EUROPEAN JOURNAL OF BEHAVIOR ANALYSIS 2011, **12**, 483 - 503 NUMBER 2 (WINTER 2011)

Training Structures and the Formation of Equivalence Classes

Erik Arntzen and Steffen Hansen
Oslo and Akershus University College

Different training structures have produced different outcomes of equivalence yields when using a simultaneous training protocol. According to the discrimination analysis by R. R. Saunders and Green (1999), the number of simple discriminations required in conditional discrimination differs for the different training structures. Hence, for the MTO (many-to-one) structure, all the simple discriminations are required during training of the conditional discriminations, while they are not in the OTM (one-to-many) or LS (linear series) training structures. The differences in the number of discriminations will be greater when the number of members in the classes increases; thus, there should be a more pronounced difference in the outcome on equivalence tests. The purpose of the present experiment was to train 3 potential classes of 3 members each and 3 potential classes of 6 members each with MTO, OTM, and LS training structures. Thirty adult participants were randomly assigned to 6 groups, that is, 3 or 6 members with MTO, OTM, or LS. The results showed that there were small differences in the outcome following MTO or OTM. However, the equivalence outcome was lowest following the LS training structure. The data from the follow-up tests showed that none of the participants in the LS groups responded in accordance with stimulus equivalence, while 6 of 10 did so in the OTM groups and 3 of 8 did in the MTO groups. Moreover, the reaction time data showed that there was an increase from directly trained trials to the symmetry and equivalence trials, with a more pronounced increase for equivalence trials.

Key words: training structures, MTO, LS, OTM, large classes, stimulus equivalence

In behavior analysis, the most common way to train the necessary prerequisites for testing for derived relations is to use conditional discrimination procedures. These procedures can produce emergent relations, like stimulus equivalence. Stimulus equivalence is defined as responding in accordance with reflexivity, symmetry, and transitivity (Sidman, 1994; Sidman & Tailby, 1982). Three different training structures are typically used when establishing conditional relations: many-to-one (MTO), one-to-many (OTM), and linear series (LS) (e.g., K. J. Saunders, Saunders, Williams, & Spradlin, 1993). In an MTO training structure, many sample stimuli are trained to one comparison stimulus, while in OTM, one sample stimulus is trained to many comparisons. In an LS training structure, the nodal stimulus shift in one conditional relation from functioning as a comparison to a sample in the next condtional relation (Fields & Verhave, 1987). Even if the present experiment is concerned with basic research questions, knowledge about the differential effectiveness is important for teaching programs.

The literature is not consistent with respect to the equivalence outcome of the different training structures. However, when using a simultaneous training and testing protocol

Author note: Correspondence concerning this manuscript should be addressed to Erik Arntzen, Oslo and Akershus University College, Department for Behavioral Science, PO Box 4 St. Olavs Plass, 0130 Oslo, Norway. E-Mail: erik.arntzen@equivalence.net.

(for a discussion about protocols, see Fields, Landon-Jimenez, Buffington, & Adams, 1995; Imam, 2006), the LS training structure has been shown to be the least effective structure (e.g., Arntzen, Grondahl, & Eilifsen, 2010; Arntzen & Holth, 1997, 2000). The MTO training structure has been shown to be superior to the OTM training structure in some studies (Fields, Hobbie-Reeve, Adams, & Reeve, 1999; Hove, 2003; K. J. Saunders et al., 1993; R. R. Saunders, Chaney, & Marquis, 2005; R. R. Saunders, Drake, & Spradlin, 1999; R. R. Saunders & McEntee, 2004), OTM has shown to be more effective in others (Arntzen & Holth, 1997, 2000), and, finally, others have found no differences or very small differences between OTM and MTO (e.g., Arntzen et al., 2010; Arntzen & Nikolaisen, 2011; Arntzen & Vaidya, 2008; Smeets & Barnes-Holmes, 2005).

Reaction time is not a part of the definition of stimulus equivalence; however, it could be an important additional measure in training conditional discriminations and emergent relations (e.g., Dymond & Rehfeldt, 2001; Pilgrim & Galizio, 2000).

R. R. Saunders and Green (1999) hypothesized in their discrimination analysis that the probability of responding in accordance with stimulus equivalence would differ as a function of training structure. Saunders and Green's hypothesis was rooted in the analysis of successive simple discriminations and simultaneous simple discriminations that were contained in the conditional discrimination training and testing within the different training structures. For example, if the training structures are going to produce three 3-member equivalence classes, 36 simple discriminations are needed in the training as prerequisites for testing. According to their analysis, only the MTO training structure incorporates all 36 simple discriminations that are required for repeated positive results on tests for equivalence class formation. On the other hand, in the OTM and LS training structures, only 27 of 36 discriminations are presented during training. Thus, when the number of classes and class sizes increase, the total number of simple discriminations in each of the three training structures increases exponentially. Saunders and Green's analysis suggests, therefore, that an expansion of the number of members in each stimulus class, as well as an expansion of the number of stimulus classes, would make the difference in responding in accordance with stimulus equivalence even more pronounced. For example, if the training structures are going to produce three 6-member classes, the OTM and LS training structures include only 90 of the 153 simple discriminations needed as prerequisites for a positive outcome on the tests for derived relations. Therefore, the MTO structure, having trained all the required simple discriminations, should consistently show better results on tests for equivalence class formation.

In Fields et al.'s (1999) study, 70 undergraduate students were randomly assigned to one of four experimental groups, and potentially two 5- and 7-member equivalence classes were established. Furthermore, for half of the participants, an MTO training structure was used to establish the conditional discriminations, while for the remaining participants an OTM training structure was used. In addition, each conditional baseline relation was trained within a single training block, and every trial type was randomly presented an equal number of times. Fifty-six of 70 participants acquired the baseline discrimination relations. The test block consisted of baseline conditional relations, with programmed consequences, and probes for symmetry and equivalence, without programmed consequences. The results showed no effects of training structure for the two 5-member equivalence classes. However, a larger proportion of participants showed the emergence of 7-member

equivalence classes after MTO than after OTM. In addition, the baseline conditional relations for the 7-member classes were learned more slowly during MTO than during OTM. In other words, the result is in accordance with Saunders and Green's (1999) discrimination analysis. Nonetheless, in each experimental group, Fields et al. trained two stimulus classes only, and, as is evident in the following studies, in order to study the effects of rejection control and component simple discriminations, there is a need for experiments with three or more stimulus classes per experimental condition.

R. R. Saunders et al. (2005) looked at equivalence class establishment by senior citizens. Eighteen participants were introduced to a two-, three-, and four-choice matching-to-sample format in which 18 conditional relations were trained using the MTO, OTM, and LS structures. Results showed that training in a three- and four-choice MTS format, compared to a two-choice MTS format, did not increase the probability of equivalence class formation, except slightly in the MTO structure. Also noteworthy, the OTM and MTO structures notwithstanding, the LS training structure did show a decrease in the formation of equivalence relations as both the number of stimuli per class and the number of classes increased. Otherwise, there were no significant differences among the three structures with regard to the number of training trials needed to pass the training phase.

Attempting to replicate and extend the R. R. Saunders et al. (2005) findings, Arntzen et al. (2010) investigated the differential effects of different training structures in the establishment of conditional discriminations with potentially three 3-member classes and three 4-member classes. Twelve participants were recruited and trained using a single-subject design. Furthermore, for all three training structures, participants were first trained to produce three 3-member classes and, subsequently, three 4-member classes. The training was introduced in two phases and the testing in one phase only In Phase 1, the number of comparison stimuli increased from one to two, and in Phase 2 all comparisons were presented. In Phase 3, participants were tested for emergent relations, that is, symmetry, transitivity (for LS), and equivalence.

In line with the R. R. Saunders et al. (2005) study, the results showed no substantial differences in emergent relations between the OTM and MTO training structures as class size increased from three to four members. Furthermore, the LS structure once again resulted in the lowest score on tests for equivalence class formation. In addition, the OTM structure indicated slightly better results than the MTO structure when the tests were arranged for the potential three 3-member classes. However, there was no such difference when testing for equivalence formation between the three potential classes of four stimuli each. Also, the OTM structure indicated fewer training trials than the MTO structure in order for participants to reach criterion for testing. Regarding reaction time, Arntzen et al. (2010) found that the reaction time to comparison stimuli was slower for equivalence trials than symmetry trials. Compared to baseline trials, response speed was slower for both symmetry and equivalence trials.

The purpose of the present study was to replicate and extend the Arntzen et al. (2010) findings. First, the present study investigated the differential probabilities of equivalence class formation as a function of increasing number of class members, that is, 3 and 6. Second, this study attempted to replicate earlier findings on reaction time, which have shown an increase from baseline to testing as well as a greater increase in reaction time on equivalence than symmetry tests. Finally, inspired by Rehfeldt and Hayes (2000), the present experiment also

looked at the differential effects of training structures on class formation in a follow-up test between 2 and 4 weeks after the initial training and testing.

Method

General Overview

Participants were randomly placed in one of six different groups. The first three groups conformed to the MTO, OTM, or LS training structure, all with the potential to produce three 3-member equivalence classes.

The last three groups consisted of the MTO, OTM, and LS training structures, with the potential to produce three 6-member equivalence classes.

Participants

A total of 30 adults participated in the study, 23 of whom were college students (18 female, 5 male). The remaining 7 participants—4 males and 3 females—were recruited through personal contacts (i.e., Participants 2620, 2625, 2626, 2627, 2628, 2629, and 2630) and worked in full-time positions. The average age was 36.8 years for the MTO group with 3 potential classes and 3 stimuli each (2601, 2602, 2613, 2618, and 2619), 26.4 years for the MTO group with 3 potential classes and 6 stimuli each (2606, 2607, 2612, 2624, and 2630), 24.8 years for the OTM group with 3 potential classes and 3 stimuli each (2605, 2609, 2618, 2621, and 2622), 26 years for the OTM group with 3 potential classes and 6 stimuli each (2603, 2610, 2616, 2623, and 2627), 34 years for the LS group with 3 potential classes and 3 stimuli each (2604, 2608, 2617, 2619, and 2626), and 28.4 years for the LS group with 3 potential classes and 6 stimuli each (2611, 2614, 2615, 2620, and 2625). Participants were selected based on their willingness to participate and their lack of knowledge of research within the specific area of stimulus equivalence. Before the study, the participants were asked to read and sign a consent form. In addition, they were informed that they could withdraw from the study at any time. Upon completion of training and testing, each participant was debriefed on the purpose of the study, and an article on stimulus equivalence was offered.

Apparatus and Software

An HP EliteBook 8740w, with the processor specifications Intel Core i5 CPU M520 @ 2.40GHz 2.40 GHz, 2.98GB RAM, was used to run the experiments. The computer was equipped with a mouse and was connected to a 17-in. monitor. Software developed by Psych Fusion Ltd in collaboration with the first author was used. The software presented all stimuli and recorded responses to sample and comparison. In addition, reaction time was recorded.

Setting

The experimental sessions were conducted in two different laboratories. All tests occurred in rooms with the door closed, in order to eliminate external noise as well as to make sure that no interruptions would occur. For the students enrolled in the university, the test room was equipped with a table, a chair, and a bookcase. The remaining seven adults were tested in a room that was equipped with a chair, a table, two bookcases, and a bed. The test rooms were of different sizes. The rooms were 3 x 3 meters and 3 x 5 meters. One of the rooms had one window; however, in both rooms, participants were positioned facing a blank wall, with a blank wall on either side of them.

Stimulus Material

Eighteen different stimuli were presented on the computer screen, of which the first 9 were used in the training to produce three 3-member equivalence classes. Because it was a between-subjects design (i.e., each participant was only subjected to one of the six different experimental conditions), it was possible to use the same stimuli within all six different training structures.

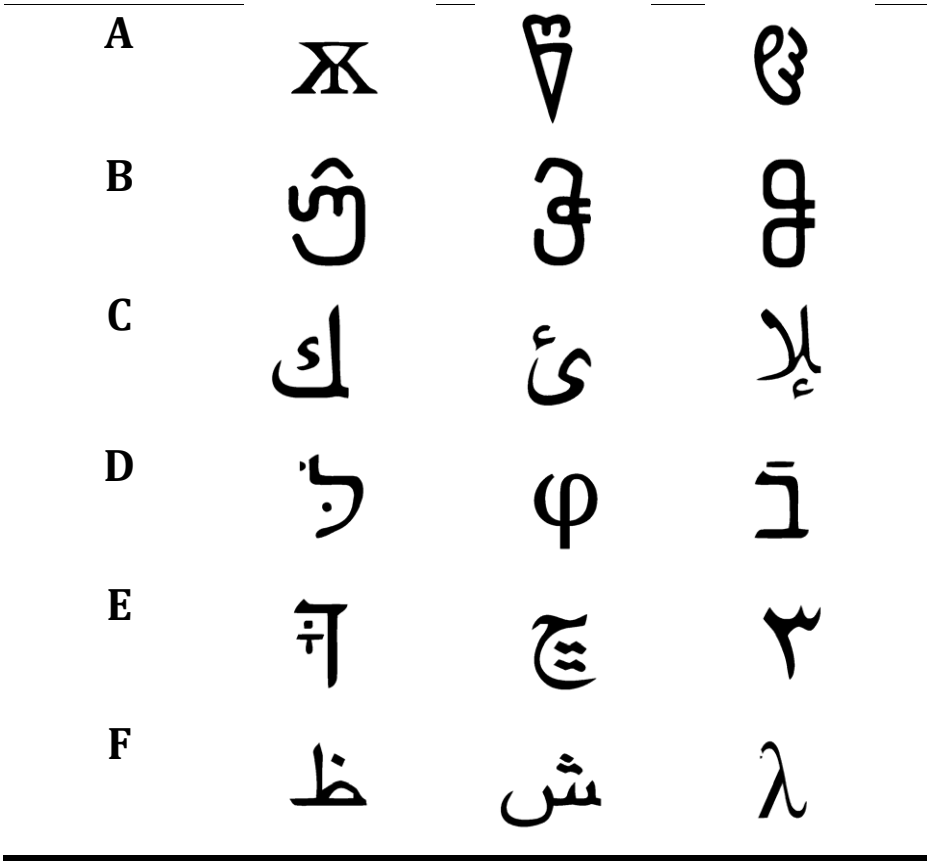

Figure 1. The stimuli A through C were used in the three 3-member class conditions, where as stimuli A through F were used in the three 6-member class conditions.

The visual stimuli consisted of letters from the Greek, Japanese, Arabic, and Cyrillic alphabets (see Figure 1).

Procedure

General information to participants. During presentation and recruitment of potential participants, students were told that this was a study on the psychology of learning, in which they would have to respond to stimuli that were presented on a computer screen. In addition, they were informed that the duration of the experiment would be somewhere between 45 min and 3 hrs. Furthermore, a follow-up phase was to take place approximately 2 to 3 weeks after the initial experimental session. As an incentive, participants were told that they would receive 50 kroner (approximately $8) upon completion of the follow-up phase.

Instructions. When a participant had positioned him- or herself in front of the computer, he or she was told to press the "Start" square at the lower half of the screen. When this square was clicked, the following appeared:

A stimulus will appear in the middle of the screen. Click on this by using the computer mouse. Three other stimuli will then appear. Choose one of these by using the computer mouse. If you choose the stimulus that we have defined as correct, words like *good*, *super*, and so on will appear on the screen. If you press a wrong stimulus, the word *wrong* will appear on the screen. The number of correct responses you have made will be counted

and displayed at the bottom of the screen. During some stages of the experiment, the computer will not tell you if your choices are correct or wrong. However, based on what you have learned, you can get all the tasks correct. Please do your best to get everything right. Good luck! Press Start to begin the experiment.

Training and testing. All trials began with the presentation of the sample stimulus in the middle of the screen. When the participant clicked on the sample stimulus, three comparison stimuli would randomly appear in three of the four corners on the computer screen. The sample stimulus would remain on the computer screen until the participant clicked on one of the comparison stimuli. After each trial, the mouse cursor would automatically be reset to a predetermined position on the computer screen after an intertrial interval that lasted 1,000 ms. Depending on whether the participant chose what the experimenter had defined as either the correct or incorrect relation between sample and comparison stimuli, a word such as *Excellent*, *Great*, or *Wrong* would appear in the middle of the computer screen for 500 ms.

An overview of the training and testing block phases, trials, and programmed consequences for the three training structures with, respectively, the potential three 3-member equivalence classes, and three 6-member equivalence classes are presented in Tables 1 through 6.

Trial types in the MTO training structure. As shown in Table 1, the acquisition phase for the potential three 3-member equivalence classes consisted of two blocks in which the conditional relations AC and BC were established on a serialized basis (i.e., AC trials were trained first, then BC trials). After training the first two conditional relations, a mixed-trial block would randomly train all the conditional relations (e.g., AC/BC trials). The trained trial types were as follows: A1C1C2C3, A2C1C2C3, A3C1C2C3, B1C1C2C3, B2C1C2C3, and B3C1C2C3. In all strings, the first code is the sample stimulus and the underlined comparison is the correct stimulus choice. Trials were randomly presented five times each before the next training block was introduced. To move on to the next training block, a participant was required to have at least 14 out of 15 trials correct. In the maintenance phase (i.e., Blocks 3 to 6 in Table 1), the participants were introduced to four blocks with mixed training trials. Moreover, all trial types were presented five times each on a random basis. Thus, the number of trials was 30. As in the training phase, the participants needed at

Table 1. *Number of Trials and Probability of Programmed Consequences in the training and testing blocks used to establish 3-Member Equivalence Classes with MTO structure*

Blocks	Probability for programmed consequences (%)	Minimum trials	Criterion
Training			
1. AC	100	15	14/15
2. BC	100	15	14/15
3. Mixed Trials AC, BC	100	30	27/30
4. Mixed Trials AC, BC	75	30	27/30
5. Mixed Trials AC, BC	25	30	27/30
6. Mixed Trials AC, BC	0	30	27/30
Testing			
Test block with *DT trials* (AC, BC),	0	30	27/30
SYM trials (CA, CB), and	0	30	27/30
EQ trials (AB, BA) randomly intermixed	0	30	27/30

Note. DT = Direct trained; SYM = symmetry; EQ = equivalence.

Table 2. *Number of Trials and Probability of Programmed Consequences in the training and testing blocks used to establish 6-Member Equivalence Classes with MTO structure*

Blocks	Probability of programmed consequences (%)	Minimum trials	Criterion
Training			
1. AF	100	15	14/15
2. BF	100	15	14/15
3. Mixed Trials AF, BF	100	30	27/30
4. CF	100	15	14/15
5. Mixed Trials AF, BF, CF	100	60	54/60
6. DF	100	15	14/15
7. Mixed Trials AF, BF, CF, DF	100	75	68/75
8. EF	100	15	14/15
9. Mixed Trials AF, BF, CF, DF, EF	100	90	81/90
10. Mixed Trials AF, BF, CF, DF, EF	75	75	68/75
11. Mixed Trials AF, BF, CF, DF, EF	25	75	68/75
12. Mixed Trials AF, BF, CF, DF, EF	0	75	68/75
Testing			
Test block with *DT trials* (AF, BF, CF, DF, EF),	0	75	68/75
SYM trials (FA, FB, FC, FD, FE), and	0	75	68/75
EQ trials (AB, BA, BC, CB, CD, DC, DE, ED, AC, CA, BD, DB, CE, EC, AD, DA, BE, EB, AE, EA) randomly intermixed	0	300	270/300

Note. Due to a programming error, the number of trials in block 5 increased with 30 trials instead of 15 trials. Thus, the extra number of trials is presented in blocks 5, 7, and 9. The 15 extra trials are not presented in blocks 10–12. DT = Direct trained; SYM = symmetry; EQ = equivalence.

least 90% correct (27 out of 30) in order to move on to the next block. In addition, the probability of a programmed consequence in the first maintenance block was 100%, in the second block the probability of a programmed consequence was 75%, in the third block it was 25%, and in the last block the probability was zero.

Following the training and maintenance phases, the participants (see the lower part of Table 1) were introduced to a testing block that consisted of a random intermix of 30 baseline, 30 symmetry, and 30 equivalence trials. The trial types were A1C1C2C3, A2C1C2C3, A3C1C2C3, B1C1C2C3, B2C1C2C3, B3C1C2C3 (i.e., directly trained trials), C1A1A2A3, C2A1A2A3, C3A1A2A3, C1B1B2B3, C2B1B2B3, C3B1B2B3 (i.e., symmetry trials), A1B1B2B3, A2B1B2B3, A3B1B2B3, B1A1A2A3, B2A1A2A3, and B3A1A2A3 (i.e., equivalence trials).

Finally, in order to pass the test phase, a participant had to have at least 90% of the trials correct in each of the different test trial types. Furthermore, there were no programmed consequences in the testing phase. If a participant failed the test, he or she would return to another cycle of training and testing.

For the potential three 6-member equivalence classes, the acquisition phase consisted of eight blocks, as shown in Table 2. As with the potential three 3-member equivalence classes, one conditional relation was trained first before moving on to the next (i.e., first AF then BF, etc.). Again, after training the first two conditional relations (e.g., AF and BF), a mixed-trial block would randomly train all the conditional relations that had been trained thus far. Hence, the training of a new conditional relation (e.g., CF, DF, and EF) would always be followed by a mixed-trial block in which all trained relations were intermixed and presented five times each. The training trials were A1F1F2F3, A2F1F2F3, A3F1F2F3, B1F1F2F3, B2F1F2F3, B3F1F2F3, C1F1F2F3, C2F1F2F3, C3F1F2F3, D1F1F2F3, D2F1F2F3, D3F1F2F3, E1F1F2F3, E2F1F2F3, and E3F1F2F3.

In the maintenance phase (i.e., Blocks 9 to 12 in Table 2), the participants were introduced to four blocks with mixed training trials. Moreover, all trained conditional relations were presented five times each on a random basis. In addition, the probability of a programmed consequence was thinned, and the criterion to move to next block is as described above.

Following a training and maintenance phase, the participants were introduced to a testing block consisting of a mix of 75 baseline, 75 symmetry, and 300 equivalence trials. The lists of tested conditional relation trials are comprehensive; however, the pattern (see Table 2) is similar to the tested conditional relation trials presented for the three 3-member classes above. Furthermore, the criterion for finishing the experimental session or restarting a second training and test cycle was the same as above.

Trial types in the OTM training structure. The acquisition phase for the potential three 3-member equivalence classes consisted of two blocks in which the conditional relations AB and AC were established on a serialized basis, as shown in Table 3. Again, after training the first two conditional relations (e.g., AB and AC), a mixed-trial block would randomly train all the conditional relations. The training trials were as follows: A1B1B2B3, A2B1B2B3, A3B1B2B3, A1C1C2C3, A2C1C2C3, and A3C1C2C3. The number of trials presented and the mastery criterion were the same as for the MTO.

The maintenance phases for the potential three 3-member equivalence classes (i.e., training Blocks 3 to 6 in Table 3) and the potential three 6-member equivalence classes (i.e., Blocks 9 to 12 in Table 4) were introduced in the same manner as the maintenance phases in the MTO structures.

In the testing phase, the trial types were A1B1B2B3, A2B1B2B3, A3B1B2B3, A1C1C2C3, A2C1C2C3, A3C1C2C3 (i.e., directly trained trials), B1A1A2A3, B2A1A2A3, B3A1A2A3, C1A1A2A3, C2A1A2A3, C3A1A2A3 (i.e., symmetry trials), B1C1C2C3, B2C1C2C3, B3C1C2C3, C1B1B2B3, C2B1B2B3, and C3B1B2B3 (i.e., equivalence trials).

Furthermore, as depicted in Table 4, the potential three 6-member equivalence classes consisted of eight blocks. Again, following the training of the first two conditional relations (e.g., AB and AC), a mixed-trial block would randomly train all the conditional relations that had been trained thus far. The training trials were A1B1B2B3, A2B1B2B3, A3B1B2B3, A1C1C2C3, A2C1C2C3,

Table 3. *Number of Trials and Probability of Programmed Consequences in the training and testing blocks used to establish 3-Member Equivalence Classes with OTM structure*

Blocks	Probability of programmed consequences (%)	Minimum trials	Criterion
Training			
1. AB	100	15	14/15
2. AC	100	15	14/15
3. Mixed Trials AB, AC	100	30	27/30
4. Mixed Trials AB, AC	75	30	27/30
5. Mixed Trials AB, AC	25	30	27/30
6. Mixed Trials AB, AC	0	30	27/30
Testing			
Test block with *DT trials* (AB, AC),	0	30	27/30
SYM trials (BA, CA), and	0	30	27/30
EQ trials (BC, CB) randomly intermixed	0	30	27/30

Note. DT = Direct trained; SYM = symmetry; EQ = equivalence.

Table 4. *Number of Trials and Probability of Programmed Consequences in the training and testing blocks used to establish 6-Member Equivalence Classes with OTM structure*

Blocks	Probability of programmed consequences (%)	Minimum trials	Criterion
Training			
1. AB	100	15	14/15
2. AC	100	15	14/15
3. Mixed Trials AB, AC	100	30	27/30
4. AD	100	15	14/15
5. Mixed Trials AB, AC, AD	100	60	54/60
6. AE	100	15	14/15
7. Mixed Trials AB, AC, AD, AE	100	75	68/75
8. AF	100	15	14/15
9. Mixed Trials AB, AC, AD, AE, AF	100	90	81/90
10. Mixed Trials AB, AC, AD, AE, AF	75	75	68/75
11. Mixed Trials AB, AC, AD, AE, AF	25	75	68/75
12. Mixed Trials AB, AC, AD, AE, AF	0	75	68/75
Testing			
Test block with *DT trials* (AB, AC, AD, AE, AF),	0	75	68/75
SYM trials (BA, CA, DA, EA, FA), and	0	75	68/75
EQ trials (BC, CB, CD, DC, DE, ED, EF, FE, BD, DB, CE, EC, DF, FD, BE, EB, CF, FC, BF, FB) randomly intermixed	0	300	270/300

Note. Due to a programming error, the number of trials in block 5 increased with 30 trials instead of 15 trials. Thus, the extra number of trials is presented in blocks 5, 7, and 9. The 15 extra trials are not presented in blocks 10–12. DT = Direct trained; SYM = symmetry; EQ = equivalence.

A3C1C2C3, A1D1D2D3, A2D1D2D3, A3D12D3, A1E1E2E3, A2E1E2E3, A3E1E2E3, A1F1F2F3, A2F1F2F3, and A3F1F2F3.

In the potential three 6-member equivalence classes, the testing block consisted of a mix of 75 baseline, 75 symmetry, and 300 equivalence trials. The lists of tested conditional relation trials are comprehensive; however, the pattern (see Table 4) is similar to that in the tested conditional relation trials presented for the potential three 3-member equivalence classes above.

Trial types in the LS training structure. The procedures employed in the LS training structure (see Tables 5 and 6) mirror those of the MTO and OTM training structures. Thus, as described in Table 5, the acquisition phase for the potential three 3-member equivalence classes consisted of following trained conditional relations: A1B1B2B3, A2B1B2B3, A3B1B2B3, B1C1C2C3, B2C1C2C3, and B3C1C2C3. Furthermore, as shown in Table 6, and similar to the MTO and OTM training structures, the acquisition phase for the potential three 6-member equivalence classes in the LS structure consisted of following conditional relations: A1B1B2B3, A2B1B2B3, A3B1B2B3, B1C1C2C3, B2C1C2C3, B3C1C2C3, C1D1D2D3, C2D1D2D3, C3D12D3, D1E1E2E3, D2E1E2E3, D3E1E2E3, E1F1F2F3, E2F1F2F3, and E3F1F2F3. Again, as described above, a participant had to have 90% of the trials correct in order to move on to the next block.

From training to maintenance, this phase, for the potential three 3-member equivalence classes (see Table 5) as well as the potential 6-member equivalence classes (see Table 6), was introduced and carried out using the same protocol as with the MTO and OTM training structures.

Finally, in the testing phase, the conditional relations, tested for with three 3-member classes, were A1B1B2B3, A2B1B2B3, A3B1B2B3, B1C1C2C3, B2C1C2C3, B3C1C2C3 (i.e., baseline trials), B1A1A2A3, B2A1A2A3, B3A1A2A3, C1B1B2B3, C2B1B2B3, C3B1B2B3 (i.e., symmetry trials), A1C1C2C3, A2C1C2C3, A3C1C2C3, C1A1A2A3, C2A1A2A3, and C3A1A2A3 (i.e., equivalence trials). In the potential three 6-member equivalence classes, the

Table 5. *Number of Trials and Probability of Programmed Consequences in the training and testing blocks used to establish 3-Member Equivalence Classes with LS structure*

Blocks	Probability of programmed consequences (%)	Minimum trials	Criterion
Training			
1. AB	100	15	14/15
2. BC	100	15	14/15
3. Mixed Trials AB, BC	100	30	27/30
4. Mixed Trials AB, BC	75	30	27/30
5. Mixed Trials AB, BC	25	30	27/30
6. Mixed Trials AB, BC	0	30	27/30
Testing			
Test block with *DT trials* (AB, BC),	0	30	27/30
SYM trials (BA, CB), and	0	30	27/30
TRA/EQ trials (AC, CA) randomly intermixed	0	30	27/30

Note. DT = Direct trained; SYM = symmetry; TRA/EQ = transitivity/equivalence.

Table 6. *Number of Trials and Probability of Programmed Consequences in the training and testing blocks used to establish 6-Member Equivalence Classes with LS structure*

Blocks	Probability of programmed consequences (%)	Minimum trials	Criterion
Training			
1. AB	100	15	14/15
2. BC	100	15	14/15
3. Mixed Trials AB, BC	100	30	27/30
4. CD	100	15	14/15
5. Mixed Trials AB, BC, CD	100	60	54/60
6. DE	100	15	14/15
7. Mixed Trials AB, BC, CD, DE	100	75	68/75
8. EF	100	15	14/15
9. Mixed Trials AB, BC, CD, DE, EF	100	90	81/90
10. Mixed Trials AB, BC, CD, DE, EF	75	75	68/75
11. Mixed Trials AB, BC, CD, DE, EF	25	75	68/75
12. Mixed Trials AB, BC, CD, DE, EF	0	75	68/75
Testing			
Test block with *DT trials* (AB, BC, CD, DE, EF),	0	75	68/75
SYM trials (BA, CB, DC, ED, FE), and	0	75	68/75
TRA/EQ trials (AC, CA, BD, DB, CE, EC, DF, FD, AD, DA, BE, EB, CF, FC, AE, EA, BF, FB, AF, FA) randomly intermixed	0	300	270/300

Note. Due to a programming error, the number of trials in block 5 increased with 30 trials instead of 15 trials. Thus, the extra number of trials is presented in blocks 5, 7, and 9. The 15 extra trials are not presented in blocks 10–12. DT = Direct trained; SYM = symmetry; TRA/EQ = transitivity/equivalence.

testing block consisted of a mix of 75 baseline, 75 symmetry, 150 transitivity, and 150 equivalence trials. As with the preceding training structures, the pattern of the tested conditional relation trials (see Table 6) is similar to that of the tested conditional relation trials presented for the potential three 3-member equivalence classes above. As for passing criteria in the LS training structure, please refer to the passing criteria for the MTO and OTM structures.

Equivalence, transitivity, symmetry, and baseline trials. Regarding the potential three 3-member equivalence classes, to respond in accordance with equivalence or transitivity, participants had to score correctly on 27 out of 30 trials, or at least have 90% correct. On baseline and symmetry trials, a participant had to score 27 out of 30 trials correctly (i.e., a score of more than 90% correct) to pass these relations. In the potential three 6-member equivalence classes, to respond in accordance with equivalence and transitivity, either a score of 135 correct out of 150 or at least 90% correct was needed. In testing for baseline relations or responding in accordance with symmetry, a participant would have to have scored 68 out of 75 trials correctly (i.e., more than 90% correct) to pass these relations.

Reaction time. Reaction time was measured from when the comparison stimulus was presented to when the participant clicked on it. The median was calculated from the individual reaction times, and, based on that, the mean of the median reaction times was then calculated for each group. The reaction times were measured from the last five training trials, the first five test trials for baseline relations, the first five symmetry and equivalence trials, and the last five symmetry and equivalence trials.

Correct and incorrect responses. For each participant, the number of both correct and incorrect training trials was recorded. If a participant failed the first test, the number of additional training trials conducted, both correct and incorrect, was recorded as well.

Retention test. Between 2 and 4 weeks after the initial training and testing, participants were asked to participate in a follow-up test. The follow-up test was the same one that each participant had had a few weeks earlier, the only difference being the random mix-up of all the different test trials (i.e., for the potential three 3-member equivalence classes, a random mix-up of 30 baseline trials, 30 symmetry trials, and 30 equivalence trials for MTO and OTM, respectively, and for LS, a random mix-up of 30 baseline trials, 30 symmetry trials, 15 transitivity trials, and 15 equivalence trials; for the potential three 6-member equivalence classes, a random mix-up of 75 baseline trials, 75 symmetry trials, and 300 equivalence trials for MTO and OTM, respectively, and for LS, a random mix-up of 75 baseline trials, 75 symmetry trials, 150 transitivity trials, and 150 equivalence trials). Furthermore, each participant went through the follow-up test only one time. Passing or failing the follow-up test depended on the same criteria as those in effect during initial testing.

Results

Overall, results showed that participants from the OTM groups required fewer trials to pass the training phase. Also, 10 of 10 participants from the OTM, 8 of 10 participants from the MTO, and 2 of 10 participants from the LS groups responded in accordance with stimulus equivalence. Furthermore, participants responded faster to the last five baseline trials during training than they did to the first five directly trained trials during testing; responded faster to symmetry trials than equivalence trials; and on a follow-up retention test 9 participants (i.e., 3 from MTO and 6 from OTM) responded in accordance with stimulus equivalence.

Number of Training Trials

The training structures with the potential three 3-member equivalence classes are referred to as MTO-3x3, OTM-3x3, and LS-3x3, and the potential three 6-member equivalence classes are referred to as MTO-3x6, OTM-3x6, and LS-3x6.

On average, participants in the MTO-3x3 group needed 471 trials to pass the training phase, and participants in the MTO-3x6 group needed 759 (see Table 7). Moreover, compared to the MTO-3x3 group, the difference between the lowest and highest number of training trials was substantially lower for the MTO-3x6 group.

Participants from the OTM-3x3 group required fewer than half the number of responses that participants from the MTO-3x3 group did.

Table 7. *MTO Groups: Number of Trials and Errors for Each Participant and Scores on Baseline (DT), Symmetry (SYM), and Equivalence (EQ) Trials*

Participant	Group	Trials	Errors	Test #	DT	SYM	EQ
2602	MTO-3x3	255	49	1	**30**	**30**	**30**
2613	MTO-3x3	435	87	1	**30**	**30**	**29**
2601	MTO-3x3	600	208	1	**30**	**29**	**29**
2629	MTO-3x3	645	222	1	**28**	**29**	**30**
2628	MTO-3x3	420	101	1	**27**	**28**	21
2628*	MTO-3x3	120	1	2	**30**	**30**	**30**
2606	MTO-3x6	675	52	1	**75**	**75**	**298**
2630	MTO-3x6	705	59	1	**75**	**75**	**300**
2607	MTO-3x6	795	97	1	**75**	**74**	**299**
2612	MTO-3x6	840	77	1	**74**	**75**	**299**
2624	MTO-3x6	780	111	1	61	46	135
2624*	MTO-3x6	300	13	2	61	53	167

Note. The first column in each table shows the participant numbers, thereafter training condition, number of training trials, number of errors during training, test number one or two (i.e., if failing first test), and responding during testing for direct trained relations, symmetry, and equivalence. Maximum number of each trial type for the 3x3 Group was 30 (DT, SYM, and EQ), and 75 (DT and SYM) and 300 (EQ) for the 3x6 Group. Numbers in bold indicate that participants met the criteria for the condition. DT = Direct trained trials; SYM = symmetry trials; EQ = equivalence trials; OTM = one-to-many; MTO = many-to-one; LS = Linear Series; 3x3 = 3 classes with 3 members; 3x6 = 3 classes with 6 members.

** = Tested twice due to failing first test.*

Furthermore, they recorded the fewest number of errors in order to pass the training phase (see Table 8). In general, the OTM-3x3 group averaged 198 training trials, and participants in the OTM-3x6 group, compared to the MTO-3x6 group, required fewer training trials to pass the training phase—only 669.

The results for the LS3x3 group show that participants needed 297 trials, on average, to pass the training phase—a number that is somewhat in between the OTM-3x3 group (i.e., 198 trials) and the MTO-3x3 group (i.e., 471 trials) (see Table 9).

Table 8. *OTM Groups: Number of Trials and Errors for Each Participant and Scores on Baseline (DT), Symmetry (SYM), and Equivalence (EQ) Trials*

Participant	Group	Trials	Errors	Test #	DT	SYM	EQ
2621	OTM-3x3	195	19	1	**30**	**30**	**30**
2622	OTM-3x3	225	23	1	**30**	**30**	**30**
2609	OTM-3x3	195	11	1	**30**	**30**	**29**
2618	OTM-3x3	195	12	1	**30**	**30**	**29**
2605	OTM-3x3	180	9	1	**30**	**28**	**29**
2610	OTM-3x6	630	44	1	**75**	**75**	**300**
2627	OTM-3x6	645	31	1	**75**	**75**	**300**
2616	OTM-3x6	720	94	1	**75**	**75**	**297**
2623	OTM-3x6	660	50	1	**74**	**75**	**297**
2603	OTM-3x6	690	48	1	**73**	**74**	**291**

Note. The first column in each table shows the participant numbers, thereafter Group, number of training trials, number of errors during training, test number one or two (i.e., if failing first test), and responding during testing for direct trained relations, symmetry, and equivalence. Maximum number of each trial type for the 3x3 Group was 30 (DT, SYM, and EQ), and 75 (DT and SYM) and 300 (EQ) for the 3x6 Group. Numbers in bold indicate that participants met the criteria for the condition. DT = Direct trained trials; SYM = symmetry trials; EQ = equivalence trials; OTM = one-to-many; MTO = many-to-one; LS = Linear Series; 3x3 = 3 classes with 3 members; 3x6 = 3 classes with 6 members.

** = Tested twice due to failing first test.*

The last group, LS-3x6, averaged 735 training trials, ranging from 645 to 870 (see Table 9). Again, as was the case with the 3x3 conditions, participants from the LS-3x6 group required fewer trials to criterion than participants from the MTO-3x6 group (i.e., 759) and more trials to criterion than participants from the OTM-3x6 (i.e., 669).

Stimulus Equivalence Class Formation

In the MTO-3x3 group, all participants passed tests for directly trained relations as well as symmetry (see Table 7). However, four out of five—all but Participant 2628—responded in accordance with stimulus equivalence. Participant 2628 went through a second training and testing phase, and those results indi-

cated passing tests for directly trained relations as well as symmetry and stimulus equivalence. In the MTO-3x6 group, all participants but one, 2624, passed the test. Participant 2624 went through another training and testing phase but failed that time as well.

In both the OTM-3x3 and the OTM-3x6 group, all five participants passed with almost perfect scores on tests for directly trained relations, symmetry, and equivalence (see Table 8).

Table 9. *LS Groups: Number of Trials and Errors for Each Participant and Scores on Baseline (DT), Symmetry (SYM), and Transitivity/Equivalence (TRA/EQ) Trials*

Participant	Group	Trials	Errors	Test #	DT	SYM	TRA/EQ
2617	LS-3x3	210	26	1	**30**	**30**	**30**
2626	LS-3x3	345	57	1	**30**	**30**	**30**
2608	LS-3x3	375	93	1	**30**	**30**	10
2608*	LS-3x3	119	1	2	**30**	**30**	10
2619	LS-3x3	255	53	1	**30**	**30**	10
2619*	LS-3x3	119	0	2	**30**	**29**	10
2604	LS-3x3	300	52	1	**30**	**30**	20
2614	LS-3x6	870	98	1	**72**	66	190
2614*	LS-3x6	300	3	2	**75**	**75**	**280**
2615	LS-3x6	735	90	1	**71**	67	211
2615*	LS-3x6	300	9	2	**73**	**73**	**297**
2611	LS-3x6	750	73	1	49	45	95
2611*	LS-3x6	374	17	2	30	30	96
2620	LS-3x6	720	88	1	**72**	**71**	236
2625	LS-3x6	645	54	1	**70**	67	157

*Note. The first column in each table shows the participant numbers, thereafter Group, number of training trials, number of errors during training, test number one or two (i.e., if failing first test), and responding during testing for direct trained relations, symmetry, and transitivity/equivalence. Participants 2604, 2620, and 2625 preferred not continue with a new cycle of retraining and testing. Maximum number of each trial type for the 3x3 Group was 30 (DT, SYM, and TRA/EQ), and 75 (DT and SYM) and 300 (TRA/EQ) for the 3x6 Group. Numbers in bold indicate that participants met the criteria for the condition. DT = Direct trained trials; SYM = symmetry trials; TRA/EQ = transitivity/equivalence trials; LS = Linear Series; 3x3 = 3 classes with 3 members; 3x6 = 3 classes with 6 members. * = Tested twice due to failing first test.*

In the LS-3x3 group, only two of five Participants—2617 and 2626—passed the tests within the three different categories (see Table 9). Furthermore, three participants—2604, 2608, and 2619—passed tests of directly trained relations and symmetry; however, none of them responded in accordance with stimulus equivalence. Participants 2608 and 2619 were available for a second training and testing phase but again failed to respond in accordance with stimulus equivalence. Finally, in the LS-3x6 group, none of the five participants responded in accordance with stimulus equivalence. The test results for directly trained relations indicated that four out of five participants (2614, 2615, 2620, and 2625) passed, and only one of them, 2620, responded in accordance with symmetry. Moreover, three out of five group members—2611, 2614, and 2615—were available for a second training and testing phase. This time, two out of three (2614 and 2615) passed the test.

Reaction Times to Comparison Stimuli

Figure 2 shows the mean of the median reaction times to comparison stimuli on tests for the last five directly trained trials in the

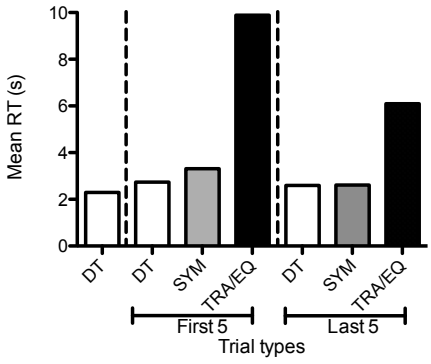

Figure 2. Mean of medians reaction times to comparison stimuli across all groups. The white bars indicate the mean of the median reaction times for the directly trained (DT) trials, the light grey bars indicate the mean median reaction times symmetry (SYM) trials, and the black bar indicate the mean median reaction times for the transitivity/equivalence (TRA/EQ) trials.

training; the first five directly trained, symmetry, and transitivity/equivalence trials; and the last five directly trained, symmetry, and equivalence trials in the test. Because there were no differences across groups, the figure shows reaction times for all participants. Figure 2 shows that there was a minimal increase from the directly trained trials in the training to test, a larger increase to symmetry trials, and a substantial increase to the transitivity/equivalence trials. For all trial types, there was a decrease from the first to the last part of the test. The relative difference was most pronounced for the equivalence trials.

Figure 3 shows reaction times to correct and incorrect comparison stimuli as a function of trial type for the LS-3x6 group. Furthermore, trial types are divided accord-

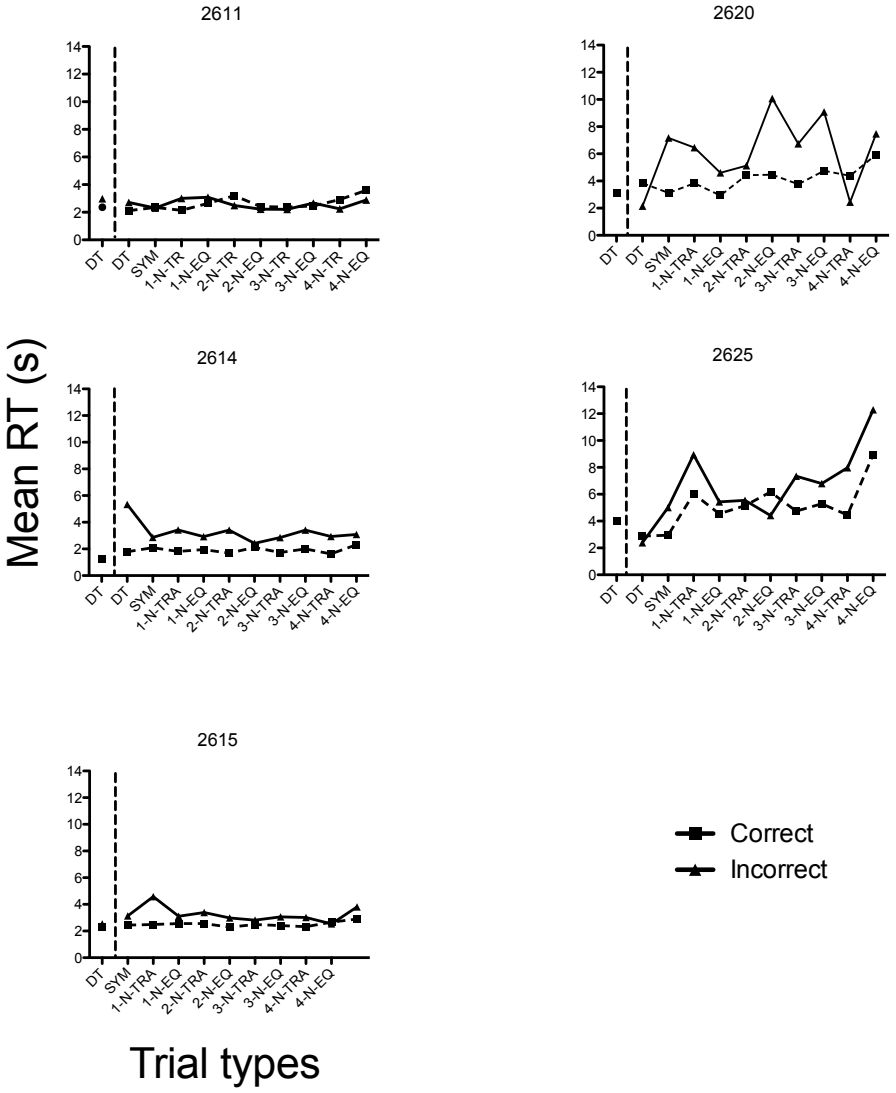

Figure 3. Reaction times to correct and incorrect comparison stimuli are shown as a function of direct trained (DT), symmetry (SYM), transitivity (TRA), and equivalence (EQ) trials. Transitivity and equivalence trials are divided according to nodal number of stimuli, that is 1-N is 1-node, 2-N is 2-node, etc. Furthermore, for direct training trials before testing three of the participants had no errors.

ing to number of nodes between sample and comparison stimuli. Three of the participants (2614, 2620, and 2625) had no incorrect scores in the conditional discrimination training. Participant 2611 notwithstanding, a clear pattern emerged in the reaction time to comparison stimuli. When incorrect, Participants 2614 and 2615 responded about 0.5 to 1 s more slowly to comparison stimuli, while Participants 2620 and 2625 responded between 0.5 and 4 s more slowly under the same conditions. For three of the participants (2611, 2614, and 2615), there was no increase in RT as a function of nodal number. However, for Participants 2620 and 2625 the pattern was somewhat different. For Participant 2620, there is an increase in reaction time for incorrect trials as function of increasing number of nodes. For correct comparisons, the RT pattern shows an increase for some of the nodes. For Participant 2625, there is an increase across number of nodes, most pronounced for the correct comparisons.

Table 10. *Follow-Up Test - MTO Groups: Scores on Baseline (DT), Symmetry (SYM), and Equivalence (EQ) Trials*

Participant	Group	Week	DT	SYM	EQ
2602	MTO-3x3	1	**30**	**30**	**30**
2628	MTO-3x3	3	24	25	**30**
2629	MTO-3x3	3	19	23	7
2613	MTO-3x3	4	20	20	10
2624	MTO-3x6	1	64	65	206
2607	MTO-3x6	2	**73**	**73**	**287**
2630	MTO-3x6	2	**73**	**73**	**282**
2606	MTO-3x6	4	65	65	218

Note. The first column in each table shows the participant numbers, thereafter week number for follow-up test, and the last three columns show responding during follow-up test for directly trained relations, symmetry, and equivalence. Participants 2601 and 2612 were not able to come back to laboratory for the follow-up tests. Maximum number of each trial type for the 3x3 Group was 30 (DT, SYM, and EQ), and 75 (DT and SYM) and 300 (EQ) for the 3x6 Group. Numbers in bold indicate that participants met the criteria for the condition. Rows with gray background indicate that the participant met the criteria for all three conditions on the follow-up test. DT = Direct trained trials; SYM = symmetry trials; EQ = equivalence trials; MTO = many-to-one; 3x3 = 3 classes with 3 members; 3x6 = 3 classes with 6 members.

Follow-Up Retention Test

Tables 10 through 12 show that a follow-up retention test was completed between 2 and 4 weeks after initial training and testing, except for Participants 2624 and 2625.

Table 11. Follow-Up Test - OTM Conditions: Scores on Baseline (DT), Symmetry (SYM), and Equivalence (EQ) Trials

Participant	Group	Week	DT	SYM	EQ
2605	OTM-3x3	2	**30**	**30**	**30**
2621	OTM-3x3	2	**29**	**30**	**28**
2609	OTM-3x3	2	**27**	25	**27**
2622	OTM-3x3	2	**27**	26	23
2618	OTM-3x3	4	**29**	**29**	23
2603	OTM-3x6	2	**74**	**74**	**300**
2627	OTM-3x6	2	**75**	**74**	**291**
2610	OTM-3x6	3	**74**	**75**	**298**
2623	OTM-3x6	3	51	46	181
2616	OTM-3x6	4	**72**	**75**	**293**

Note. The first column in each table shows the participant numbers, thereafter week number for follow-up test, and the last three columns show responding during follow-up test for directly trained relations, symmetry, and equivalence. Maximum number of each trial type for the 3x3 Group was 30 (DT, SYM, and EQ), and 75 (DT and SYM) and 300 (EQ) for the 3x6 Group. Numbers in bold indicate that participants met the criteria for that specific condition. Rows with gray background indicate that the participant met the criteria for all three conditions on the follow-up test. DT = Direct trained trials; SYM = symmetry trials; EQ = equivalence trials; OTM = one-to-many; 3x3 = 3 classes with 3 members; 3x6 = 3 classes with 6 members.

Out of 27 participants who took the follow-up test, 16 passed the initial stimulus equivalence test; only 9 of the 16 responded in accordance with equivalence class formation on the follow-up test. Categorizing the participants who responded in accordance with stimulus equivalence on the follow-up test reveals that 1 participant represented the MTO-3x3 group, 2 represented the MTO-3x6 group, another 2 represented the OTM-3x3 group, and the last 4 represented the OTM-3x6 group. It is noteworthy that on the follow-up test none of the participants from the LS groups responded in accordance with stimulus equivalence.

Table 12. Follow-Up Test - LS Groups: Scores on Baseline (DT), Symmetry (SYM), and Equivalence (EQ) Trials.

Participant	Group	Week	DT	SYM	TRA/EQ
2608	LS-3x3	2	**29**	**27**	10
2604	LS-3x3	3	26	22	21
2626	LS-3x3	3	14	15	10
2619	LS-3x3	4	21	21	26
2625	LS-3x6	1	57	61	133
2620	LS-3x6	2	68	55	198
2611	LS-3x6	2	54	48	93
2614	LS-3x6	2	57	59	179
2615	LS-3x6	2	58	53	175

Note. The first column in each table shows the participant number, thereafter week number for follow-up test, and the last three columns show responding during follow-up test for directly trained relations, symmetry, and transitivity/equivalence. Participant 2617 were not able to come back to laboratory for the follow-up tests. Maximum number of each trial type for the 3x3 Group was 30 (DT, SYM, and TRA/EQ), and 75 (DT and SYM) and 300 (TRA/EQ) for the 3x6 Group. Numbers in bold indicate that participants met the criteria for that specific condition. Rows with gray background indicate that the participant met the criteria for all three conditions on the follow-up test. DT = Direct trained trials; SYM = symmetry trials; TRA/EQ = transitivity/equivalence trials; LS = Linear Series; 3x3 = 3 classes with 3 members; 3x6 = 3 classes with 6 members.

Discussion

The purpose of the current study was to replicate and extend Arntzen et al.'s (2010) findings on differential probabilities of equivalence class formation when using three different training structures in the development of potentially 3-member and 6-member equivalence classes. Furthermore, in addition to comparing the number of trials and errors among the different training structures, the present study attempted to replicate earlier findings on reaction times to comparison stimuli. Finally, this study also looked at the differential effects of training structures and the formation of equivalence classes in a follow-up retention test 2 to 4 weeks after the initial training and testing.

The results replicated Arntzen et al.'s (2010) findings, whereby they showed that the difference between the OTM training structure and the MTO training structure is relatively small in training with three and six members. The formation of equivalence classes is lowest for the LS training structure, which is in accordance with previous studies (Arntzen et al., 2010; Arntzen & Holth, 1997, 2000). Furthermore, Arntzen and Vaidya (2008) argued that adults might generate better outcomes with the OTM structure due to a more complex behavioral repertoire that, in contrast to children, makes them better prepared to respond to simultaneous and successive discriminations. Regarding the LS training structure, it seems to be another variable that could influence the outcome, that is, the switching of all the nodal stimuli from being a comparison in one relation to a sample in another relation (see the discussions in Arntzen, 2011; Sidman, 2011).

The LS training structure generated the lowest outcomes on stimulus equivalence tests, and this was more pronounced in the

3x6 group—a result that is in accordance with recent findings (Arntzen et al., 2010; R. Saunders et al., 2005) as well as with R. R. Saunders and Green's (1999) discrimination analysis. In the LS training structure, again, all the simple discriminations necessary for repeated positive results on tests for symmetry are trained, but not all are trained for tests on either transitivity or equivalence. In accordance with the discrimination analysis, we see that in the 3x3 group, all five participants responded in accordance with symmetry (see Table 9). On the other hand, only two out of the five responded in accordance with transitivity and equivalence. In this regard, Saunders and Green offered an explanation for the notion that some participants responded in accordance with stimulus equivalence in the smaller class. First, although not presented explicitly during baseline training, some successive discriminations, which were required on the tests for transitivity and equivalence, were established. Referring to the component simple discrimination training, linear-series training does not require discrimination of comparison stimuli from different samples, as these are never presented together.

The component simple discrimination analysis notwithstanding, it could also be argued that poorer performances in the LS structure, when training more than six conditional discriminations, result from the fact that some nodal stimuli share a common learning history, while others do not. For example, as stimuli A and C have the nodal stimulus B in common, and stimuli D and F have the nodal stimulus E in common, we note that the stimuli B and E have no sample or comparison stimuli in common. Hence, considering the one-nodal training structures MTO and OTM, whether these can be compared to the LS structure when training stimulus classes with more than one node is arguable. For that reason, it is speculated that some form of mediated behavior is necessary for consistent positive results with the LS training structure when training more than three members.

The OTM training structure was the most efficient method for training baseline conditional relations, if one considers the number of training trials necessary to reach criterion among the 3x3 groups. Although not as pronounced in the 3x6 groups, the OTM structure still required fewer training trials than the other two. These findings are in accordance with the Arntzen et al. (2010) study and support R. R. Saunders and Green's (1999) discrimination analysis, suggesting that fewer training trials are needed with the OTM structure, as fewer component simple discriminations are required. Also in accordance with the Arntzen et al. study and Saunders and Green's discrimination analysis, participants that were trained with the three-member MTO structure required almost twice as many training trials compared to participants in the OTM group. Again, when using the MTO training structure, all component simple discriminations, necessary for consistent positive results on tests for symmetry, transitivity, and equivalence, are trained. Finally, and also in accordance with the predictions from the discrimination analysis, participants that were trained with the three-member LS training structure needed far fewer training trials than participants that were trained with the MTO structure, though still more trials than participants that were trained with the OTM structure.

In measuring reaction times to comparison stimuli, the present study replicated previous findings showing, in general, that reaction times (1) increased for the first five test trials for directly trained relations compared to the last five trials during training, (2) increased more quickly to the last five symmetry trials compared to the first five, and (3) increased more quickly to the last five equivalence trials compared to the first five. Moreover, reaction times to comparison stimuli increased substantially more for equivalence trials than for symmetry trials. These results are also in accordance with previous findings (e.g., Arntzen et al., 2010; Arntzen & Holth, 1997, 2000;

Arntzen & Lian, 2010; Eilifsen & Arntzen, 2009; Holth & Arntzen, 1998, 2000; R. R. Saunders et al., 2005; Spencer & Chase, 1996). Some of the reason for the slightly increased reaction times for the directly trained trials could be related to the fact that in the test, trials were intermixed so the participants were exposed to novel trials.

Furthermore, the current study replicated previous studies (Arntzen et al., 2010; Arntzen & Holth, 1997, 2000) that found that the relative differences in reaction times to comparison stimuli between the last five baseline trials and the first five equivalence trials were higher in the MTO training structure compared to the OTM training structure with three-member classes. As a possible explanation, Arntzen et al. (2010) compared the differences in sample stimulus and comparison stimulus relations between the MTO and OTM training structures. They reasoned that during OTM training, one sample stimulus is related to more than one comparison stimulus, whereas in MTO training, one comparison stimulus is related to more than one sample stimulus. Hence, with the MTO structure a participant could learn to predict the comparison stimulus, depending on the choice of sample stimulus, and this sort of problem-solving behavior could therefore explain the increase in reaction times to comparison stimuli when testing for equivalence relations.

In contrast to other studies (Bentall, Jones, & Dickins, 1999; Fields, Adams, Newman, & Verhave, 1990), the present study did not find that reaction times to comparison stimuli in the LS training structure increased as a function of an increase in nodal number for all 5 participants in the LS-3X6 group. Also noteworthy, the present study found that reaction times to comparison stimuli were longer for incorrect responses compared to correct responses for 4 out of 5 participants. Thus, it could be that participants responded more slowly when conditional relations were weakly established, and that they did not respond randomly and more quickly. Moreover, the study did not show an increase in reaction times to comparison stimuli as a function of number of nodes between sample stimulus and comparison stimulus for all participants. Hence, 3 out of 5 participants did not show any increase in reaction time, which does not support Spencer and Chase's (1996) argument about an inverse relationship between response speed and nodal number.

On the other hand, the findings of the current study support the results of other research studies (e.g., Imam, 2006) and suggest that a more equalized presentation of the trained conditional relations would help reduce the nodal number effects. An even number of presentations of the trained conditional relations, however, brings about overtraining (e.g., Driskell, Willis, & Copper, 1992; Minister, Elliffe, & Muthukumaraswamy, 2011) and raises questions, therefore, about whether such an approach camouflages the real purpose of the study—identifying the differences in stimulus equivalence outcome as a function of the component simple discriminations that are embedded within the three different training structures. Nonetheless, by attempting to equalize the number of training trials per conditional relation in the current study, in an effort to replicate the Imam (2006) study, could it have been the case that we actually reduced the nodal number effects due to overtraining? This does not seem to have been the case, as not all participants in the LS-3x6 group passed the test for directly trained relations (results that indicate that the phenomenon of overtraining should not have occurred).

Finally, the follow-up retention results showed that 9 of the 16 participants who responded in accordance with equivalence class formation on the initial test also responded in accordance with equivalence class formation 2 to 4 weeks later. Interestingly, 6 of 9 participants who responded in accordance with stimulus equivalence were from the OTM groups, and, once again,

this result supports the idea that it is easier to remember conditional relations and emergent relations when fewer component simple discriminations are required during training (R. R. Saunders & Green, 1999).

In contrast to the Arntzen et al. (2010) study, the current experiment was carried out using a between-subjects design, and thus individual learning histories could have obscured the results. Future studies should either attempt to replicate our findings using a single-subject design or increase the number of participants per experimental group. Furthermore, with respect to the differences in the LS training structure as compared to the MTO and OTM structures, future studies should compare the formation of equivalence classes to populations with different behavioral repertoires in order to learn more about the effects that mediating behaviors have on equivalence class formation when stimulus classes contain more than three members.

In sum, the present study explored the effects of differences in training structures on the formation of equivalence classes, and we found that the OTM training structure generated higher outcomes on tests for emergent relations in three potential classes of three- as well as six stimuli with a typically adult population. Moreover, the LS training structure generated the lowest outcomes on tests for emergent relations, and this was more pronounced in the potential six-member equivalence class. Furthermore, reaction times to comparison stimuli replicated previous findings showing an increase in reaction time from the last five training trials to the first five test trials, longer reaction times in the symmetry trials compared to directly trained trials, and even longer reaction time in the equivalence trials compared to symmetry trials. Finally, participants trained with the OTM structure performed better in two out of three tests for equivalence class formation 2 to 4 weeks after initial training and testing.

References

Arntzen, E. (2011). Comments on Sidman's Remarks. *European Journal of Behavior Analysis, 12*, 351–353. Retrieved from http://www.ejoba.org/

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. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47*, 309–320. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Arntzen, E., & Holth, P. (2000). Differential probabilities of equivalence outcome in individual subjects as a function of training structure. *The Psychological Record, 50*, 603–628. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Arntzen, E., & Lian, T. (2010). Trained and derived relations with pictures as nodes. *The Psychological Record, 60*, 659–677. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Arntzen, E., & Nikolaisen, S. L. (2011). Establishing equivalence classes in children using familiar and abstract stimuli and many-to-one and one-to-many training structures. *European Journal of Behavior Analysis, 12*, 105–120. Retrieved from http://www.ejoba.org/

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. Retrieved from http://www.eahb.org/NewSitePages/Bulletin-Homepage.htm

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. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Driskell, J. E., Willis, R. P., & Copper, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology, 77*, 615-622. doi: 10.1037//0021-9010.77.5.615

Dymond, S., & Rehfeldt, R. A. (2001). Supplemental measures and derived stimulus relations. *Experimental Analysis of Human Behavior Bulletin, 19*, 8–12. Retrieved from http://www.eahb.org/NewSitePages/BulletinHomepage.htm

Eilifsen, C., & Arntzen, E. (2009). On the role of trial types in tests for stimulus equivalence. *European Journal of Behavior Analysis, 10*, 187–202. Retrieved from http://www.ejoba.org/

Fields, L., Adams, B. J., Newman, S., & Verhave, T. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53*, 345–358. doi: 10.1901/jeab.1990.53-345

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. Retrieved from http://thepsychologicalrecord.siuc.edu/

Fields, L., Landon-Jimenez, D. V., Buffington, D. M., & Adams, B. J. (1995). Maintained nodal-distance effects in equivalence classes. *Journal of the Experimental Analysis of Behavior, 64*, 129–145. doi: 10.1901/jeab.1995.64-129

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. Journal of the Experimental Analysis of Behavior, 48, 317–332. doi: 10.1901/jeab.1987.48-317

Holth, P., & Arntzen, E. (1998). Stimulus familiarity and the delayed emergence of stimulus equivalence or consistent nonequivalence. *The Psychological Record, 48*, 81–110. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

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. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Hove, O. (2003). Differential probability of equivalence class formation following a one-to-many versus a many-to-one training structure. *The Psychological Record, 53*, 617–634. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

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

Minister, T., Elliffe, D., & Muthukumaraswamy, S. D. (2011). Emergent stimulus relations depend on stimulus correlations and not on reinforcement contingencies. *Journal of the Experimental Analysis of Behavior, 95*, 327–342. doi: 10.1901/jeab.2011.95-327

Pilgrim, C., & Galizio, M. (2000). Stimulus equivalence and unit of analysis. In J. C. Leslie & D. E. Blackman (Eds.), *Experimental and applied analysis of human behavior* (pp. 111–126). Reno, NV: Context Press.

Rehfeldt, R. A., & Hayes, L. J. (2000). The long-term retention of generalized equivalence. *The Psychological Record, 50*, 405–428. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

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. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Saunders, R. R., Chaney, L., & Marquis, J. G. (2005). Equivalence class establish-

ment with two-, three-, and four-choice matching to sample by senior citizens. *The Psychological Record, 55*, 539–559. doi: http://thepsychologicalrecord.siuc.edu/

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. doi: 10.1901/jeab.1999.71-195

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., & McEntee, J. E. (2004). Increasing the probability of stimulus equivalence with adults with mild mental retardation. *The Psychological Record, 54*, 423–435. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

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

Sidman, M. (2011). Reply to commentaries on "Remarks" columns. *European Journal of Behavior Analysis, 12*, 355–370. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

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

Spencer, T. J., & Chase, P. N. (1996). Speed analysis of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65*, 643–659. doi: 10.1901/jeab.1996.65-643