Establishing Equivalence Classes in Children Using Familiar and Abstract Stimuli and Many-to-One and One-to-Many Training Structures

Erik Arntzen and Silje Lunde Nikolaisen
Akershus University College

In the present experiment, the effects of responding in accordance with stimulus equivalence were tested using one-to-many (OTM) and many-to-one (MTO) training structures and familiar and abstract stimuli. Sixteen typically developing children, ages 8 to 9 years, were trained and tested for formation of two 5-member equivalence classes in four conditions: OTM–familiar, OTM–abstract, MTO–familiar, and MTO–abstract. In the the familiar conditions, the nodes were familiar stimuli while the other stimuli were abstract stimuli. In the abstract condition, all stimuli were abstract stimuli. The results showed that the conditions with the familiar stimulus sets were more effective in formation of equivalence classes than the conditions with abstract stimuli. Furthermore, the results showed small discrepancies between the training structures in the formation of equivalence classes, and OTM was more effective than MTO in the first condition. In the conditional discrimination training, OTM was faster in establishing the conditional relations than MTO. The results suggest the use of familiar stimuli to establish equivalence relations and the use of OTM for effective establishment of conditional relations.

Key words: stimulus equivalence, emergent relations, meaningfulness, familiar stimuli, abstract stimuli, training structures

---

Conditional discrimination and matching-to-sample procedures have been used to train relations between stimuli and subsequently assess the emergence of derived relations, that is, relations that have not been directly trained (Sidman, 1994; Sidman & Tailby, 1982). For example, when training of AB relations, the trial for the AB relation in Class 1 would include that A1 is presented as the sample while B1, B2, and B3 are presented as comparisons, where B1 is the correct choice (i.e., one trial type is A1<u>B1</u>B2B3). Likewise, when training the AB relations in Classes 2 and 3. When training BC relations, the trial for the BC relation in Class 1 would include that B1 is presented as the sample while C1, C2, and C3 are presented as comparisons, where C1 is correct choice (i.e., one trial type is B1<u>C1</u>C2C3). Likewise, when training the BC relations in Classes 2 and 3. When the correct stimulus is chosen to match another stimulus in the same class, the response is reinforced, and the relations is called a directly trained relation or a baseline relation. When the conditional relations have been established, we can test for the emergence of untrained or derived relations.

One interesting aspect of stimulus equivalence is that, after a few conditional relations are trained, several new relations emerge without any direct training. Stimulus equivalence concerns the notion that classes of stimuli have the properties of reflexivity, symmetry, and transitivity.

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

When training AB and BC conditional relations, reflexivity means that the participant must show that each stimulus in the classes bears a relation to itself. In other words it involves interchangeability of the stimuli. In this preparation, then, symmetry would be demonstrated when the participant responds in a consistent manner when presented with BA and CB relations. Transitivity means that the participant responds according to the experimenter-defined AC relations. The overall equivalence test assesses symmetry and transitivity which means that the participant responds according to the experimenter-defined CA relations. When assessing the emergence of an equivalence class, the participant must respond in a class-consistent manner to all of the derived relations mentioned above. Furthermore, the participant must demonstrate the properties that define equivalence and, thus, show that the stimuli in the set are acting as members of an equivalence class (Sidman, 1992). Reflexivity is usually not tested for, because most humans can relate identical stimuli with each other. Such testing may cause participants to look for patterns of physical similarity, which in turn may interfere with tests for symmetry, transitivity, and equivalence relations. The stimuli used in stimulus equivalence studies become mutually interchangeable as a result of training a certain number of conditional discriminations. The stimuli used in studies on stimulus equivalence do not bear any physical similarities. When the stimuli are abstract, there is very little chance that the participant has had a learning history with the stimuli (Sidman, 1994; Sidman & Tailby, 1982).

Imam (2006) described three different training and testing protocols, that is simple-to-complex, complex-to-simple, and simultaneous protocol. In the present study we have used a simultaneous protocol. Furthermore, three training structures, one-to-many (OTM), many-to-one (MTO), and linear series (LS), have been used in conditional discrimination procedures (e.g., K. J. Saunders, Saunders, Williams, & Spradlin, 1993). With three members in a class, for the LS structure the A stimuli are related conditionally to the B stimuli, and then the B stimuli are related to the C stimuli; the test for emergent relations looks at BA, CB, AC, and CA relations. With OTM, the training involves AB and AC relations, and testing is of the relations BA, CA, BC, and CB. With MTO, the A and B stimuli are conditionally related to the C stimuli, and the test for emergent relations evaluates CA, CB, AB, and BA relations (Arntzen & Vaidya, 2008). When the simultaneous protocol has been used some studies have found the MTO training structure to be the most effective in producing formation of stimulus equivalence (e.g., Fields, Hobbie-Reeve, Adams, & Reeve, 1999; R. R. Saunders, Drake, & Spradlin, 1999; R. R. Saunders, Wachter, & Spradlin, 1988; Spradlin & Saunders, 1986). Other studies have found the OTM training structure to be more effective (e.g., Arntzen & Holth, 1997, 2000a). The LS training structure has been shown to be the least effective in producing equivalence formation in most studies (e.g., Arntzen, Grondahl, & Eilifsen, 2010; Arntzen & Holth, 1997, 2000a). Others have found no discrepancies in responding to equivalence relations when using either MTO or OTM training (e.g., Smeets & Barnes-Holmes, 2005). Most of the studies that have found MTO to be superior to OTM have used a two-choice format and children or persons with developmental disabilities as participants. However, the studies that have shown the superiority of OTM have used a three-choice format and adult participants.

Saunders and Green (1999) published a discrimination analysis to which they provided a coherent account of the differences in yield produced by the three training structures. In their logic analysis, they predicted that MTO would produce higher yields of equivalence relations than would OTM and LS. Their argument for the superiority of MTO is that, in the establishment of conditional discrimination, each conditional discrimination requires successive discrimination between samples and choice discrimination between comparisons. The establishment of these simple discriminations must occur in the MTO procedure while they might be formed under OTM and LS even if they are not required. Because of the discrepancies in the findings in the formation of equivalence classes, we wanted in the present study to compare MTO and OTM with a larger number of members in each class.

Both meaningful and abstract stimuli have been used in experiments to determine which are the most effective in formation of equivalence relations. Meaningfulness has been defined as stimuli that generate real words or concepts (Lyddy, Barnes-Holmes, & Hampson, 2000). The use of meaningful stimuli can involve, for example, pictures that are already known to the participant. Research using different sets of stimuli that are either abstract or familiar has shown that familiar stimuli result in higher equivalence yields (e.g., Arntzen, 2004; Arntzen & Holth, 2000b; Arntzen & Lian, 2010). Arntzen (2004) found that introducing familiar stimuli early in training resulted in the formation of equivalence classes by a higher proportion of participants in a group than when the meaningful stimuli were introduced late in the training. The use of the all-abstract-stimuli condition resulted in the lowest yields of equivalence class formation. In another study, Arntzen and Lian (2010) studied equivalence class formation in children. The participants were trained with MTO to form three 3-member classes in two conditions, one group with all abstract stimuli and one group with familiar stimuli as nodes or comparisons. Half of the participants started with the all-abstract-stimuli condition followed by the condition with familiar stimuli. The other participants were exposed to the conditions in the reverse order. The data showed that responding in accordance with stimulus equivalence was more probable with familiar stimuli as nodes than with abstract stimuli only.

Saunders et al. (1999) studied responding in accordance with stimulus equivalence in children as a function of MTO and OTM training. After training of two 5-member classes, more children showed higher yields of stimulus equivalence following MTO than OTM. Smeets and Barnes-Holmes (2005) extended the study by including some familiar stimuli. They studied derived relations as a function of MTO and OTM training with different stimulus material, abstract stimuli only and familiar stimuli, in 16 preschool children. They set up four different conditions: OTM–familiar, OTM–abstract, MTO–familiar, and MTO–abstract. The pictures in the familiar-stimuli condition were comparisons in the MTO training and samples in the OTM training. Each child was tested in two different conditions, for example, MTO–abstract first and OTM–familiar second. Two 5-member classes were used. Eight children received the MTO structure and the other half the OTM; half of the children in each group received familiar stimuli and the other half abstract stimuli. Eight of the children participated in a second experiment with opposite conditions from the first experiment. The data showed no difference in equivalence yields as a function of training structures. The authors, however, found that using abstract stimuli only was more effective in establishing equivalence classes than using familiar stimuli. The authors concluded that the children's learning history was the reason for differences in responding to equivalence class formation. The results from Smeets and Barnes-Holmes are in contrast to most of the studies that have used pictures as one of the stimulus sets or as one of the stimuli in a set of otherwise abstract stimuli. Furthermore, one should predict that the use of pictures to form equivalence classes should increase with children as participants regardless of training structure. Therefore, it seems reasonable to arrange experiments with children as participants in which the effects of pictures and abstract stimuli on equivalence formation are studied further.

The purpose of the present study was to replicate the study by Smeets and Barnes-Holmes (2005). The first goal was to determine if the use of MTO as a training structure would give a higher probability of responding in accord with stimulus equivalence than the use of OTM. The second goal was to determine whether the probability of responding in accordance with stimulus equivalence was more likely when the training structure involved familiar stimuli as nodes as opposed to abstract stimuli as nodes. One predicted outcome of the present study was that responding in accordance with stimulus equivalence would be more likely with familiar stimuli than with abstract stimuli (e.g., Arntzen, 2004; Arntzen & Lian, 2010). This was predicted given that the participants would likely have a learning history with the familiar stimuli that they do not have with the abstract ones, meaning that they could name the familiar stimuli. Furthermore, that the difference in reaction time from training to test should be more pronounced for the all-abstract-stimuli condition compared to the familiar-stimuli condition. Another prediction was that MTO would be more effective in producing responding to equivalence relations than OTM. This is related to the discrimination analysis by Saunders and Green (1999) in which all of the simple discriminations required for the test are trained in the MTO. On the other hand, OTM does not include all of the simple discriminations in the training and also requires more successive discriminations among comparisons, which should make it less likely to establish the emergent relations. Finally, MTO should take more time to train.

Method

Participants

Sixteen children—8 boys and 8 girls from the age of 8 years, 5 months to 9 years, 10 months—participated in the study (see Table 1). The children were assigned to four groups and then assigned to different conditions. The participants in the first three groups were located at one school, and the fourth group was located at another school. The first 12 participants were recruited through teachers at their school; they received information with a consent form that their parents filled out. The parents of the participants in the last group were contacted directly by the experimenter with the same information and consent form. The consent form contained information about the project and described the task. The approximate length of the experiment was 3 hrs, depending on the children's performance. The parents and children were informed that the experiment would be carried out on two successive days, and the children were informed that they could withdraw from the experiment at any time. None of the children had participated in any experiment previously, and they had no experience with the type of training and testing that they received. When the children had completed the tasks, they received a medal with their name on it as a reward, as well as stickers they had chosen previous to the experimental session.

Setting and Apparatus

The experimental sessions were carried out at the children's schools in different classrooms or a quiet room outside the schools. Some children were tested alone, but most of them were tested in pairs. When two children were tested in the same condition at the same time, they sat about 5 m apart with their backs to each other. The training and testing were controlled by a software program written by Psych Fusion in collaboration with the first author. Two different computers were used; Computer 1 was an HP Compaq nc6320 with a 15-in. screen, and Computer 2 was an HP Compaq nw 8440 with a 15.4-in. screen. The participants who used Computer 1 were Participants 7801, 7803, 7806, 7808, 7811, 7818, 7819, 7823, and 7827. The other participants used Computer 2.

Stimuli

The stimuli used in the experiment were the same as those used in the Smeets and Barnes-Holmes (2005) study (see Figure 1).

In two conditions both familiar and abstract stimuli were used, and in the other two conditions all stimuli were abstract stimuli. Eight abstract stimuli were divided into two different classes and were used in all conditions. The background on the computer was white, the abstract stimuli were black, and the familiar stimuli were colored, with a red heart and a yellow smiley face.

Design

The design employed four different conditions, and each participant was assigned to two conditions. Two training structures (OTM and MTO) and two sets of stimuli (one with all abstract stimuli and one with familiar stimuli as nodes) were used. This resulted in four conditions: OTM–familiar, OTM–abstract, MTO–familiar, and MTO–abstract (see Table 1). In the OTM training structure, A stimuli served as samples or nodes and B, C, D, and E stimuli served as comparisons. In the MTO training structure, A, B, C, or D stimuli served as samples and E stimuli served as comparisons or nodes.

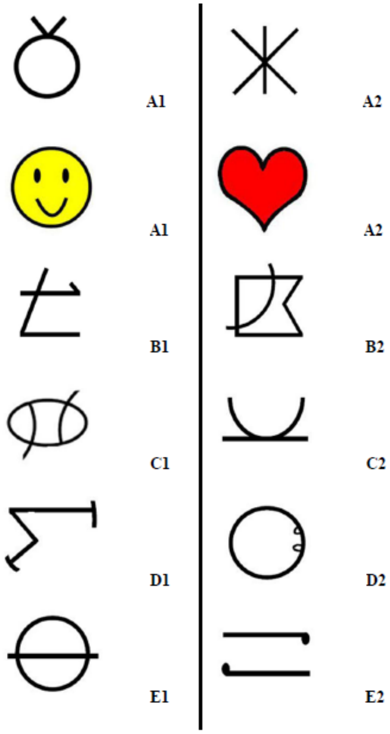

Figure 1. The stimuli used in the experiment. Note. A1 and A2 are either an abstract or familiar stimulus.

Procedure

If a participant received, for example, OTM–familiar in the first condition, he or she would receive an opposite condition with the two other variables the second day, that is, MTO–abstract. The first group received the OTM–familiar condition the first day and the MTO–abstract condition the second day (Participants 7801, 7802, 7803, and 7804). The second group received the OTM–abstract condition followed by the MTO–familiar condition (Participants 7806, 7808, 7818, and 7819), the third group received the MTO–familiar condition followed by the OTM–abstract condition (Participants 7811, 7812, 7814, and 7815), and the last group received the MTO–abstract condition followed by the OTM–familiar condition (Participants 7820, 7823, 7825, and 7827). All participants received the two conditions two days in a row.

Instructions. Before the training began, the participants were given instructions inNorwegian on the computer, and the experimenter read the instructions to the participants. The instructions were as follow:

> When the trial begins, a picture will appear on the top of the screen. Click on this picture. Then, two pictures will appear in the corners at the bottom of the screen. Choose one of these pictures. If you choose the correct one, the words “right,” “correct,” “great,” and so on will appear on the screen. If you choose the wrong picture, the word “wrong” will appear on the screen. With this information you will find out which picture is the correct choice. After a while, you will not be given information about whether you have chosen the right or the wrong picture. Do the best you can to get everything right. Thank you for participating and good luck.

In addition, the participants were given information about the numbers displayed

Table 1. *Demographic Characteristics of the Participants and Participants' Experimental Conditions*

Participant number	Gender	Age (Years, Months)	Conditions	
7801	F	8, 10	OTM–FAM	MTO–ABS
7802	F	8, 5	OTM–FAM	MTO–ABS
7803	M	8, 10	OTM–FAM	MTO–ABS
7804	F	8, 7	OTM–FAM	MTO–ABS
7806	M	8, 7	OTM–ABS	MTO–FAM
7808	M	8, 11	OTM–ABS	MTO–FAM
7818	M	8, 8	OTM–ABS	MTO–FAM
7819	F	8, 6	OTM–ABS	MTO–FAM
7811	F	8, 9	MTO–FAM	OTM–ABS
7812	M	9, 3	MTO–FAM	OTM–ABS
7814	F	9, 4	MTO–FAM	OTM–ABS
7815	F	9, 3	MTO–FAM	OTM–ABS
7820	F	9, 2	MTO–ABS	OTM–FAM
7823	M	8, 9	MTO–ABS	OTM–FAM
7825	M	9, 10	MTO–ABS	OTM–FAM
7827	M	9, 5	MTO–ABS	OTM–FAM

at the bottom right-hand corner of the screen for each correct response. For every 50th correct response, the participants could choose a sticker to place on a given medal. The numbers on the screen were shown only when the participants received programmed consequences. During training and testing, no other instructions were given, but the instructions already given could be repeated if the participants were uncertain about the details. The duration of the consequences was 1,000 ms, and the intertrial interval (ITI) was 500 ms.

Training. Two 5-member classes were trained and tested with a simultaneous protocol. Each trial started with the presentation of a sample stimulus at the top of the screen. When the participant clicked on the sample, two comparison stimuli were presented at the bottom corners of the screen. Every condition (i.e., OTM–familiar, OTM–abstract, MTO–familiar, and MTO–abstract) contained two phases of training and then a test phase for symmetry and equivalence (see Table 2 and 3). The first phase contained training on a serialized basis with 100% reinforcement in four different blocks. In each of these blocks, two relations were trained between sample stimuli and comparison stimuli, for example, A1B1A2 and A2B2A1. Each relation was presented four times, resulting in eight trials in a block. In the OTM training, AB, AC, AD, and AE relations were trained. The trials in the different blocks were A1B1B2, A2B1B2, A1C1C2, A2C1C2, A1D1D2, A2D1D2, A1E1E2, and A2E1E2. In MTO training, AE, BE, CE, and DE relations were trained. The trials in the different blocks were A1E1E2, A2E1E2, B1E1E2, B2E1E2, C1E1E2, C2E1E2, D1E1E2, and D2E1E2. When AB or AC relations (OTM or MTO, respectively) in the first block met the criterion that was set at 100%, or 8/8 correct responses, participants moved on to the next block, AC or BC, depending on the training structure. When participants reached the criterion of 100%, or 8/8 correct responses, in each of these blocks, they moved on to the last block in the first phase, which was a mix of training trials with 100% programmed consequences. All eight trained relations were trained four times in each mix, with 100% probability of programmed consequences, giving 32 different trial types. The criterion in the mixed block was a minimum of 30/32 correct responses.

Table 2. *Overview of the Training and Testing Blocks, Trials, and Fading Feedback for the OTM training structure*

Blocks	Trials	PPC	Minimum trials	Criterion
Training				
1. AB	A1B1B2, A2B1B2	100%	8	8/8
2. AC	A1C1C2, A2C1C2	100%	8	8/8
3. AD	A1D1D2, A2D1D2	100%	8	8/8
4. AE	A1E1E2, A2E1E2	100%	8	8/8
5. Mixed trials	A1B1B2, A2B1B2, A1C1C2, A2C1C2, A1D1D2, A2D1D2, A1E1E2, and A2E1E2	100%	32	30/32
6. Mixed trials	same as above	75%	32	30/32
7. Mixed trials	same as above	50%	32	30/32
8. Mixed trials	same as above	25%	32	30/32
9. Mixed trials	same as above	0%	32	30/32
Testing	A1B1B2, A2A1B2, A1C1C2, A2C1C2, A1D1D2, A2D1D2, A1E1E2, and A2E1E2 (dt trials).		32	29/32
Test block with direct trained trials (dt), symmetry, and equivalence trials randomly intermixed	B1A1A2, B2A1A2, C1A1A2, C2A1A2, D1A1A2, D2A1A2,E1A1A2, and E2A1A2 (sym trials).		32	29/32
	B1C1C2, B2C1C2, C1D1D2, C2D1D2, D1E1E2,D2E1E2, B1D1D2, B2D1D2, C1E1E2, C2E1E2, B1E1E2, B2E1E2, C1B1B2, C2B1B2, D1C1C2, D2C1C2, E1D1D2, E2D1D2, E1C1C2, E2C1C2, D1B1B2, D2B1B2, E1B1B2, and E2B1B2 (eq trials).		96	87/96

Note. The direct trained trials were intermixed with symmetry and equivalence trials in the test block. None of the trials in the test block were followed by any programmed consequence. PPC=Probability of Programmed Consequences.

The second phase of training contained four mixed blocks with thinning of probability of consequences (i.e., 75, 50, 25, and 0%). When participants had reached the respective criteria in all of the blocks in Phase 2, they received the test for symmetry and equivalence.

Testing. The test contained trials for directly trained relations (dt), relations for symmetry (sym), and relations for equivalence (eq). None of the test trials were followed by programmed consequences. The different trial types were presented in random order; that is dt, sym, and eq were randomly intermixed in one test block. All of the eight directly trained relations were tested four times each, which resulted in 32 dt trials. Symmetry relations in the OTM structure were BA, CA, DA, and EA (B1A1A2, B2A1A2, C1A1C2, C2C1A2, D1A1A2, D2A1A2, E1A1A2, and E2A1A2) and in the MTO structure were EA, EB, EC, and ED (E1A1A2, E2A1A2, E1B1B2, E2B1B2, E1C1C2, E2C1C2, E1D1D2, and E2D1D2). Thus there were eight relations in each condition, and each symmetry relation was tested four times, resulting in 32 sym trials. The relations tested for equivalence in the OTM structure were BC, CB, BD, DB, BE, EB, CD, DC, CE, EC, DE, and ED (B1C1C2, B2C1C2, C1B1B2, C2B1B2, B1D1D2, B2D1D2, D1B1B2, D2B1B2, B1E1E2, B2E1E2, E1B1B2, E2B1B2, C1D1D2, C2D1D2, D1C1C2, D2C1C2, C1E1E2, C2E1E2, E1C1C2, E2C1C2, D1E1E2, D2E1E2, E1D1D2, and E2D1D2). For the MTO structure the relations were: AB, BA, AC, CA, AD, DA, BC, CB, BD, DB, CD, and DC (A1B1B2, A2B1B2, B1A1A2, B2A1A2, A1C1C2, A2C1C2, C1A1A2, C2A1A2, A1D1D2, A2D1D2, D1A1A2, D2A1A2, B1C1C2,

Table 3. *Overview of the Training and Testing Blocks, Trials, and Fading Feedback for the MTO Training Structure*

Blocks	Trials	PPC	Minimum trials	Criterion
Training				
1. AE	A1E1E2, A2E1E2	100%	8	8/8
2. BE	B1E1E2, B2E1E2	100%	8	8/8
3. CE	C1E1E2, C2E1E2	100%	8	8/8
4. DE	D1E1E2, D2E1E2	100%	8	8/8
5. Mixed trials	A1E1E2, A2E1E2, B1E1E2, B2E1E2, C1E1E2, C2E1E2, D1E1E2, and D2E1E2	100%	32	30/32
6. Mixed trials	same as above	75%	32	30/32
7. Mixed trials	same as above	50%	32	30/32
8. Mixed trials	same as above	25%	32	30/32
9. Mixed trials	same as above	0%	32	30/32
Testing	A1E1E2, A2E1E2, B1E1E2, B2E1E2, C1E1E2, C2E1E2, D1E1E2, and D2E1E2 (dt trials).		32	29/32
Test block with direct trained trials (dt), symmetry, and equivalence trials randomly intermixed	E1A1A2, E2A1A2, E1B1B2, E2B1B2, E1C1C2, E2C1C2,E1D1D2, and E2D1D2 (sym trials).		32	29/32
	A1B1B2, A2B1B2, B1A1A2, B2A1A2, D1E1E2,D2E1E2, B1D1D2, B2D1D2, C1E1E2, C2E1E2, B1E1E2, B2E1E2, C1B1B2, C2B1B2, D1C1C2, D2C1C2, E1D1D2, E2D1D2, E1C1C2, E2C1C2, D1B1B2, D2B1B2, E1B1B2, and E2B1B2 (eq trials).		96	87/96

Note. The direct trained trials were intermixed with symmetry and equivalence trials in the test block. None of the trials in the test block were followed by any programmed consequence. PPC=Probability of Programmed Consequences.

B2C1C2, C1B1B2, C2B1B2, B1D1D2, B2D1D2, D1B1B2, D2B1B2, C1D1D2, C2D1D2, D1C1C2, and D2C1C2). This resulted in 24 relations tested for equivalence, with each relation tested four times, giving a total of 96 trials. Altogether, a total of 160 trials were tested, including 32 emergent relations, sym and eq. When the test was completed, the following text was presented on the screen: "Thank you for participating. You can now get your experimenter."

Recordings. The program recorded the results from the training and testing. All blocks were recorded, including the trial number, the presence or absence of programmed consequences, and the type of consequences. Relations between sample stimuli and comparison stimuli were registered, as well as what the participants chose as the comparisons. The program recorded the number of correct and incorrect responses and the reaction time from sample stimulus to a response, measured as the time from comparison onset to the participant clicking on one of the comparisons.

The accuracy criterion necessary to advance from one training phase to the next was a minimum of eight of 8 trials for the four blocks, then thirty out of 32 for the next five blocks. In the test block, definition of responding in accordance with symmetry and equivalence was defined as correct responding above 90%.

Results

Number of Trials and Errors in Baseline Relations

The individual data are shown in Tables 4 and 5. Participant 7815 had the highest number of trials in both the first and second conditions, with 1,728 total trials in MTO–familiar

Table 4. *Number of Trials and Errors for Each Participant and Scores on Direct Trained, Symmetry, and Equivalence Trials in Condition 1*

Participant	Condition 1	Trials	Incorrect	Dt	Sym	Eq
7801	OTM–Fam	264	14	**31/32**	**31/32**	**91/96**
7802	OTM–Fam	232	12	**31/32**	**31/32**	**95/96**
7803	OTM–Fam	264	27	**32/32**	**32/32**	**89/96**
7804	OTM–Fam	272	22	**32/32**	**29/32**	**89/96**
7806	OTM–Abs	872	234	**32/32**	**32/32**	**93/96**
7808	OTM–Abs	264	33	**32/32**	**30/32**	70/96
7818	OTM–Abs	328	30	**32/32**	**30/32**	**95/96**
7819	OTM–Abs	256	15	**32/32**	**32/32**	48/96
7811	MTO–Fam	328	56	**30/32**	**32/32**	**92/96**
7812	MTO–Fam	248	17	**32/32**	**30/32**	**93/96**
7814	MTO–Fam	312	40	**31/32**	**32/32**	**91/96**
7815	MTO–Fam	1728	381	**32/32**	**32/32**	**92/96**
7820	MTO–Abs	960	207	**30/32**	**32/32**	**88/96**
7823	MTO–Abs	736	148	21/32	14/32	51/96
7825	MTO–Abs	376	65	**29/32**	22/32	45/96
7827	MTO–Abs	536	86	**31/32**	**31/32**	85/96

Note. Numbers in bold indicate participants who met the criteria for equivalence relations. Dt = direct trained trials; Sym = symmetry trials; Eq = equivalence trials; OTM = one-to-many; MTO = many-to-one; fam = familiar and abstract stimuli; abs = abstract stimuli only.

Table 5. *Number of Trials and Errors for Each Participant and Scores on Baseline, Symmetry, and EquivalenceTrials During Condition 2*

Participant	**Condition 2**	**Trials**	**Errors**	**Dt**	**Sym**	**Eq**
7801	MTO–Abs	200	2	**32/32**	**31/32**	**90/96**
7802	MTO–Abs	216	10	**30/32**	**32/32**	**95/96**
7803	MTO–Abs	200	2	**32/32**	**32/32**	**96/96**
7804	MTO–Abs	512	45	20/32	21/32	40/96
7806	MTO–Fam	200	2	**32/32**	**31/32**	**95/96**
7808	MTO–Fam	256	13	**30/32**	**32/32**	**91/96**
7818	MTO–Fam	600	111	**32/32**	**31/32**	**91/96**
7819	MTO–Fam	232	7	**32/32**	**32/32**	66/96
7811	OTM–Abs	216	11	28/32	28/32	**89/96**
7812	OTM–Abs	208	3	**32/32**	**30/32**	**93/96**
7814	OTM–Abs	216	9	**32/32**	**32/32**	**94/96**
7815	OTM–Abs	1112	417	**32/32**	**30/32**	**96/96**
7820	OTM–Fam	440	40	**32/32**	**32/32**	**92/96**
7823	OTM–Fam	776	131	25/32	28/32	**87/96**
7825	OTM–Fam	216	9	**32/32**	**32/32**	48/96
7827	OTM–Fam	368	40	**31/32**	28/32	84/96

Note. Numbers in bold indicate participants who met the criteria for equivalence relations. Dt = Direct trained trials; Sym = symmetry trials; Eq = equivalence trials; OTM = one-to-many; MTO = many-to-one; Fam = familiar and abstract stimuli; Abs = abstract stimuli only.

and 1,112 total trials in the OTM–abstract condition. Participant 7802 had the lowest number of trials in the first condition, with a total of 232 trials. In the second condition, Participants 7801, 7803, and 7806 had the lowest number of trials, with a total of 200. All participants had fewer responses to criterion in Condition 2 as compared to Condition 1, except for Participants 7804, 7818, and 7823.

Participant 7815 had the highest number of errors in training in both conditions, with

381 errors in Condition 1 and 471 errors in Condition 2. The participant with the lowest number of errors in training in Condition 1 was Participant 7802, with 12 errors, and the participants in Condition 2 were Participants 7801, 7803, and 7806, each with only 2 errors.

As can be seen in the left-side panel in Figure 2, the median number of trials to criterion for the conditional discrimination training was highest for the fourth group, who started with MTO–abstract (i.e., 636 trials and 404 trials). The lowest number of trials was for the first group, who started with the condition OTM–familiar (i.e., 264 trials). For all groups there was a reduction in the median number of trials from Condition 1 to Condition 2.

The fourth group, MTO–abstract and OTM–familiar, had the highest median number of errors in Condition 1 (i.e., 148 errors) and Condition 2 (14 errors; see the right-side panel in Figure 2). The first group, OTM–familiar and MTO–abstract, had the lowest number of errors in both conditions, with 14 and 2 errors, respectively. For all groups there was a reduction in the median number of errors from Condition 1 to Condition 2.

Emergent Relations

As shown in Tables 3 and 4, 11 of 16 participants reached the criterion for equivalence relations in the first condition and 12 of 16 reached it in the second condition. For the symmetry relations, 14 of 16 reached the criterion in the first condition and 12 of 16 in the second condition. The scores show that in the all-abstract condition, six participants did not respond according to equivalence relations, while three 3 participants failed to reach the criterion for equivalence relations with familiar stimuli. Fifteen of 16 participants and thirteen of 16 participants had the trained conditional discriminations intact during testing in the first condition and second condition, respectively.

Participants 7808, 7819, 7823, 7825, and 7827, all with abstract stimuli, did not reach the criterion for equivalence relations in the first condition. Two of these participants had the OTM training structure and three had MTO. Two of these five participants, Participants 7823 and 7825, both in the MTO–abstract condition, did not reach the criterion for symmetry relations either. Participant 7823 did not have the baseline relations intact during the test.

In the second condition, Participants 7804, 7819, 7825, and 7827 did not reach

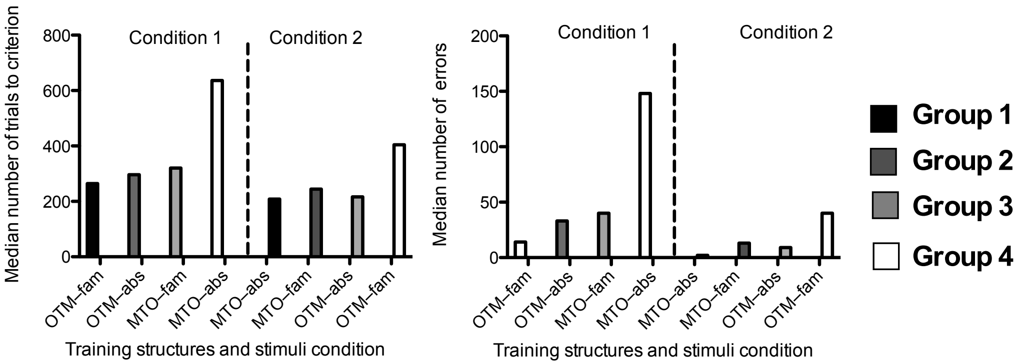

Figure 2. In the left-side panel, the number of trials to criterion before testing is shown for the different combinations of training structures and stimuli. The left side of this panel is for the first condition and the right side is for the second condition. In the right-side panel, the number of errors is shown for the different combinations of training structures. The left side of this panel is for the firs condition and the right side is for the second condition.

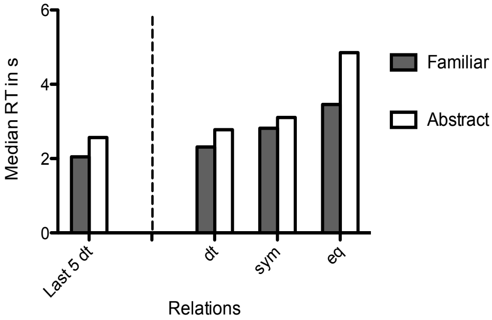

Figure 3. The median reaction time in seconds is shown for last five direct trained trials and for the first five direct trained, symmetry, and equivalence trials in the test block. The grey bars are for familiar stimuli and the white bars are for the abstract stimuli.

the criterion for equivalence relations. Participant 7804 did not respond in accord with stimulus equivalence in the MTO–abstract condition; the three others did not reach the criterion for equivalence with the familiar stimuli set, one participant with MTO and two with OTM. Participants 7804 and 7827 also did not reach the criterion for symmetry. Two other participants did not reach the criterion for symmetry in the second condition, Participant 7811 in the OTM–abstract condition and Participant 7823 in the OTM–familiar.

In three of four conditions with familiar stimuli, all participants responded in accordance with stimulus equivalence, whereas in none of the conditions with abstract stimuli only did all participants respond in accordance with stimulus equivalence.

Reaction Time

In the current experiment, reaction time was calculated as the median of the last five direct trained trials and the first five trials in the test block for direct trained, symmetry, and equivalence trials, respectively. Median reaction times for the conditions with familiar stimuli are plotted against the reaction times for conditions with abstract stimuli for the first conditions. The increase from the last five direct trained trials to equivalence trials is also the most pronounced increase.

Discussion

The present study was a replication of Smeets and Barnes-Holmes' (2005) experiment with some modifications, which will be discussed in depth later in this section. In the present experiment, the purpose was to study responding in accordance with stimulus equivalence as a function of familiar or abstract stimuli only in combination with different training structures, MTO or OTM. The results showed that familiar stimuli were more effective in establishing responding in accordance with stimulus equivalence than were abstract stimuli only, independent of training structures. When participants received familiar stimuli in the first condition, more of them responded in accordance with equivalence relations in the second condition compared to participants who received abstract stimuli only. This could be due to learning history or order effect, which means that the order of presentation of the procedure may have influenced the results (Shadish, Cook, & Campbell, 2002). However, such carry-over effects are quite challenging to control for when using an experimental design with repeated exposure to conditions. Furthermore, there were small differences between the training structures, but responding in accordance with stimulus equivalence was higher with the OTM training structure than with MTO. Nevertheless, in the second condition, an equal number of participants reached criterion for equivalence relations with different training structures.

With respect to training the conditional discriminations, there was no significant difference in number of trials or number of errors when comparing participants with familiar stimuli and those with abstract stimuli only, except for the fourth group with the MTO–abstract and OTM–familiar conditions. During conditional discrimination training in the first condition with OTM, there were fewer trials and fewer errors than for MTO. In addition, the groups who received OTM in the first condition also showed fewer trials and fewer errors in the second condition than the two other groups. All groups responded with fewer trials and fewer errors in the second condition. Most participants who started with the familiar stimuli had fewer trials in both the first and the second conditions than the participants who started with abstract stimuli only. Some of the participants, who received abstract stimuli in the first condition and had many trials and errors, repeated this pattern in the second condition. One participant who first received MTO–familiar and then OTM–abstract increased the number of trials and errors; the other participants had a low number of both trials and errors. Smeets and Barnes-Holmes (2005) reported no differences with respect to training trials, although it is difficult to interpret this result because they only reported blocks of trials.

Most of the participants who first received the OTM condition responded to the criterion for establishing the conditional discriminations with fewer trials in both conditions than those who started with MTO. In their discrimination analysis, Saunders and Green (1999) also predicted that MTO would be more time consuming in the conditional discrimination training than OTM. This prediction was in line with the current findings. The participants who started with OTM also had fewer trials and errors with MTO in the second condition, which could be due to an order effect. Arntzen and Holth (2000a) and Fields et al. (1999) also found that MTO is more time consuming than OTM in training because it requires more discriminations. Furthermore, the prediction is that MTO should be more effective in establishing equivalence relations, but this prediction was not supported by the present results. The findings are in accordance with other studies which have tested the differences between the training structures (Arntzen et al., 2010; Arntzen & Holth, 1997, 2000a). The discrimination analysis also predicted that the difference should be more pronounced with a higher number of class members, which is not in accordance with the present study and the Smeets and Barnes-Holmes (2005) study with the use of two 5-member classes. Also, this notion has not been supported when the number of members in the classes has been increased from three to four with abstract stimuli (e.g., Arntzen et al., 2010; Arntzen & Vaidya, 2008). However, Fields et al. (1999) did find that MTO was superior to OTM in establishing two 5-member and 7-member classes with nonsense syllables. Hence, future research should focus on large classes and with three classes or more and with different types of stimuli.

Smeets and Barnes-Holmes (2005) did find that the all-abstract-stimuli condition was more effective than the familiar-stimuli condition when testing for stimulus equivalence. Nevertheless, the present results indicate that the use of familiar stimuli is more effective in establishing conditional discriminations and responding in accordance with emergent relations. The present findings are consistent with results from other studies using familiar stimuli (Arntzen, 2004; Arntzen & Holth, 2000b; Arntzen & Lian, 2010; Bentall, Dickins, & Fox, 1993; Mandell & Sheen, 1994).

The reason for differences in outcome can be related to different procedural discrepancies between the present study and the study by Smeets and Barnes-Holmes' (2005). In their experiment, they used specific instructions that involved naming of some stimuli used in the conditional discrimination. Teaching the participants to name the abstract stimuli may have been a good experimental strategy in the sense that "familiar" and "abstract" stimuli have names and if they showed different performances it should be assigned to other features of the stimuli, which still remain to be identified. The instructions involved telling the children which stimuli to choose, and the conditional discrimination started with instructions in both the first and the second trial. The instruction in Smeets and Barnes-Holmes' study in the all-abstract-stimuli condition was as follows: "This (A1) is an apple, this (B1) is a nose, and that (B2) is a flag. Point to the nose" (p. 286). In the familiar-stimuli condition the instruction was "This (A1) is a smiley face, this (B1) is a nose, and that (B2) is a flag. Point to the nose" (p. 286). The children were required to name four stimuli from the beginning of training in one condition. Hence, Arntzen and Lian (2010) suggested that familiar stimuli can yield mediating behavior, which can affect training of conditional relations and tests for emergent relations. One explanation of the higher yields with the familiar stimuli in the current study could be that they were caused by the fact that the stimuli were nameable. Another type of data that could indicate some mediating behavior is the reaction time date. The median reaction time (RT) results in the present experiment showed that there was an increase from the last five direct trained trials to test for all groups and a more pronounced increase for the all-abstract-stimuli condition than for the familiar-stimuli condition, which is consistent with other the findings with children as participants in our lab (e.g., Arntzen & Lian, 2010).

The results from the current study and from Arntzen and Lian (2010) call for experiments in which the functions of different stimuli are studied in more detail. Because the participants in the Smeets and Barnes-Holmes study were also given names for some of the abstract stimuli, it can be questioned whether the use of abstract stimuli versus familiar stimuli made any difference for the results. Naming of a stimulus in one class may lead to all of the stimuli being named the same when the class controls the same response (Mandell, 1997).

Sidman (1994) discussed the possibility that some words or expressions may be part of an equivalence class in the participant's repertoire and therefore interfere with the variables one is trying to study: "Then too, if one teaches labels for stimuli by using terms like goes with, *is the name of, means, matches, is the same as*, and so forth, the instructions rather than the baseline conditional discriminations may be responsible for subsequently manifested equivalence relations" (pp. 305–306). However, such instructions have been used for different reasons. For example, in a study by Devany, Hayes, and Nelson (1986), participants were instructed to "touch the one that goes with this one." Therefore, a precise definition of instructions is of great importance in interpreting the findings of a study. Pilgrim, Jackson, and Galizio (2000) studied the effects of instructions and naming of sample stimuli in 3- to 7-year-old children. The results showed that general instructions were not of importance in establishing conditional discriminations. On the other hand, specific instructions such as "[w]hen this one is in the middle, pick this one," in addition to requiring participants to name sample stimuli, had a facilitative effect on establishment of conditional discriminations for most of the participants. However, a number of variables can influence the effect of a specific set of instructions, including when they are given and their content. Arntzen, Vaidya, and Halstadtro (2008) conducted an experiment to find out more about these variables. The specific instructions were presented when the children were in the first phase of conditional discrimination training, AB trials, without any increase in correct responding. The children were instructed that "these stimuli belong together" while the experimenter pointed to the stimuli. The data showed facilitation of the establishment of the first relation and the second relation for all participants in the study.

In addition to the difference in instructions, there were other procedural differences between the current study and Smeets and Barnes-Holmes (2005). First, the current study was performed on a computer, whereas the previous study was performed using a table-top procedure and recording. Even if the Smeets and Barnes-Holmes' study was carried out by experimenters with extensive training, it is possible that the participants' responding came under incorrect stimulus control. However, it is not obvious how that could have influenced only one of the training structures or the different stimulus sets. The present concern is more with the experimental control in the Smeets and Barnes-Holmes' study. Second, as described in the procedural section in Smeets and Barnes-Holmes, the stimuli were presented on stimulus cards, with one sample stimulus and two comparison stimuli on each card. It is not clear if there was a requirement of a response to sample stimulus or not. Previous experiments have shown that when comparisons are presented without any response to sample stimulus, the number of trials to criterion can increase (Lian & Arntzen, 2010). However, it is not mentioned in the Smeets and Barnes-Holmes study if the stimulus positions were random for the comparisons. Third, Smeets and Barnes-Holmes suggest that having symmetry and equivalence trials intermixed in a testing block could influence the responding on equivalence trials. The order of test trials could be an important procedural variable, but it is not obvious how this should influence the outcome of training structures. This calls for future experiments to find out the effect of the order types of test trials. Fourth, the Smeets and Barnes-Holmes study included a pretraining phase with identity matching. Identity matching could actually hinder arbitrary matching because the task is to find the similar stimuli in a stimuli array, which is not the case for arbitrary matching.

As in Saunders et al. (1999) and Smeets and Barnes-Holmes (2005), we used a two-choice format in the MTS procedure. Sidman (1987) argued against the use of two-choice conditional discrimination. When there are only two comparison stimuli to choose between, the chosen stimulus is either a correct or an incorrect match with the sample stimulus. Relating the sample to the comparison can come under the wrong stimulus control, if, for example, the participant rejects a comparison that is incorrect instead of finding the correct comparison to the sample. Sidman also expressed concern that conditional relations are not established with a two-choice format. However, Boelens (2002) argued that the criterion for conditional discrimination should be high enough to prevent Sidman's argument from being valid. In the present experiment, the criterion for conditional discrimination is set to 100% in the first phase and 93% in the following phase, with gradual thinning of reinforcement. Finally, according to Sidman, with the use of two choices, variables other than those established in conditional discrimination can control responding to emergent relations. If the classes are large and the person can respond correctly to all emergent relations, the chance for other variables to control behavior decreases. On the other hand, Smeets and Barnes-Holmes used a criterion of 83% for equivalence trials and 87% for symmetry trials, which is quite low when it is only a two-choice situation.

There are some limitations with the present experiment. First, we have a quite small number of participants in each group. The experiment was designed like this since we wanted to replicate the Smeets and Barnes-Holmes (2005) study. However, future research should be designed with a larger number of participants. Second, there is an age-difference between the children in present study and the Smeets and Barnes-Holmes study. Actually, one should predict that the use of pictures should have been even more effective with younger children. Therefore, future research should include younger children.

In summary, the present results showed that the the familiar-stimuli conditions were more effective in producing responding in accordance with stimulus equivalence relations than all-abstract-stimuli conditions. Furthermore, the results showed little difference between using the MTO and OTM training structures with respect to responding in accord with emergent relations.

However, OTM was more effective in the first condition. The present findings on the stimuli types , familiar and abstract stimuli, diverged from those of Smeets and Barnes-Holmes' (2005). This could be because they instructed participants to name the abstract stimuli, which could facilitate stimulus naming in both the all-abstract-stimuli and familiar-stimuli conditions. Finally, the variability in the findings related to different training structures can be related to differences in procedures, like the use of specific instructions, pre-training or not, number of comparisons, and the use of different stimuli. Further research should focus on different variables which can be influence the equivalence outcome with the use of different training structures.

References

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

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., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47*, 309–320.

Arntzen, E., & Holth, P. (2000a). Differential probabilities of equivalence outcome in individual subjects as a function of training structure. *The Psychological Record, 50*, 603–628.

Arntzen, E., & Holth, P. (2000b). Probability of stimulus equivalence as a function of class size vs. number of classes. *The Psychological Record, 50*, 79–104.

Arntzen, E., & Lian, T. (2010). Trained and derived relations with pictures as nodes. *The Psychological Record, 60*, 659–677.

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.

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.

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

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. doi: 10.1901/jeab.1986.46-243

Fields, L., Hobbie-Reeve, S. A., Adams, B. J., & Reeve, K. F. (1999). Effects of training directionality and class size on equivalence class formation by adults. *The Psychological Record, 49*, 703–724.

Imam, A. A. (2006). Experimental control of nodality via equal presentations of conditional discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of the Experimental Analysis of Behavior, 85*, 107–124. doi: 10.1901/jeab.2006.58-04

Lian, T., & Arntzen, E. (2010). Matching-to-sample tasks and observing responses. *Poster presented at the annual convention of the Association for Behavior Analysis, San Antonio, TX.*

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. doi: 10.1016/S0166-4115(97)80134-3

Mandell, C., & Sheen, V. (1994). Equivalence class formation as a function of the pronounceability of the sample stimulus. *Behavioural Processes, 32*, 29–46. doi: 10.1016/0376-6357(94)90025-6

Pilgrim, C., Jackson, J., & Galizio, M. (2000). Acquisition of arbitrarily conditional discriminations by young normally developing children. *Journal of the Experimental Analysis of Behavior, 73*, 177–193.

Saunders, K. J., Saunders, R. R., Williams, D. C., & Spradlin, J. E. (1993). An interaction of instructions and training design on stimulus class formation: Extending the analysis of equivalence. *The Psychological Record, 43*, 725–744.

Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of the Experimental Analysis of Behavior, 71*, 195–214.

Saunders, R. R., & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior, 72*, 117–137. doi: 10.1901/jeab.1999.72-117

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. doi: 10.1901/jeab.1988.49-95

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin Company.

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

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

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*, 5–22. doi: 10.1901/jeab.1982.37-5

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. doi: 10.1016/j.beproc.2004.12.009

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. doi: 10.1016/0270-4684(86)90005-4