The Psychological Record, 2010, 60, 659–678

TRAINED AND DERIVED RELATIONS WITH PICTURES VERSUS ABSTRACT STIMULI AS NODES

Erik Arntzen and Torunn Lian

Akershus University College

Earlier studies have shown divergent results concerning the use of familiar picture stimuli in demonstration of equivalence. In the current experiment, we trained 16 children to form three 3-member classes in a many-to-one training structure. Half of the participants were exposed first to a condition with all abstract stimuli and then to a condition with new abstract stimuli as samples and 3 picture stimuli as comparisons (and nodes). The other participants were given the 2 conditions in the reverse order. The results, regardless of order, showed that the condition with picture stimuli as nodes was more effective in producing responding in accord with equivalence than stimulus sets with abstract stimuli only. In addition, more participants responded in accord with equivalence when they were trained with picture stimuli first. Reaction time to the comparison stimuli showed a greater increase with abstract stimuli than with pictures as nodes.

Key words: stimulus equivalence, pictures, familiar stimuli, children, nodes

Stimulus equivalence is defined as responding after preliminary training of conditional discriminations in accord with novel, unreinforced tests of reflexivity, symmetry, and transitivity (e.g., Sidman & Tailby, 1982). When stimulus equivalence is achieved, the stimuli within an equivalence class can be described as mutually substitutable (Green & Saunders, 1998). Stimulus equivalence has been demonstrated with both verbally competent people (Dugdale & Lowe, 2000), including adults and typically developing children (e.g., Arntzen & Vaidya, 2008; Lipkins, Hayes, & Hayes, 1993; Pilgrim, Chambers, & Galizio, 1995; Sidman & Tailby, 1982), and those with developmental disabilities or autism (Arntzen, Halstadtro, Bjerke, & Halstadtro, 2010; LeBlanc, Miguel, Cummings, Goldsmith, & Carr, 2003). It has also been demonstrated after training with different stimulus modalities such as olfactory (Annett

The current study was presented as a paper at the Association for Behavior Analysis International conference in Chicago in 2008.

The authors are very grateful to Dag-Erik Eilertsen for help with the statistical analyses and to Lill-Beathe Halstadtro for assistance in organizing the lab facilities in one of the laboratories. Furthermore, we want to thank three anonymous reviewers and David Dickins for valuable comments to earlier versions of the article.

Correspondence concerning this article should be addressed to Erik Arntzen, Akershus University College, PO Box 423, 2001 Lillestrom, Norway. E-mail: erik.arntzen@equivalence.net

& Leslie, 1995), haptic (Belanich & Fields, 1999), tactile (O'Leary & Bush, 1996), and gustatory (L. J. Hayes, Tilley, & Hayes, 1988). Furthermore, stimulus equivalence has been demonstrated with a variety of visual stimulus materials, such as different abstract stimuli (e.g., Sidman & Tailby, 1982), pictures (e.g., Arntzen, 2004), consonant–vowel–consonant syllables (CVCs; e.g., Fields et al., 1997), and three-dimensional objects (e.g., Devany, Hayes, & Nelson, 1986). Also, stimulus equivalence has been shown after simultaneous matching and delayed matching to sample with different retention intervals (Arntzen, 2006; Arntzen, Galaen, & Halvorsen, 2007). Furthermore, some studies have shown that certain stimuli with which the participants have a learning history could have an inhibitory effect on equivalence class formation (Holth & Arntzen, 1998b; Leslie, Tierney, Robinson, Keenan, & Watt, 1993; Plaud, 1995). Hence, as pointed out by Sidman (1992), in conditional discrimination experiments derived stimulus relations may arise of reasons other than the experimental conditions explicitly arranged.

Three different training structures have been used in training conditional discriminations: one-to-many (OTM), many-to-one (MTO), and linear series (LS). In OTM, one sample is trained in relation to at least two comparisons, while in MTO at least two samples are trained in relation to one comparison. In LS, one sample is trained first to one comparison, and then the comparison is trained as a sample to another comparison (Fields & Verhave, 1987). The differential effects of the different training structures have been discussed (e.g., Arntzen, Grondahl, & Eilifsen, 2010), but because the MTO training structure has in some cases been demonstrated to be the most effective training structure with children (Arntzen & Vaidya, 2008; Saunders, Wachter, & Spradlin, 1988; Spradlin & Saunders, 1986), we used the MTO structure in the current study.

Some studies have shown that the prerequisites for stimulus equivalence and the formation of equivalence classes are established more readily with the use of pictures as at least one of the stimulus sets (Arntzen, 2004; Holth & Arntzen, 1998a). For example, Holth and Arntzen (1998a) explored the effects of using abstract stimuli and familiar picture stimuli on responding in accord with equivalence. In Experiment 1 of that study, 50 adult participants were exposed to five different combinations of Greek letters and picture stimuli. The participants were trained in an LS training structure. For one group the stimulus set included Greek letters only; for the other four groups the stimulus sets included Greek letters and picture stimuli (i.e., pictures as either B or C stimuli or as both A and C stimuli). The results showed that the stimulus set of Greek letters only gave the lowest probability of responding in accord with equivalence and that the probability of responding in accord with equivalence increased significantly when picture stimuli served as nodes (B stimuli) or as A and C stimuli. Arntzen (2004) investigated differential probabilities of responding in accord with equivalence as a function of the position of picture stimuli and nonsense syllables in three 5-member classes with an MTO training structure. Fifty adult students were assigned to five experimental groups: (a) Greek and Arabic letters only, (b) pictures as A stimuli, (c) pictures as E stimuli, (d) CVCs as A stimuli, or (e) Greek and Arabic letters only with keyboard touch as the response requirement, which differed from the other groups in that they had a mouse click as the response requirement. The results showed that all participants with pictures as A stimuli responded in accord with equivalence and that there was a significantly lower probability of responding in accord with equivalence in all the other groups. The results from these studies indicate that the use of picture stimuli is effective in establishing responding in accord with equivalence, especially when the picture stimuli are introduced first.

However, other researchers have presented inconclusive data and have made contradictory conclusions regarding the effect of familiar picture stimuli in establishing responding in accord with equivalence. Smeets and Barnes-Holmes (2005) explored the probability of responding in accord with equivalence as a function of abstract stimuli and familiar picture stimuli as nodes in two different training structures (OTM and MTO). Sixteen 5-year-old children were trained to form two 5-member classes. The results showed that experimental conditions involving abstract stimuli produced responding in accord with equivalence more readily than conditions with familiar picture stimuli.

Experimental variances and contradictory results, in addition to the conclusions described here, call for further analysis and experimental investigation. Pursuing this research path could be important in at least two ways. First, training procedures in stimulus equivalence research can sometimes be time consuming. Thus, determining which training procedures more effectively establish conditional discriminations could ease future research. Second, research in this area could be of interest in educational settings (Cautilli, Hancock, Thomas, & Tillman, 2002; Stromer, Mackay, & Stoddard, 1992).

Both number of trials to criterion and reaction time to comparison stimuli could be important dependent variables, even if indices of derived relations have been the most critical recording. Reaction-time data are sometimes used to evaluate whether it is reasonable to infer that choice responses during test conditions are under direct control of the sample stimulus. Hence, Dymond and Rehfeldt (2001) suggested using supplementary measures in the study of derived relations to shed light on the nature of derived relations and the variables that control the occurrence of such behavior. It is important to mention that one difficulty here is the fact that reaction time is a sensitive measure and could easily be influenced by any stimulus event or by another response, such as looking at a watch, sneezing, and other events irrelevant to the task.

The purpose of the current experiment was to study in typically developing children (a) if the different types of stimulus sets, with and without familiar picture stimuli as nodes, influenced the number of trials to criterion during training of conditional discrimination and (b) the effects of responding in accord with equivalence as a function of different stimulus sets (abstract vs. familiar). Also, we wanted to study the differences in number of trials and equivalence formation between participants who were exposed to a condition with familiar picture stimuli before a condition with the abstract stimuli and participants who were exposed to the reversed order. Furthermore, we wanted to study the differences in reaction times from training to test for the abstract- versus familiar-picture-stimuli condition, as well as any differences between symmetry and equivalence tests.

Method

Participants

Sixteen typically developing children, ages 7 to 12 years, were voluntarily recruited through personal contacts (details about the participants are shown in Table 1). The participants were nine boys and seven girls. They had no former experience as participants in an experiment. When recruited, the parents and children were told that the experiment was about learning and that they would be presented with stimuli presented on a computer. The parents filled out a consent form before the start of the experiment. The parents and children were also told that the approximate duration of the experiment was 3 hours, that the actual length depended on how correctly they responded, and that the children would be offered some breaks. They were told that they could withdraw from the experiment at any time, and that the experimenter was not going to provide any cues or instructions after the onset of the experiment. When the experimental session was finished, each participant was thanked and debriefed. They were given a small gift, such as a comic book or a set of stickers, as a reward for participating.

Table 1
Participants, Age, Gender, and Training Order

Participant	Age	Gender	Training Order	
1351	12-6	F	Abstract stimuli	Familiar picture stimuli
1353	10-0	M	Abstract stimuli	Familiar picture stimuli
1355	11-3	M	Abstract stimuli	Familiar picture stimuli
1356	8-7	M	Abstract stimuli	Familiar picture stimuli
1358	10-9	M	Abstract stimuli	Familiar picture stimuli
1380	9-0	M	Abstract stimuli	Familiar picture stimuli
1381	8-9	F	Abstract stimuli	Familiar picture stimuli
1383	9-0	M	Abstract stimuli	Familiar picture stimuli
1352	7-10	F	Familiar picture stimuli	Abstract stimuli
1359	8-6	M	Familiar picture stimuli	Abstract stimuli
1360	8-2	M	Familiar picture stimuli	Abstract stimuli
1371	9-7	F	Familiar picture stimuli	Abstract stimuli
1375	10-8	F	Familiar picture stimuli	Abstract stimuli
1376	10-7	F	Familiar picture stimuli	Abstract stimuli
1377	10-11	F	Familiar picture stimuli	Abstract stimuli
1379	9-1	M	Familiar picture stimuli	Abstract stimuli

Note. "Abstract stimuli" means that all nine stimuli were abstract stimuli, while "familiar picture stimuli" means that three of the stimuli were familiar picture stimuli (the nodes) and six were abstract stimuli.

Setting and Apparatus

The experimental sessions were conducted individually in three different laboratories, using a software program for the training and testing of conditional discriminations made by Psych Fusion Software and developed in collaboration with the first author. The software was presented on two Compaq nc6320 personal computers with 1828-MHz Intel Centrino® processors.

Stimuli

The stimuli used in the current experiment were abstract and familiar picture stimuli as shown in Figure 1 and Figure 2. The abstract stimuli were printed in black and the picture stimuli in color, both on a white background.

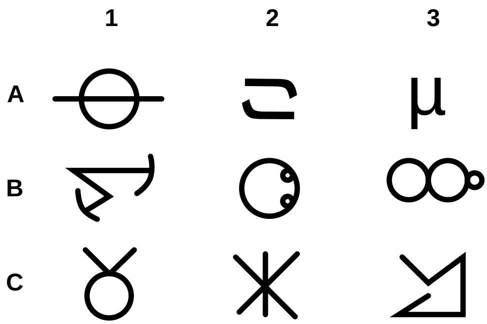

Figure 1. Stimulus set with abstract stimuli only. The numbers indicate the different classes, while the letters indicate the members in the classes. The C stimuli were always the nodes.

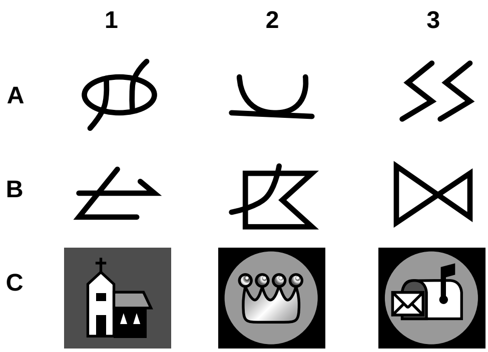

Figure 2. Stimulus set with picture stimuli as nodes. The numbers indicate the different classes, while the letters indicate the members in the classes. The C stimuli were always the nodes.

Procedure

The training structure was an MTO (AC, BC) structure in which the C stimuli served as nodes in both stimulus sets. The participants were randomly assigned to two different sequences: (a) abstract stimuli as nodes followed by familiar picture stimuli as nodes and (b) familiar picture stimuli as nodes followed by abstract stimuli as nodes. Each participant finished the experiment within a day (see Table 1).

Instructions. At the beginning of the experiment, the following instructions in Norwegian were displayed on the screen and read aloud by the experimenter:

> You are going to do several tasks on the computer. Use the mouse to click the stimulus in the middle of the screen. Once you have clicked the stimulus, several other stimuli will appear. Choose one stimulus and click on it. If you click the correct one, the word "correct" will appear on the screen. If you click on an incorrect stimulus the word "incorrect" will appear on the screen. After several trials neither "correct" nor "incorrect" will show up on the screen when you click a sign. It is important to get as many correct as possible. Good luck!

No further instructions or cues were given during the experimental sessions.

Baseline training and testing. Each trial started with the presentation of a sample stimulus in the middle of the screen. A mouse click to the sample stimulus was followed by the presentation of comparison stimuli. Comparison stimuli appeared in random position in three corners of the screen. While baseline relations were being established, a correct comparison choice was followed by the word "correct" and an incorrect comparison choice was followed by the word "incorrect" displayed in the middle of the screen. The feedback duration was 2 s and the intertrial interval was 1 s. The mouse marker position was reset above the sample stimulus between each trial. After incorrect comparison choices the particular baseline relation was repeated. Six relations were trained during baseline in an MTO training structure.

At the beginning of training, baseline relations were introduced one by one (as shown in Table 2). Hence, to minimize the number of incorrect choices initially, the number of comparison stimuli was gradually increased from one to three. For the first block a minimum of nine trials were presented, with A1C1, A2C2, and A3C3 presented three times each, followed by a block with a minimum of nine trials with B1C1, B2C2, and B3C3 presented three times each. In the third block, the trial types were mixed, with a minimum of 18 trials. In the fourth block, the trial types were mixed with a minimum of 18 trials, with two comparisons presented at a time (i.e., A1C1C2, A1C1C3, A2C2C1, A2C2C3, A3C3C1, A3C3C2, B1C1C2, B1C1C3, B2C2C1, B2C2C3, B3C3C1, and B3C3C2). The fifth block was a mixing of all trial types, with all three comparisons presented in a minimum of 18 trials (i.e., A1C1C2C3, A2C1C2C3, A3C1C2C3, B1C1C2C3, B2C1C2C3, and B3C1C2C3). When all baseline relations were trained at a minimum of 90% correct in a training block consisting of 18 trials, feedback was reduced to 75% of trials for the next training block and to 50%, 25%, and 0% for the following blocks, respectively. In each block, at least 90% correct responses were required to proceed to the next block. The last block, with no feedback, was followed by a test block. The test block consisted of 18 symmetry trials and 18 equivalence trials randomly intermixed. Symmetry trials were C1A1A2A3, C2A1A2A3, C3A1A2A3, C1B1B2B3, C2B1B2B3, and C3B1B2B3. Equivalence trials were A1B1B2B3, A2B1B2B3, A3B1B2B3, B1A1A2A3, B2A1A2A3, and B3A1A2A3. No feedback was delivered during the tests for emergent relations.

Table 2
Sequence of Training and Test Phases Provided for Both Stimulus Sets

Blocks	Trials	Feedback	Minimum trials	Criterion
Training				
1. Separate	A1C1, A2C2, A3C3	100%	9	9/9
2. Separate	B1C1, B2C2, B3C3	100%	9	9/9
3. Mixed trials, 1 comparison	A1C1, A2C2, A3C3, B1C1, B2C2, B3C3	100%	18	17/18
4. Mixed trials, 2 comparisons	A1C1C2, A1C1C3, A2C2C1, A2C2C3, A3C3C1, A3C3C2, B1C1C2, B1C1C3, B2C2C1, B2C2C3, B3C3C1, and B3C3C2	100%	18	17/18
5. Mixed trials, 3 comparisons	A1C1C2C3, A2C1C2C3, A3C1C2C3, B1C1C2C3, B2C1C2C3, and B3C1C2C3	100%	18	17/18
6. Mixed trials, 3 comparisons	same as above	75%	18	17/18
7. Mixed trials, 3 comparisons	same as above	50%	18	17/18
8. Mixed trials, 3 comparisons	same as above	25%	18	17/18
9. Mixed trials, 3 comparisons	same as above	0%	18	17/18
Testing				
Test block with symmetry and equivalence trials randomly intermixed	C1A1A2A3, C2A1A2A3, C3A1A2A3, C1B1B2B3, C2B1B2B3, and C3B1B2B3 A1B1B2B3, A2B1B2B3, A3B1B2B3, B1A1A2A3, B2A1A2A3, and B3A1A2A3		36	17/18 17/18

Recordings. The software recorded all of the data, including trial number, which stimulus relation was trained or tested, number of responses to sample stimuli, reaction time to sample and comparison stimuli, correct/incorrect comparison choices, and whether feedback was provided or not. Thus, the duration of the experiment, number of baseline and test trials, and symmetry and equivalence indices were calculated by the software.

The accuracy criterion necessary to advance from one training phase to the next was a minimum of 94.4% correct comparison choices on three consecutive training blocks, constituting at least 17 of 18 correct choices. The equivalence criterion similarly was at least 17 of 18 correct, for symmetry or equivalence trials separately.

Statistical Analyses

For both the number of trials and the emergent relations, the data were analyzed separately by repeated-measures ANOVAs with one repeated factor (stimulus type) and one group factor (order). For statistical analyses of reaction time, data were organized in five-trial blocks—each block comprised one baseline measure for the last five training trials and one test measure for test trials. Each measure was computed as a mean of five trials. For reaction times the difference between symmetry and equivalence and possible interaction effects were of interest, and test condition was included in the analyses, resulting in a three-way ANOVA design with two repeated factors (test condition and stimulus type) and one between-subjects factor (order). Reaction times were measured in differences from baseline.

Results

Number of Trials

For the participants who started with the abstract-stimuli-only condition, the number of trials for the six baseline relations in both training phases was higher in the first condition (abstract stimuli only) than in the picture-as-nodes condition (as shown in the upper panel of Table 3). In the first condition, Participant 1355 had the highest number of baseline trials during the first phase of training (Blocks 1–5), with 360 trials, and Participant 1358 had the lowest number of trials, with 126 trials. During the phases in which the feedback was reduced (Blocks 6–9), none of the participants had more responses to criterion than the minimum. In the second condition, the familiar-stimuli-as-nodes condition, three of the eight participants (Participants 1380, 1381, and 1383) had more trials to criterion than in Phase 1. During the mixing phase, no extra trials were needed to meet the criterion.

As shown in the lower panel of Table 3, which depicts data for the participants starting with the familiar-stimuli-as-nodes condition, Participant 1359 had the highest number of baseline trials in the first phase of training, with 126 trials, while Participants 1352 and 1360 had 90 trials. For five of the eight participants, the number of trials was equal to criterion. In Phase 2 no participants had any extra trials to meet the criterion. In the subsequent condition with all abstract stimuli, Participant 1352 had the highest number of training trials, with 180 trials, and four other participants (Participants 1375, 1359, 1360, and 1379) had a higher number of trials than the criterion. For three participants (Participants 1371, 1376, and 1377), the number of training trials was equal to criterion. In Phase 2 no participants had any extra trials to meet the criterion.

The group data show that the number of trials to establish baseline relations was lower under the picture-as-nodes condition than under the abstract-stimuli-only condition regardless of order of stimulus sets in these participants. For the participants starting with the abstract-stimuli-only condition, the mean number of trials to criterion was 297 in the first condition and 171 in the second condition. For participants who started with the picture-as-nodes condition, the mean number of trials to criterion was 171 in the first condition and 202 in the second condition. The statistical analyses showed that there was a significant effect of stimulus type, that is, the number of trials in the first phase was significantly lower for the familiar picture stimuli than for abstract stimuli, $F[1, 14] = 27.95$, $p = 0.000$, and there was also an interaction between the effects of stimulus type and order, $F[1, 14] = 10.66$, $p = 0.006$.

Table 3
Individual Data for Both Experimental Conditions with Number of Total and Incorrect Trials in Baseline Training and Responding in Accord with Equivalence During Testing

Participant	Abstract stimuli						Familiar picture stimuli					
	Acquisition		Maintenance		Testing		Acquisition		Maintenance		Testing	
	Total	Incorrect	Total	Incorrect	Sym	Eq	Total	Incorrect	Total	Incorrect	Sym	Eq
1351	180	27	72	0	**17/18**	**17/18**	72	0	72	0	**18/18**	**18/18**
1353	162	27	72	1	9/18	11/18	72	1	72	0	**18/18**	**18/18**
1355	360	99	72	0	**18/18**	**18/18**	72	0	72	0	**18/18**	**18/18**
1356	144	17	72	0	16/18	10/18	72	0	72	0	**18/18**	**18/18**
1358	126	17	72	1	16/18	7/18	72	0	72	0	**17/18**	**17/18**
1380	198	53	72	0	13/18	6/18	90	8	72	4	**18/18**	**18/18**
1381	252	61	72	0	8/18	10/18	90	6	72	0	**18/18**	**18/18**
1383	234	42	72	5	**17/18**	14/18	108	12	72	0	**18/18**	**18/18**
	Familiar picture stimuli						Abstract stimuli					
	Acquisition		Maintenance		Testing		Acquisition		Maintenance		Testing	
	Total	Incorrect	Total	Incorrect	Sym	Eq	Total	Incorrect	Total	Incorrect	Sym	Eq
1352	90	2	72	4	**18/18**	16/18	180	19	72	4	**17/18**	7/18
1375	72	0	72	0	**18/18**	**18/18**	90	12	72	1	**18/18**	**17/18**
1376	72	2	72	2	**18/18**	**18/18**	72	1	72	1	**18/18**	**18/18**
1359	126	12	72	2	**18/18**	16/18	144	15	72	2	**18/18**	**18/18**
1360	90	2	72	2	16/18	**18/18**	162	21	72	1	13/18	12/18
1371	72	0	72	0	**18/18**	**18/18**	72	0	72	1	**18/18**	**18/18**
1377	72	0	72	0	**18/18**	**18/18**	72	0	72	1	**18/18**	**17/18**
1379	72	3	72	0	**18/18**	**18/18**	108	15	72	4	**18/18**	**17/18**

Note. Acquisition includes Blocks 1 through 5, in which there was a gradual increase in the number of comparisons. Maintenance includes Blocks 6 through 9, in which the feedback was gradually decreased. Sym = Symmetry trials. Eq = Equivalence trials. Results in bold indicate trials in which the accuracy criterion of 90% correct was reached.

Emergent Relations

As shown in the upper panel of Table 3, Participants 1351 and 1355 responded in accord with equivalence in both conditions. Participant 1383 responded in accord with symmetry in the abstract-stimuli-only condition and in accord with equivalence in the pictures-as-nodes condition. In the abstract-stimuli-only condition, Participants 1358 and 1380 had the lowest equivalence yields, with seven and six correct choices, respectively. Furthermore, Participants 1353 and 1381 both responded at a chance level but had higher yields on equivalence trials than on symmetry trials. When the familiar-picture-stimuli condition was introduced, all participants responded in accord with equivalence. In the familiar-picture-stimuli condition, seven participants responded correctly on all test trials, while Participant 1358 made one incorrect response for both symmetry and equivalence trials. For the participants starting with the familiar-picture-stimuli condition, seven participants made correct choices on all symmetry trials, but Participants 1352 and 1359 failed on test-for-equivalence trials (as shown in the lower panel of Table 3). Participant 1360 made two incorrect choices on test for symmetry but made correct choices on all equivalence trials. In the abstract-stimuli condition, six participants responded in accord with equivalence. On tests for symmetry, seven participants responded to the criteria. Participant 1352 made several incorrect choices on tests for equivalence relations. Participant 1360 failed on tests for both trial types, with 13 correct choices on symmetry trials and 12 correct choices on equivalence trials.

The results from all participants regardless of order of stimulus sets showed that, for the abstract-stimuli-only condition, eight of 16 of the participants responded in accord with equivalence given abstract stimuli only. Furthermore, 13 of 16 participants responded in accord with equivalence in the picture-as-nodes condition.

For the participants who started with the abstract-stimuli-only condition, two participants responded in accord with equivalence in the first condition and all participants responded in accord with equivalence in the condition with familiar picture stimuli as nodes. For the participants who started with the familiar-stimuli-as-nodes condition, six participants responded in accord with equivalence when familiar picture stimuli were nodes, and six participants responded in accord with equivalence when all stimuli were abstract stimuli. The statistical analyses showed that there was no effect of order. However, there was an effect of stimulus type and an interaction between stimulus type (familiar stimuli vs. abstract stimuli) and order. For symmetry the main effect of stimulus type was significant, $F[1, 14] = 11.67$, $p = .004$, as was the effect of interaction between stimulus type and order, $F[1, 14] = 11.67$, $p = .004$. For equivalence the main effect of stimulus type was significant, $F[1, 14] = 9.00$, $p = .01$, as was the effect of interaction between stimulus type and order, $F[1, 14] = 9.00$, $p = .01$.

Reaction Times

For the participants starting with the abstract-stimuli-only condition, the median reaction time to comparison stimuli increased from the last five training trials to the first five test trials for both symmetry and equivalence trials in both conditions (see Figure 3). The typical pattern for reaction time to comparison stimuli was not seen for symmetry trials (i.e., the decrease in reaction time to comparison stimuli from the first five test trials to the last five test trials), while the median reaction time for equivalence trials did show this typical pattern. For participants first starting with the picture-as-nodes condition, there was a small increase in median reaction time from the last five training trials to the first five symmetry trials in the first condition, while reaction time remained the same in the second condition with abstract stimuli only (see Figure 4). In sum, for the equivalence trials, median reaction time increased for both picture stimuli and abstract stimuli. In these participants, the typical pattern of decreased reaction time to comparison stimuli was seen only for equivalence trials in the second condition with abstract stimuli only. The statistical analyses showed that there was an effect of stimulus type, $F(1, 14) = 8,55$, $p = .001$, which means that reaction time was longer for the structures with abstract stimuli only than those with familiar picture stimuli. Furthermore, there was an effect of test conditions (trial type), $F(1, 14) = 58,45$, $p = .001$, indicating that reaction time was longer for equivalence trials than symmetry trials.

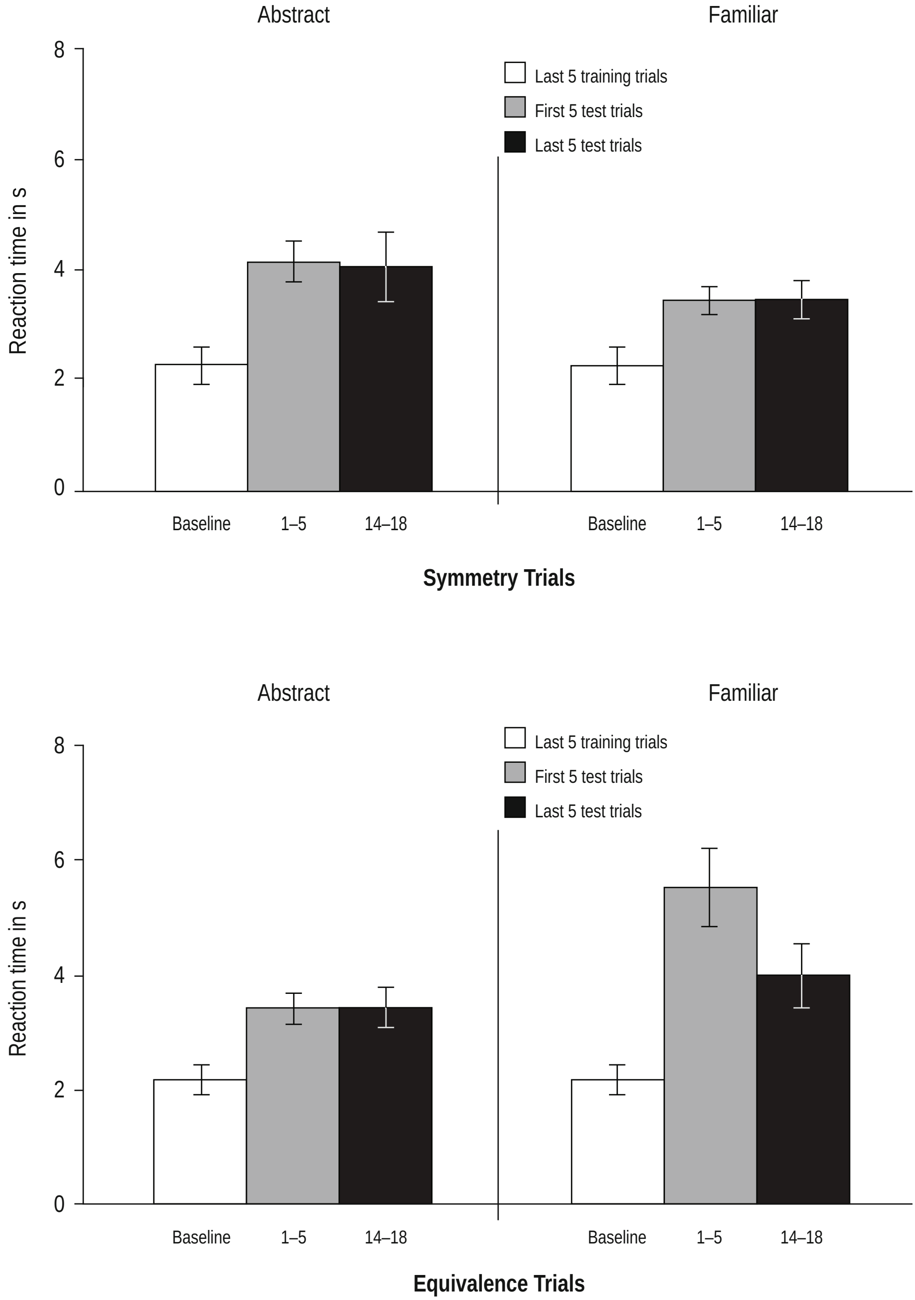

Figure 3. Reaction times for the last five training trials, the first five test trials, and the last five test trials for the participants starting with the abstract condition. The upper panel shows data for symmetry responding, while the lower panel shows data for equivalence responding. Standard error of the mean is shown for each bar.

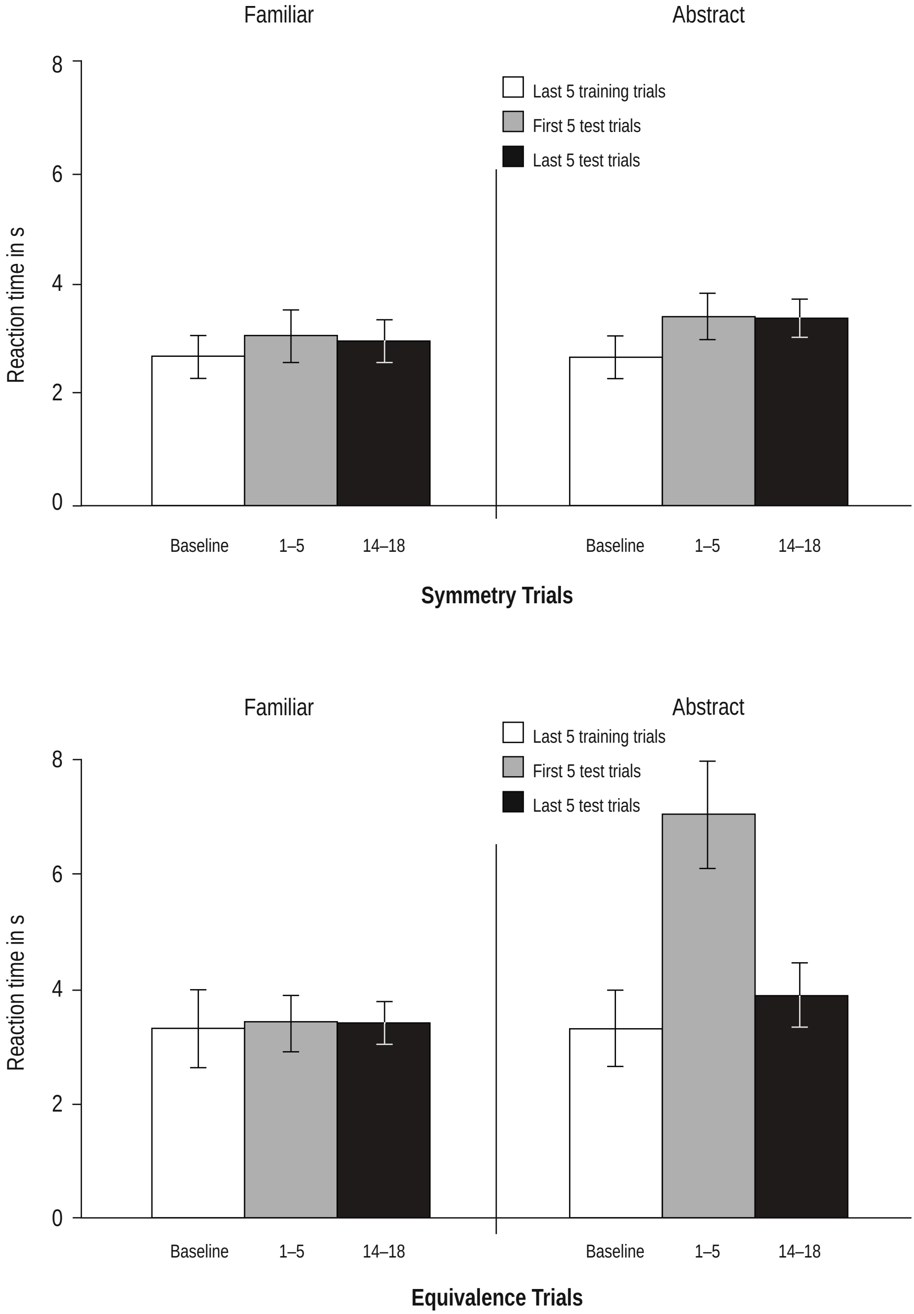

Figure 4. Reaction times for the last five training trials, the first five test trials, and the last five test trials for the participants starting with the picture condition. The upper panel shows data for symmetry responding, while the lower panel shows data for equivalence responding. Standard error of the mean is shown for each bar.

Discussion

The purpose of the present experiment was to investigate the effects of responding in accord with equivalence as a function of pictures versus abstract stimuli as nodes. The results showed that pictures were more effective than abstract stimuli both in training of the conditional discriminations (i.e., the number of responses to criterion) and in tests for responding in accord with equivalence. Furthermore, none of the 16 typically developing children reached the test criteria in the minimum number of trials before testing in the condition with all abstract stimuli, while five participants reached the criteria for testing in a minimum number of trials in the familiar-stimuli-as-nodes condition. Hence, the condition with pictures as nodes established baseline relations in fewer trials and with fewer errors than the condition with all abstract stimuli. In addition, when the participants were trained first with familiar picture stimuli, more of the participants responded in accord with equivalence in the condition with all abstract stimuli than when abstract stimuli were presented as the first condition. Finally, the typical increase in reaction time from training to test was most pronounced for equivalence trials and in the condition with all abstract stimuli.

In the current study we found that for conditional discrimination training with familiar stimuli, the number of trials to criterion was lower than with abstract stimuli only. Furthermore, when the participants had an experimental history with familiar stimuli, they established the conditional discriminations faster. The experimental design also controlled for the possibility that this effect could be related to the effect of order per se. As mentioned earlier, it is important to find procedures that are the most effective in establishing conditional discriminations. The results from the current experiment indicate that the use of pictures as one of the stimulus sets could be very effective in training conditional discriminations in children. However, Lyddy, Barnes-Holmes, and Hampson (2000) conducted an experiment with adult participants on the effect of different degrees of meaningfulness of CVC syllables, defined as the degree to which the nonsense syllables evoked real word or concept associations, on performance in MTS tasks and found that low-meaningfulness stimuli produced fewer errors than did high-meaningfulness stimuli. Hence, they suggested that high-meaningfulness stimuli might produce conflict with trained relations. Because the findings on pictures and meaningful stimuli seem to diverge, we think it will be important to conduct future research on this issue. For example, parametric studies on different aspects of stimuli could be useful. One such study could evaluate results when training conditional discriminations and stimulus equivalence outcomes as a function of stimuli along a dimension from abstract to familiar (recognizable and nameable picture stimuli) as a possible way to study the meaningfulness of stimuli.

The results of derived relations in the current study showed that there was significantly more responding in accord with equivalence with pictures as nodes than with abstract stimuli. Furthermore, more of the participants starting with the picture condition responded in accord with equivalence in the abstract condition than the participants who had the abstract condition first. The facilitating effect of familiar picture stimuli is consistent with other studies (Arntzen, 2004; Holth & Arntzen, 1998a). The pictures used in the current study are nameable and, therefore, could be responsible for the greater number of participants in the familiar picture condition responding in accord with equivalence as compared to the abstract-stimuli-only condition. Another related issue is the pronounceability of stimuli (Mandell, 1997; Mandell & Sheen, 1994). For example, Mandell and Sheen (1994) found that participants who experienced easy-to-pronounce stimuli (five-letter phonological pseudowords) produced stimulus equivalence more quickly and with fewer errors than they did with difficult-to-pronounce stimuli. Furthermore, Randell and Remington (1999) found that classes of visual stimuli in which the names of the stimuli rhymed produced baseline relations and responding in accord with equivalence more consistently than the control conditions. The control conditions constituted visual stimuli in which the names that participants gave to the comparison stimuli in each baseline trial rhymed, but the sample and the comparison names did not. The result of this study indicated that the phonological properties of naming responses influenced the equivalence class formation.

Another measure that has been used in research on the meaningfulness of stimuli is reaction time to stimuli presented. The main findings have been that the reaction time is shorter when stimuli are semantically related than when they are not (e.g., Meyer & Schvaneveldt, 1971). One interpretation of the findings from the current study could be that the differences in reaction time between the two conditions, with and without pictures, are related to a certain degree to the meaningfulness of stimuli. For example, when the participants started with the abstract-stimuli-only condition, there was a much higher increase in reaction time from baseline trials to the first test trials on both symmetry and equivalence trials compared to reaction times for the participants starting with the familiar-stimuli-as-nodes condition. In addition, the increase in reaction time from baseline trials to test trials was always most pronounced for equivalence trials. This is probably related to the fact that for both conditions in the test for responding in accord with equivalence, none of the stimuli were seen together before, and in the test for responding in accord with symmetry, the comparisons from training were presented as samples and samples from training were presented as comparisons. The differences in reaction times on symmetry trials for the two conditions, with and without pictures, could be directly related to the fact that pictures are presented on test trials in one of the conditions and not in the other one. Thus, for the condition with pictures as nodes, the differences in reaction times for equivalence versus symmetry could be related to the fact that in the first type of test no meaningful stimuli or pictures were present, but they were present in the second test. Hence, the differences in reaction times between symmetry trials and equivalence trials in the current study are in accord with other findings (e.g., Bentall, Dickins, & Fox, 1993; Bentall, Jones, & Dickins, 1999; S. C. Hayes & Bissett, 1998; Spencer & Chase, 1996).

It seems reasonable to assume that the use of familiar picture stimuli could facilitate some mediating behavior, which could be important during the establishment of conditional relations and also during testing for equivalence classes. In the current experiment, all the data related to the dependent variables, such as number of trials to criterion, number of participants responding in accord with equivalence, and reaction time support the effectiveness of pictures or meaningful stimuli. However, for both conditions, with and without pictures, there was an increase in reaction time from training to test. Different accounts have been set forth to explain the initial increase in reaction time during testing, including some sort of naming (Horne & Lowe, 1996) or "precurrent" responses to a problem situation (Holth & Arntzen, 2000). We think that there is a need for more molecular analyses of stimulus equivalence and, therefore, future research should, for example, focus on time restrictions for responding during testing and also variations of titrating limited hold (Tomanari, Sidman, Rubio, & Dube, 2006).

As mentioned earlier, the findings from Smeets and Barnes-Holmes (2005) were not in favor of pictures or meaningful stimuli as nodes, rather the opposite. Several important issues regarding the Smeets and Barnes-Holmes study could account for the discrepancy. First, Smeets and Barnes-Holmes provided the following instruction during the first two trials in the second phase of the experiment: "This (pointing to abstract stimulus A1) is an apple, this (pointing to abstract stimulus B1) is a nose, and this (pointing to abstract stimulus B2) is a flag" (p. 286). Participants in the condition with the OTM training structure and familiar picture stimuli were then asked to point to the flag. Corresponding instructions were provided for the MTO conditions. However, such a procedure could actually teach the participants to name the stimuli. Hence, several studies have shown that naming stimuli during equivalence tasks have facilitated responding in accord with equivalence (e.g., Eikeseth & Smith, 1992), even though it is not evident that naming is a necessary condition for the emergence of responding in accord with equivalence (e.g., Randell & Remington, 1999; Sidman, 1994). In the current and earlier studies in our lab, we have been very careful to avoid giving any instructions that might influence the explicitly arranged experimental conditions (see, e.g., Arntzen, Vaidya, & Halstadtro, 2008, on the role of instruction in stimulus equivalence research). As mentioned earlier, one possible interpretation of the results in the present study is that familiar picture stimuli facilitate precurrent verbal behavior. It could be argued that teaching naming as part of the experimental procedure might have cancelled out the effects of using familiar picture stimuli in the Smeets and Barnes-Holmes study. Another possible explanation for the marginal differentiation between abstract and familiar picture stimuli in the Smeets and Barnes-Holmes study is that their familiar picture stimuli did not differ significantly from the abstract stimuli. For example, the A1 stimulus was quite similar to the D2 stimulus. On the other hand, the familiar picture stimuli used in our lab were quite different from the abstract stimuli.

Second, the current study used a three-choice MTS format, whereas Smeets and Barnes-Holmes (2005) provided a two-choice MTS format. Carrigan and Sidman (1992) and Sidman (1987, 2000) have warned against the use of the two-choice format in equivalence research because both rejecting the "wrong" comparison stimulus and selecting the "right" comparison stimulus lead to correct comparison choice and, therefore, some variables other than established baseline relations could be responsible for emergent responding. Other researchers (Boelens, 2002; Wilkinson & McIlvane, 2001) defend the use of the two-choice MTS format and argue that Sidman's objection is a purely technical difficulty that can be handled within the experimental procedure itself, for example, by requiring high percentage-correct responding in training conditions to ensure baseline relations are established, and then by replicating the experiment with different stimuli or counterbalancing stimuli pairings during test trials in a group design to increase the likelihood that the baseline training is causally related to test outcomes (Boelens, 2002). Although the two-choice format can be justified in addressing certain research issues as emphasized by Boelens, we would argue that, in the search for variables that influence the acquisition of baseline trials and, consequently, responding in accord with equivalence, Sidman's critique is highly relevant.

Finally, the current study employed computer-administered stimuli, whereas Smeets and Barnes-Holmes (2005) used manual MTS arrangements. Even though the experimenter is extensively trained, it could be argued that tabletop administration of stimuli entails weaker experimental control of variables than computer-administered stimuli, which could influence participants' performance. Future experiments should attempt to replicate the Smeets and Barnes-Holmes (2005) study and take into consideration the procedural issues mentioned here.

In summary, the results of the current study showed that the condition with familiar picture stimuli as nodes produced more responding in accord with equivalence and a lower number of trials to establish baseline relations. Though on picture training trials everyone showed few errors, on abstract training the picture-first participants showed far fewer errors than the abstract-first participants. The results support earlier findings. The experimental history with familiar picture stimuli had an effect on the training of classes with abstract stimuli, that is, there was a significant effect of having the condition with familiar picture stimuli as nodes as the first condition. The reaction time results showed that the increase from baseline to testing was most pronounced for equivalence trials and for the condition with abstract stimuli. Equivalence reaction times were higher than symmetry at first but reduced to the same level on further testing, though both remained higher than on training trials.

References

ANNETT, J. M., & LESLIE, J. C. (1995). Stimulus classes involving olfactory stimuli. *The Psychological Record, 45*, 439–450.

ARNTZEN, E. (2004). Probability of equivalence formation: Familiar stimuli and training sequence. *The Psychological Record, 54*, 275–291.

ARNTZEN, E. (2006). Delayed matching to sample and stimulus equivalence: Probability of responding in accord with equivalence as a function of different delays. *The Psychological Record, 56*, 135–167.

ARNTZEN, E., GALAEN, T., & HALVORSEN, L. R. (2007). Different retention intervals in delayed matching-to-sample: Effects of responding in accord with equivalence. *European Journal of Behavior Analysis, 8*, 177–191.

ARNTZEN, E., GRONDAHL, T., & EILIFSEN, C. (2010). The effects of different training structures in the establishment of conditional discriminations and the subsequent performance on the tests for stimulus equivalence. *The Psychological Record, 60*, 437–462.

ARNTZEN, E., HALSTADTRO, L. B., BJERKE, E., & HALSTADTRO, M. (2010). Training and testing theoretical music skills in a boy with autism using a matching-to-sample format. *Behavioral Interventions, 25*, 129–143. doi: 10.1002/bin.301

ARNTZEN, E., & VAIDYA, M. (2008). The effect of baseline training structure on equivalence class formation in children. *Experimental Analysis of Human Behavior Bulletin, 29*, 1–8.

ARNTZEN, E., VAIDYA, M., & HALSTADTRO, L. B. (2008). On the role of instruction in conditional discrimination training. *Experimental Analysis of Human Behavior Bulletin, 29*, 17–24.

BELANICH, J., & FIELDS, L. (1999). Tactual equivalence class formation and tactual-to-visual cross-modal transfer. *The Psychological Record, 49*, 75–91.

BENTALL, R. P., DICKINS, D. W., & FOX, S. R. A. (1993). Naming and equivalence: Response latencies for emergent relations. *The Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 46B*, 187–214.

BENTALL, R. P., JONES, R. M., & DICKINS, D. W. (1999). Errors and response latencies as a function of nodal distance in 5-member equivalence classes. *The Psychological Record, 49*, 93–115.

BOELENS, H. (2002). Studying stimulus equivalence: Defense of the two-choice procedure. *The Psychological Record, 52*, 305–314.

CARRIGAN, P. F., & SIDMAN, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 183–204.

CAUTILLI, J. D., HANCOCK, M. A., THOMAS, C. A., & TILLMAN, C. (2002). Behavior therapy and autism: Issues in diagnosis and treatment. *The Behavior Analyst Today, 3*, 112–125.

DEVANY, J. M., HAYES, S. C., & NELSON, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 46*, 243–257.

DUGDALE, N., & LOWE, C. F. (2000). Testing for symmetry in the conditional discriminations of language-trained chimpanzees. *Journal of the Experimental Analysis of Behavior, 73*, 5–22.

DYMOND, S., & REHFELDT, R. A. (2001). Supplemental measures and derived stimulus relations. *Experimental Analysis of Human Behavior Bulletin, 19*, 8–12.

EIKESETH, S., & SMITH, T. (1992). The development of functional and equivalence classes in high-functioning autistic children: The role of naming. *Journal of the Experimental Analysis of Behavior, 58*, 123–133.

FIELDS, L., REEVE, K. F., ROSEN, D., VARELAS, A., ADAMS, B. J., BELANICH, J., ET AL. (1997). Using the simultaneous protocol to study equivalence class formation: The facilitating effects of nodal number and size of previously established equivalence classes. *Journal of the Experimental Analysis of Behavior, 67*, 367–389.

FIELDS, L., & VERHAVE, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior, 48*, 317–332.

GREEN, G., & SAUNDERS, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229–262). New York: Plenum Press.

HAYES, L. J., TILLEY, K. L., & HAYES, S. C. (1988). Extending equivalence and membership to gustatory stimuli. *The Psychological Record, 38*, 473–482.

HAYES, S. C., & BISSETT, R. T. (1998). Derived stimulus relations produce mediated and episodic priming. *The Psychological Record, 48*, 617–630.

HOLTH, P., & ARNTZEN, E. (1998a). Stimulus familiarity and the delayed emergence of stimulus equivalence or consistent nonequivalence. *The Psychological Record, 48*, 81–110.

HOLTH, P., & ARNTZEN, E. (1998b). Symmetry versus sequentiality related to prior training, sequential dependency of stimuli, and verbal labeling. *The Psychological Record, 48*, 293–315.

HOLTH, P., & ARNTZEN, E. (2000). Reaction times and the emergence of class consistent responding: A case for precurrent responding? *The Psychological Record, 50*, 305–338.

HORNE, P. J., & LOWE, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*, 181–241.

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, 18*, 279–289.

LESLIE, J. C., TIERNEY, K. J., ROBINSON, C. P., KEENAN, M., & WATT, A. (1993). Differences between clinically anxious and non-anxious subjects in a stimulus equivalence training task involving threat works. *The Psychological Record, 43*, 153–161.

LIPKINS, R., HAYES, S. C., & HAYES, L. J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology, 56*, 201–239.

LYDDY, F., BARNES-HOLMES, D., & HAMPSON, P. J. (2000). The effect of stimulus meaningfulness on the formation of equvialence classes. *European Journal of Behavior Analysis, 1*, 71–87.

MANDELL, C. (1997). Stimulus equivalence and meaning: The influence of verbal behavior. In C. Mandell & A. McCabe (Eds.), *The problem of meaning: Behavioral and cognitive perspectives* (pp. 81–116). New York: Elsevier Science.

MANDELL, C., & SHEEN, V. (1994). Equivalence class formation as a function of the pronounceability of the sample stimulus. *Behavioural Processes, 32*, 29–46.

MEYER, D. E., & SCHVANEVELDT, R. W. (1971). Facilitation in recognition pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227–234.

O'LEARY, C. A., & BUSH, K. M. (1996). Stimulus equivalence in the tactile modality. *The Psychological Record, 46*, 509–517.

PILGRIM, C., CHAMBERS, L., & GALIZIO, M. (1995). Reversal of baseline relations and stimulus equivalence: II. Children. *Journal of the Experimental Analysis of Behavior, 63*, 225–238.

PLAUD, J. J. (1995). The formation of stimulus equivalences: Fear-relevant versus fear-irrelevant stimulus classes. *The Psychological Record, 45*, 207–222.

RANDELL, T., & REMINGTON, B. (1999). Equivalence relations between visual stimuli: The functional role of naming. *Journal of the Experimental Analysis of Behavior, 71*, 395–415.

SAUNDERS, R. R., WACHTER, J. A., & SPRADLIN, J. E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedure. *Journal of the Experimental Analysis of Behavior, 49*, 95–115.

SIDMAN, M. (1987). Two choices are not enough. *Behavior Analysis, 22*, 11–18.

SIDMAN, M. (1992). Equivalence relations: Some basic considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 15–27). Reno, NV: Context Press.

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

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

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.

SMEETS, P. M., & BARNES-HOLMES, D. (2005). Establishing equivalence classes in preschool children with one-to-many and many-to-one training protocols. *Behavioural Processes, 69*, 281–293.

SPENCER, T. J., & CHASE, P. N. (1996). Speed analysis of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65*, 643–659.

SPRADLIN, J. E., & SAUNDERS, R. R. (1986). The development of stimulus classes using match-to-sample procedures: Sample classification versus comparison classification. *Analysis and Intervention in Developmental Disabilities, 6*, 41–58.

STROMER, R., MACKAY, H. A., & STODDARD, L. T. (1992). Classroom applications of stimulus equivalence technology. *Journal of Behavioral Education, 2*, 225–256.

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*, 349–369.

WILKINSON, K. M., & MCILVANE, W. J. (2001). Methods for studying symbolic behavior and category formation: Contributions of stimulus equivalence research. *Developmental Review, 21*, 355–374.