught to observe the sample and comparison stimuli on each trial and to respond to the comparison corresponding to the sample. The training approach was a simple-to-conditional discrimination procedure derived from the work of Saunders and Spradlin (1989, 1990, 1993). Briefly, simple discriminations were first established, followed by reversals of those discriminations and gradual transformation of the simple discriminations into conditional discriminations.

This procedure was used to establish two relations with familiar stimuli (sun/moon, pear/grapes) and two with abstract stimuli (PQ, QR). The sequence of presentation was eight consecutive trials each sample, then four consecutive trials each sample and then random presentation. Finally, two conditional discriminations with abstract stimuli were taught directly with random sequence of sample presentations (RS, ST). After each new relation was trained, previously trained discriminations were reviewed. Note that the sequence of introducing conditional relations was linear (i.e., one new relation introduced at a time). This approach was used in subsequent conditional discrimination training with the arbitrary stimuli that would be used to assess potentially emergent stimulus-stimulus relations and *select* and *reject* stimulus control topographies.

AB and BC Training: Figure 2 shows the stimuli and stimulus-stimulus relations established via direct AB and BC training in this condition. Conditional discriminations were separately trained in blocks of 16 trials, initially in blocks presenting four consecutive trials with each sample (criterion = a maximum of one error) and followed by training with samples in random sequence (criterion = two consecutive blocks with 1 error). After separate AB and BC training, conditional discrimination trials were intermixed (criterion = three blocks with 1 error). Thereafter, differential consequences for MTS trials were removed in preparation for probe trials (criterion = one block of 12 trials with 1 error). If participants did not meet criterion using this procedure, we presented blocks with eight consecutive trials with each sample, then with four consecutive trials, and then with random presentation. This procedure was repeated if a participant did not reach criterion or was unavailable for testing for several days in succession. If a given participant did not reach criterion on equivalence probes (described below) with stimulus Set 1 (Figure 2, upper portion), then the training and test procedures were repeated with different stimuli (Set 2; Figure 2, lower portion).

Equivalence Probes tested for potentially emergent stimulus-stimulus relations. Relations CA and AC were tested in separate blocks (not with intermixed baseline trials). Criterion for terminating testing was 1 class-inconsistent response in a block in two consecutive probe sessions. If this criterion was not attained, then (1) AB and BC (intermixed block) relations were re-trained using the same criterion as before and (2) CA and AC relations were then re-tested. This sequence was employed a maximum of three times. If stimulus equivalence was not demonstrated, then the entire training and probe sequence was repeated from the beginning with Stimulus Set 2 (as noted above).

Stimulus control topography (SCT) probes (Figure 3) were presented to evaluate whether participants had acquired both *select* and *reject* relations on AB and BC baseline trials. Novel stimuli replaced the positive comparison on four trials of an eight-trial block and the negative comparison on the other four trials. For example, if A1 was the sample and N1 (new stimulus) and B2 were comparison stimuli, the *reject* control would be inferred if the participant selected N1. As another example, if A1 was the sample and comparison stimuli were B1 and N2, *select* control would be inferred from selection of B1. The best evidence for *select* and *reject* relational control would be differential responding on both probe types, thus showing that the participant was not merely exhibiting bias towards defined or novel stimuli.

Results

Pretraining, AB and BC Training—During pretraining, all participants except RA completed all tasks rapidly. RA made numerous errors with the familiar stimuli, and was given supplemental training with the delayed cue procedure (Touchette, 1971). RA then rapidly acquired the pretraining relations with the abstract stimuli. During training, participants met criteria after 8 (the minimum) to 35 blocks (mean: 256 trials; range: 124-556). Four participants (AL, GU, VA, JO) required one or more phases of re-training on one or more of the training relations to meet the criterion necessary to move to the probe phase.

Equivalence Probes—Table 2 shows results on the equivalence probes. Participants GU, SA and VA did not meet criterion for equivalence relations after three presentations of probes with the first stimulus set. They met criterion on the second set. Participant GU showed emergent relations immediately, and SA and VA exhibited gradual emergence of equivalence relations. All other participants met criterion with the first set of stimuli, four rapidly (JA, RA, PA and AL) and one (JO) showing gradual emergence.

SCT probes—Figure 4 shows participants' performances on these probes for AB and BC relations. *Select* controlling relations would be inferred if participants consistently selected the S+ stimulus when it was displayed with a novel stimulus (Figure 4, black bars). *Reject* controlling relations would be inferred if participants consistently selected the novel stimulus when it was displayed with a defined S− (Figure 4, striped bars).

Behavior was somewhat variable on SCT probes. Although all participants showed class formation with the first or second set of stimuli, none showed perfection in *select* and *reject* SCTs on both AB and BC probes. Some (SA-set1, VA-set 2, JA and PA) showed such responding for one relation. Using a criterion of only one inconsistent selection per relation, however, performance of four participants (VA-set 2, JA, AL and PA) was consistent with both *select* and *reject* SCTs.

Experiment 2

This experiment used a blank comparison procedure aimed at *training* AB and BC *select* and *reject* relations. SCT probes were then conducted after equivalence probes as in Experiment 1. The main question was whether training aimed at promoting development

select and *reject* relations would result in more immediate, more consistent responding on novel-stimulus SCT probes and more reliable class formation. Subsequently, we repeated the training and probe procedures with additional CD relations to assess potential class expansion and associated SCT development.

Method

Participants, settings, apparatus, and criteria for programming blocks of trials were the same in Experiment 1. Data collection spanned 6-22 weeks for preschool children, and 6-9 weeks for participants with DS. Sessions lasted about 15 minutes for preschool children and about 30 minutes for participants with DS.

Procedures

Pretraining: The purpose was to establish a blank comparison MTS baseline, initially with familiar stimuli and then with abstract stimuli. The MTS relations with familiar stimuli from Experiment 1 (sun/moon, pear/grapes) were presented initially. Thereafter, two new relations were comprised of the former comparisons as samples and two other familiar stimuli as comparisons (moon/star, pear/strawberry). These relations were presented initially without the mask and then with stimulus control shaping to establish blank comparison matching. Next, the two previously established relations with abstract stimuli (PQ, QR) were reviewed without the mask and then the blank comparison shaping procedures were implemented with those stimuli. Thereafter, the other previously trained conditional discrimination (RS) was presented in the blank comparison format (i.e., with the mask appearing on all trials). Finally, two new relations (UV, VX) were established initially in the two-comparison format and then with 50% of trials in the blank comparison format. Stimuli used in this phase are shown in the right portion of Figure 1.

AB and BC training: Relations involving abstract pictures (two stimuli per set, see Figure 5) were taught. Initial training included only AB trials. When AB criterion was met, BC training commenced. When the BC criterion was met, AB and BC trials were intermixed. When criterion was met for the intermixed block, differential consequences were removed in preparation for probes. In all training blocks, 50% of trials were in the blank comparison format (see Figure 6). Blank comparison and two-comparison trials appeared in quasi-random order. Criterion at each phase (i.e., AB, BC and AB & BC) was three consecutive 16-trial blocks with 1 error. After criterion was met in a block with AB & BC trials, such blocks were repeated without differential consequences in preparation for probe sessions. In this phase of training, criterion was at least 11 correct selections in a 12-trial block.

CA and AC equivalence probes: Two-comparison probe trials (not intermixed with baseline trials) were presented in two separated 12-trial blocks of CA and AC trials. No programmed consequences followed responses. Probes ended after two consecutive blocks with 1 class-inconsistent responses. If that criterion was not met in any block, then a session of AB & BC trials was interposed before the next test session. If criterion was not met within three test sessions, then the AB and BC training and probe sequences were repeated with a second stimulus set. Three participants (GU, AL, and RA) received training and probes with a second set.

SCT probes had the same structure as in Experiment 1, replacing one comparison stimulus (S+ on some trials and S– on others) with novel stimuli in AB and BC relations.

Class expansion training: For participants who exhibited equivalence classes, CD relations were trained subsequently (also using the mask on 50% of trials in 16-trial blocks); criterion was three consecutive blocks with 1 error. Thereafter, AB, BC, and CD trials were intermixed in a 24-trial block with 50% of trials with the blank comparison. Criterion was three consecutive blocks with 2 errors per block. When criterion was met, such blocks were presented without programmed consequences in preparation for probes (criterion: 1 error).

Class expansion and SCT probes: Separate 12-trial blocks in the two-comparison format tested for emergent equivalence relations: DA/AD/CA/AC/DB/BD. Thereafter, SCT probes were presented in 8-trial blocks with each of the trained relations (AB/BC/CD).

Results and Discussion

Pretraining—Participants initially made errors when the blank comparison procedure was used on all trials with pretraining tasks despite previously exhibited high, stable accuracy on two-comparison versions of those tasks. That outcome led to introducing the blank comparison procedure gradually across tasks and to maintaining a 50%-50% mixture of two-comparison and blank comparison trials throughout training.

AB and BC Training—Participants met acquisition criteria after 10 (the minimum) to 27 trial blocks (mean: 237 trials; range: 156-428). AL's training sequence was slightly different from that described earlier in that all training blocks were initially two-comparison trials and the 50%-50% mixture was implemented thereafter. PA also had a slightly different training sequence in that he received a repeat of the training/test series with the first stimulus set after failing to exhibit equivalence classes in initial testing.

CA and AC Equivalence—Results are shown in Table 3. PA showed equivalence class formation immediately. SA, JO and AL met criterion after three presentations of these probes and JA after four. GU and VA showed immediate emergence after training with Set 2. RA met criterion after three presentations of equivalence probes with Set 2.

SCT Probes—Results are shown in Figure 7. One preschool child (VA-Set 2) and two participants with DS (PA and RA-Set 1) showed perfect consistency with acquisition of *select* and *reject* relations during AB and BC training. Two participants made only one inconsistent selection on each of two relations: JO (one on AB *select* and one on AB *reject* probes) and SA (one on AB *reject* and one on BC *select* probes).

Probe performances of GU, AL, JA and RA showed greater inconsistency with acquisition of *select* and *reject* SCTs. GU showed a bias towards selecting the novel stimuli regardless of which SCT the probe trial tested for. AL showed virtually no differential responding on the SCT probes. RA showed consistency with acquisition of *select* and *reject* control after Set 1 but a novel-stimulus bias after Set 2.

Relationship between class formation and SCT probes—Results with four participants (PA, VA, JO, SA) seem perfectly or largely consistent with our prediction that positive class formation results would be accompanied by positive results on SCT probes. By contrast, GU showed an immediately positive equivalence-test outcome with Set 2 and AL did so (gradual emergence) with Set 1; neither showed consistent performance on SCT probes. Highly discrepant results were obtained with RA. With Set 1, RA showed SCT probes consistent with both *select* and *reject* SCTs but not equivalence relations. With Set 2, the equivalence test was positive (gradual emergence), but there was little evidence of *select/reject* relations on SCT probes (most responses were selections of novel stimuli). JA showed delayed emergence of classes and consistent SCT probes only for AB relations.

One consideration in evaluating the results with GU and RA is whether their equivalence-probe outcomes were artifacts of a two-choice training/test procedure (cf. Sidman, 1987). Note that highly above-chance scores on the Set-2 equivalence tests were preceded by highly *below-chance* scores on the Set-1 tests. With this pattern, one cannot rule out so-called generalized conditional responding (aka “arbitrary assignment,” K. Saunders & Spradlin, 1990, 1993). Thus, one might speculate these Set-2 equivalence-test outcomes were false positives. If so, there would be no logical reason for asserting that equivalence and SCT probe performances should be highly correlated.

Class Expansion Training—All participants except AL met criterion after seven (the minimum) to eight blocks. AL met criterion after 24 blocks, making errors mainly on CD trials on their first presentations in intermixed blocks; this led to separate re-training of the CD relation and re-exposure to intermixed trial blocks.

Class Expansion Probes—Table 4 shows that GU, JO, PA and AL showed immediately accurate emergent performances. JA showed class expansion after three probe presentations. Although VA and RA responded reasonably accurately on expansion probes by the third presentation, they did not meet criterion until a fourth (data not shown). We consider these results positive for class expansion. Results with SA were negative; she exhibited merely generalized conditional responding.

SCT Probes—Figure 8 shows that six participants (PA, AL, GU, VA, JO and SA) exhibited performances consistent with *select* and *reject* relations. PA showed such controlling relations on all 24 SCT probe trials. AL, GU, and VA made a single inconsistent response (96% consistency), and JO, RA, and SA made two inconsistent responses (92% consistency, the last of these perhaps via generalized conditional responding as on the equivalence probes). Only JA exhibited clearly inconsistent SCT probe results.

General Discussion

Our study sought to assess equivalence class formation following MTS training in two conditions: typical two-comparison MTS (Experiment 1) and blank comparison MTS on 50% of training trials (Experiment 2). We sought also to assess the utility of novel-stimulus probes in evaluating *select* and *reject* SCTs established by in training. During both experiments, all participants showed emergent relations (CA/AC) with at least one set of

stimuli, immediately or after repeated training and testing. In Experiment 2, seven of eight participants exhibited class expansion, three virtually immediately and four during repeated testing. Fewer participants showed immediate class formation in Experiment 2. However, participants needed fewer retests than in Experiment 1 to show class emergence. Thus, repeated training and testing appeared to establish the context for equivalence class formation (Sidman, 1990), as has happened in other studies of visual-visual MTS with preschoolers (e.g., Lazar et al., 1984; Smeets & Barnes Holmes, 2005) and persons with intellectual disabilities (e.g., Dube & McIlvane, 1995). In general, our outcomes seemed somewhat better than typical for participants with developmental limitations (e.g., Augustson & Dougher, 1991; O'Donnell & Saunders, 2003; Pilgrim et al., 2000), particularly given the visual-visual MTS procedures (see Green, 1990; Stromer & Mackay, 1996).

To evaluate the possibility that *select* and *reject* relations had been established during baseline training, we used novel stimulus probes similar to those used by Stromer and Osborne (1982). Recall that they found their participants with mild intellectual disabilities exhibited both *select* and *reject* relations on all probes. By contrast, our Experiment 1 results did not replicate those of Stromer and Osborne (1982) – even when *select* and *reject* probes were conducted after equivalence classes were verified. Also noteworthy is that novel stimulus probe results were inconsistent initially in Experiment 2 even after blank comparison training designed to promote development of *select* and *reject* SCTs. However, SCT probe results did become more consistent after class expansion.

The results of both experiments show that performance in SCT probes need not be highly correlated with equivalence outcomes. We found negative performances on SCT probes for participants whose equivalence class formation results were positive (Experiment 1: SA-Set 2, GU-Set 2, JO and RA; and for Experiment 2, JA after two-node training). We also obtained positive outcomes on SCT probes for participants who did not show class formation (Experiment 1: SA-Set 1 [BC relation], Experiment 2: RA-Set 1 [probes after one-node training], Experiment 2: SA [probes after two-nodes training]).

Based on these findings, one might be tempted to declare that equivalence class formation and development of *select/reject* SCTs are independent phenomena – not strongly related as suggested by Stromer and Osborne (1987) and McIlvane et al. (1987). Before doing so, however, we must consider the possibility that some of the present findings were artifacts of the two-choice test procedures and their possibilities for promoting generalized conditional responding as already noted.

We must consider also possible interactions between procedure and participant variables. Recall that Stromer and Osborne studied adolescents (mental ages [MAs] = ~11 years on average estimated from CA and IQ data in their Table 1). By contrast, our participants had much lower MAs, especially the preschool children and the child with DS (range: 3-5 years). Extensive literature shows that the novelty dimension of stimuli may be highly salient for low-MA individuals (cf. Stevenson, 1973) and may be preferred over familiar stimuli in discrimination learning procedures (Zeaman, 1976). Based on published data (e.g., Valenti, 1985), most of our participants would be expected to show such novelty

preferences. If so, these preferences could have emerged under the programmed probe contingencies (i.e., extinction) and led to variable probe-trial performance. One might predict also that extended experience with the procedures would render the introduction of novel probe stimuli a relatively less novel event. If so, probe performances might become more consistent – an outcome that we found.

In future research, it would be useful to explore whether we could use the blank-comparison procedure for SCT tests rather than merely for training purposes. Had we taken the additional step of developing and maintaining a highly accurate and reliable blank comparison *identity* MTS procedure with the experimental stimuli, for example, then we could have presented blank-comparison SCT probes in the arbitrary matching format (cf. McIlvane et al., 1984, 1987). By doing that, we would avoid introducing novel stimuli merely for SCT test purposes and perhaps obtain stronger correlations of outcomes of those of probes with those assessing equivalence class formation.

Acknowledgments

This research here was a bi-national collaborative effort through the Instituto Nacional de Ciência e Tecnologia Sobre Comportamento Cognição e Ensino (National Institute for Science and Technology: Studies of Behavior, Cognition, and Teaching) coordinated by Universidade Federal de São Carlos, SP, Brazil (Deisy de Souza, Project Director). Support for this work came from multiple sources including: CNPq (Grant# 573972/2008-7), FAPESP (Grant# 2008/57705-8), NICHD (HD04147), and NIMH (MH90272). The first author was supported by a doctoral scholarship (FAPESP, Grant# 2007/50992-9), and the second author by a research productivity grant from CNPq. Studies reported here were part of a doctoral dissertation submitted by the first author to Programa de Pós-Graduação em Educação Especial, Universidade Federal de São Carlos.

References

- Arntzen E, Grondahl T, Eilifsen C. The effect of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *The Psychological Record*. 2010; 60:437–462.
- Arntzen E, Holth P. Probability of stimulus equivalence as a function of training desing. *The Psychological Record*. 1997; 47:309–320.
- Augustson KG, Dougher MJ. Teaching conditional discrimination to young children: Some methodological successes and failures. *Experimental Analysis of Human Behavior Bulletin*. 1991; 9:21–24.
- Barnes-Holmes D, Staunton C, Whelan R, Barnes-Holmes Y, Commings S, Walsh D, Stewart I, Smeets PM, Dymond S. Derived stimulus relations, semantic priming, and event-related potentials: testing a behavioral theory of semantic networks. *Journal of the Experimental Analysis of Behavior*. 2005; 84:417–433. doi: 10.1901/jeab.2005.78-04. [PubMed: 16596973]
- Bortoloti R, de Rose JC. Assessment of the relatedness of equivalent stimuli through a semantic differential. *The Psychological Record*. 2009; 59:563–590.
- Bortoloti R, de Rose JC. Equivalent stimuli are more strongly related after training with delayed matching than after simultaneous matching: A study using the Implicit Relational Assessment Procedure (IRAP). *The Psychological Record*. 2012; 62:41–54.
- Carr D, Wilkinson KM, Blackman D, McIlvane WJ. Equivalence classes in individuals with minimal verbal repertoires. *Journal of the Experimental Analysis of Behavior*. 2000; 74:101–114. doi: 10.1901/jeab.2000.74-101. [PubMed: 10966098]
- Carrigan PF, Sidman M. Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior*. 1992; 58(1):183–204. doi: 10.1901/jeab.1992.58-183. [PubMed: 16812663]

- Cumming, WW.; Berryman, R. The complex discriminated operant: Studies of matching to sample and related problems. In: Mostofski, DI., editor. *Stimulus generalization*. Stanford University Press; Stanford, CA: 1965. p. 284-329.
- de Rose, JC. Controlling factors in conditional discriminations and tests of equivalence. In: Zentall, TR.; Smeets, PM., editors. *Stimulus class formation in humans and animals*. North Holland (Elsevier); Amsterdam: 1996. p. 253-277.
- de Rose JC, Hidalgo M, Vasconcellos M. Controlling relations in baseline conditional discriminations as determinants of stimulus equivalence. *The Psychological Record*. 2013; 63:85–98.
- Devany JM, Hayes SC, Nelson RO. Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior*. 1986; 46:243–257. doi: 10.1901/jeab.1986.46-243. [PubMed: 3805973]
- Dube WV. Computer software for stimulus control research with Macintosh computers. *Experimental Analysis of Human Behavior Bulletin*. 1991; 9:28–39.
- Dube WV, McIlvane WJ. Stimulus-reinforcer relations and emergent matching to sample. *The Psychological Record*. 1995; 45:591–612.
- Dube, WV.; McIlvane, WJ. Some implications of a stimulus control topography analysis for emergent behavior and stimulus class. In: Zentall, TR.; Smeets, PM., editors. *Stimulus class formation in human and animals*. North Holland (Elsevier); Amsterdam: 1996. p. 197-218.
- Dunn, LM.; Dunn, LM. *Peabody Vocabulary Test -Revised*. American Guidance Service; Circle Pines, MN: 1981.
- Fields L, Adams BJ, Verhave T, Newman S. The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior*. 1990; 53:345–358. doi: 10.1901/jeab.1990.53-345. [PubMed: 2341819]
- Green G. Differences in development of visual and auditory-visual equivalence relations. *American Journal of Mental Retardation*. 1990; 95:260–270. [PubMed: 2261159]
- Haimson B, Wilkinson KM, Rosenquist C, Ouimet C, McIlvane WJ. Electrophysiological correlates of stimulus equivalence processes. *Journal of the Experimental Analysis of Behavior*. 2009; 92:245–256. doi: 10.1901/jeab.2009.92-245. [PubMed: 20354602]
- Hayes, SC.; Barnes-Holmes, D.; Roche, B., editors. *Relational Frame Theory: A Post-Skinnerian account of human language and cognition*. Plenum Press; New York: 2001.
- Horne PJ, Lowe CF. On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*. 1996; 65:185–241. doi: 10.1901/jeab.1996.56-185. [PubMed: 16812780]
- Kato OM, de Rose JC, Faleiros PB. Topography of responses in conditional discrimination influences formation of equivalence classes. *The Psychological Record*. 2008; 58:245–267.
- Lazar RM, Davis-Lang D, Sanches L. The formation of visual stimulus equivalences in children. *Journal of the Experimental Analysis of Behavior*. 1984; 41:251–266. doi: 10.1901/jeab.1984.41-251. [PubMed: 6736856]
- Lipkens G, Hayes SC, Hayes LJ. Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology*. 1993; 56:201–239. [PubMed: 8245768]
- McIlvane, WJ. Simple and complex discrimination learning. In: Madden, GJ., editor. *APA Handbook of Behavior Analysis*. American Psychological Association; Washington, DC: 2012. p. 129-163.
- McIlvane WJ, Dube WV. Stimulus control shaping and stimulus control topographies. *The Behavior Analyst*. 1992; 15:89–94. [PubMed: 22478119]
- McIlvane WJ, Dube WV. Stimulus control topography coherence theory: foundations and extensions. *Behavior Analyst*. 2003; 26:195–213. [PubMed: 22478402]
- McIlvane WJ, Kledaras JB, Munson LC, King KAJ, de Rose JC, Stoddard LT. Controlling relations in conditional discrimination and matching by exclusion. *Journal of the Experimental Analysis of Behavior*. 1987; 48:187–208. doi: 10.1901/jeab.1987.48-187. [PubMed: 3681184]
- McIlvane, WJ.; Serna, RW.; Dube, WV.; Stromer, RL. Stimulus control topography coherence and stimulus equivalence: Reconciling test outcomes with theory. In: Leslie, J.; Blackman, DE., editors. *Experimental and applied analysis of human behavior*. Context Press; Reno NV: 2000. p. 85-110.

- McIlvane WJ, Withstandley JK, Stoddard LT. Positive and negative stimulus relations in severely retarded individuals' conditional discrimination. *Analysis and Intervention in Developmental Disabilities*. 1984; 4:235–251.
- O'Donnell J, Saunders KJ. Equivalence relations in individuals with language limitations and mental retardation. *Journal of the Experimental Analysis of Behavior*. 2003; 80:131–157. doi: 10.1901/jeab.2003.80-131. [PubMed: 13677612]
- Pilgrim C, Jackson J, Galizio M. Acquisition of arbitrary conditional discriminations by young normally developing children. *Journal of the Experimental Analysis of Behavior*. 2000; 73:177–193. doi: 10.1901/jeab.2000.73-177. [PubMed: 10784008]
- Saunders K, Spradlin JE. Conditional discrimination in mentally retarded adults: The effect of training the component simple discriminations. *Journal of the Experimental Analysis of Behavior*. 1989; 52:1–12. doi: 10.1901/jeab.1989.52-1. [PubMed: 2769172]
- Saunders K, Spradlin JE. Conditional discrimination in mentally retarded adults: The development of generalized skills. *Journal of the Experimental Analysis of Behavior*. 1990; 54:239–250. doi: 10.1901/jeab.1990.54-239. [PubMed: 2103584]
- Saunders K, Spradlin JE. Conditional discrimination in mentally retarded subjects: Programming acquisition and learning set. *Journal of the Experimental Analysis of Behavior*. 1993; 60:571–585. doi: 10.1901/jeab.1993.60-571. [PubMed: 8283149]
- Saunders RR, Drake KM, Spradlin. Equivalence class establishment, expansion and modification in preschool children. *Journal of the Experimental Analysis of Behavior*. 1999; 71:195–214. doi: 10.1901/jeab.1999.71-195. [PubMed: 10220930]
- Saunders RR, McEntee JE. Increasing probability of stimulus equivalence with adults with mild mental retardation. *The Psychological Record*. 2004; 54:423–435.
- Sidman M. Remarks. *Behaviorism*. 1979; 7:123–126.
- Sidman, M. Functional analysis of emergent verbal classes. In: Thompson, T.; Zeiler, M., editors. *Analysis and Integration of Behavioral Units*. Erlbaum; Hillsdale, NJ: 1986. p. 213-245.
- Sidman, M. Equivalence relations: Some basic considerations. In: Hayes, LR.; Hayes, SC., editors. *Dialogues on verbal behavior: Proceedings of the Third International Institute on Verbal Relations*. Context Press; Reno, NV: 1990.
- Sidman, M. Equivalence relations and behavior: A research history. Authors Cooperative; Boston, MA: 1994.
- Sidman M. Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*. 2000; 74:127–146. doi: 10.1901/jeab.2000-74-127. [PubMed: 10966100]
- Sidman M, Tailby W. Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*. 1982; 37:5–22. doi: 10.1901/jeab.1982.37-5. [PubMed: 7057129]
- Smeets PM, Barnes-Holmes D. Establishing equivalence classes in preschool children with one-to-many and many-to-one training protocols. *Behavioral Processes*. 2004; 69:281–293.
- Smeets PM, Barnes-Holmes D. Auditory-visual and visual-visual equivalence relations in children. *The Psychological Record*. 2005; 55:483–503.
- Stevenson, HW. *Children's learning*. Appleton-Century-Crofts; New York: 1973.
- Stromer, R.; Mackay, HA. Naming and the formation of stimulus classes. In: Zentall, TR.; Smeets, PM., editors. *Stimulus class formation in humans and animals*. North-Holland; Amsterdam: 1996.
- Stromer R, Osborne JG. Control of adolescents' arbitrary matching-to-sample by positive and negative stimulus relations. *Journal of the Experimental Analysis of Behavior*. 1982; 37:329–348. doi: 10.1901/jeab.1982.37-329. [PubMed: 6178786]
- Touchette PE. Transfer of stimulus control: Measuring the moment of transfer. *Journal of the Experimental Analysis of Behavior*. 1971; 15:347–354. doi: 10.1901/jeab.1971.15-347. [PubMed: 4252714]
- Valenti SS. Children's preference for novelty in selective learning: Developmental stability or change? *Journal of the Experimental Child Psychology*. 1985; 40:406–419.
- Zeaman, D. The ubiquity of novelty-familiarity (habituation) effects. In: Tighe, TJ.; Leaton, RN., editors. *Habituation: Perspectives from child development, animal behavior, and neurophysiology*. Erlbaum; Hillsdale, NJ: 1976.

Figure 1.
Stimuli used during pretraining phases. Left and right columns show those introduced in Experiment 1 and 2, respectively.

Figure 2.

Training trial types for each set of stimuli in Experiment 1. Each stimulus triad represents one trial. Stimuli at the top are samples and pairs at the bottom are comparisons.

Figure 3.

Stimulus Control Topography probe trial types. Letter N indicates novel stimulus replacing a baseline comparison stimulus. Sample stimuli are at the top of each triad of figures and comparison stimuli, at the bottom.

Figure 4.

Number of responses of each participant to defined (S+/S-) or novel comparison stimuli on SCT Probes in Experiment 1. Black and striped bars indicate *select* (sample/S+) and *reject* (sample/S-) trial types, respectively.

Figure 5. Stimuli used in Experiment 2. Letters (A, B, C D) indicate stimulus sets and numbers (1, 2) indicate potential stimulus classes to be established in training.

Figure 6.

Trials types presented during training in Experiment 2. The plus sign (+) indicates stimuli defined as correct. AB Set 2 stimuli have been used as a model of trials employed in training of all relations.

Figure 7.

Number of responses of each participant to defined (S+/S-) or novel comparison stimuli on SCT Probes administered after one-node equivalence tests for Experiment 2. Black and striped bars indicate *select* and *reject* trial types, respectively.

Figure 8.

Number of responses of each participant to defined (S+/S-) or novel comparison stimuli on SCT Probes administered after class expansion tests for Experiment 2. Black and striped bars indicate *select* and *reject* trial types, respectively.

Table 1

Participant characteristics: identification (ID), gender (male [M] or female [F]), chronological age (CA in years-months) at start of the experiment, PPVT age-equivalent score, IQ from WISC III (child w/ DS) or WAIS III (adults w/ DS).

ID	M/F	CA	PPVT	WISC III	WAIS III
Preschool children					
GU	M	5 - 2	5 - 9		
VA	M	5 - 2	5 - 8		
JO	M	5 - 1	5 - 4		
SA	F	5 - 6	5 - 8		
JA	F	5 - 1	5 - 5		
Down Syndrome					
RA	M	8 - 6	3 - 8	64	
PA	M	19 - 9	9 - 4		60
AL	F	33 - 10	8 - 8		57

Table 2

Percentage of correct responses in all presentations of stimulus equivalence probes for each participant in Experiment 1.

Participants	Probe Presentations											
	1st.		2nd.		3rd.		4th.					
	CA	AC	CA	AC	CA	AC	CA	AC	CA	AC	CA	AC
GU Set 1	75	75	67	50	75	67						
GU Set 2	100	100	100	100								
SA Set 1	50	50	67	17	50	75						
SA Set 2	58	67	58	67	83	92	100	100				
VA Set 1	58	0	67	50	25	50						
VA Set 2	83	67	83	75	92	100	100	100				
JO	67	58	88	67	92	100	92	83*				
JA	100	100	100	100								
RA	100	100	100	100								
PA	100	100	100	100								
AL	100	100	100	100								

* This participant met criterion on the sixth presentation of AC and CA probes (data not shown).

Table 3

Percentage of correct responses for each presentation of one node equivalence probes for each participant in Experiment 2.

	Probe Presentations											
	1st			2nd			3rd					
	CA	AC	CA	CA	AC	AC	CA	AC	CA	CA	AC	AC
GU Set 1	17	0	17	0	17	0	33	50				
GU Set 2	100	100	100	100	100	100						
SA	75	58	100	92	100	100						
VA Set	0	0	100	42	8	0						
VA Set 2	100	92	92	100								
JO	100	67	100	100	100	92						
JA	92	83	92	83	92	100						
RA Set 1	0	0	0	0	25	0						
RA Set 2	50	75	100	100	100	100						
PA	100	100	100	100								
AL	83	92	100	100	100	100						

Table 4
Percentage of correct responses for each presentation of class expansion probes for each participant.

	Probe Presentations											
	1st.				2nd.				3rd.			
	DA	AD	CA	AC	DB	BD	DA	AD	CA	AC	DB	BD
GU	100	100	100	100	100	100	100	100	100	100	100	100
SA	92	100	100	100	75	58	0	0	100	100	0	0
VA*	100	100	92	83	100	100	92	92	92	100	92	100
JO	100	92	100	100	100	100	100	100	100	100	100	100
JA	83	92	100	92	100	100	100	92	92	92	92	100
RA*	58	58	66	100	100	100	92	100	83	100	100	100
PA	100	100	100	100	100	100	100	92	100	100	100	100
AL	100	100	92	100	100	92	100	100	100	100	100	100

* These participants met criterion on the fourth presentation of AC and CA probes (data not shown).