Effects of Meaningful Stimuli Contained in Different Numbers of Classes on Equivalence Class Formation

Justice Mensah[1] · **Erik Arntzen**[1]

Abstract Previous experiments have investigated the function of using pictures or meaningful stimuli on equivalence class formation. For example, when attempting to form three 5-member classes (A→B→C→D→E), findings have shown that pictures used as C stimuli have increased the probability of producing equivalence class formation relative to when all stimuli in the stimulus set are abstract. The present experiment extends the literature by examining whether the formation of equivalence classes varies as a function of having three (C1, C2, and C3), two (C1 and C2), or one (C1) stimulus as a picture in a set of abstract stimuli. Hence, 60 participants were randomly assigned to 4 different experimental groups: 0-picture group or abstract group (ABS), 1-picture group (1PIC), 2-pictures group (2PIC), and 3-pictures group (3PIC). In addition, we had a reference group with abstract shapes only. The findings from the present experiment showed that 2 of 15 participants in the ABS formed classes. Also, two of 15 participants in the 1PIC formed classes, 8 of 15 participants in the 2PIC formed equivalence classes, and 12 of 15 members in the 3PIC formed classes. The statistical analysis supported the notion that equivalence class formation is a function of the number of pictures in a potential equivalence class.

Keywords Stimulus equivalence · Abstract stimuli · Meaningful stimuli · University students

✉ Erik Arntzen
erik.arntzen@equivalence.net

[1] Department for Behavioral Science, Oslo and Akershus University College, P.O. Box 4, St. Olav's Plass, 0130 Oslo, Norway

Stimulus equivalence is defined as responding to relations characterized by reflexivity, symmetry, and transitivity. Stimuli that evoke the same behavioral responses without training are said to be members of an equivalence class (Sidman & Tailby, 1982). The stimuli that eventually become members of an equivalence class can be meaningless or meaningful, as well as vary in degree of meaningfulness (Arntzen, Nartey, & Fields, 2015). In this experiment, we used the linear series (LS) training structure to explore the effect of familiar pictorial stimuli on equivalence class formation.

Previous experiments have found that the inclusion of one class member as a meaningful stimulus in a class of abstract stimuli increases the probability of responding in accordance with equivalence (e.g., Arntzen, 2004; Arntzen & Lian, 2010; Arntzen, Nartey, & Fields, 2014; Fields, Arntzen, Nartey, & Eilifsen, 2012; Holth & Arntzen, 1998; Nartey, Arntzen, & Fields, 2014). Some of the studies have used the LS to form three 5-member classes (A→B→C→D→E) and having the C stimuli as familiar pictorial stimuli. For example, Fields et al. (2012) established three 3-node 5-member equivalence classes in adult participants. The experiment was arranged as LS training structure (A→B→C→D→E). The main findings were that no participant formed classes when all the A–E stimuli were abstract shapes. However, eight of 10 participants formed classes when the A, B, D, and E stimuli were abstract, and the C stimuli were meaningful pictures. This finding has been replicated in a series of studies (Arntzen et al., 2014; Arntzen, Nartey, et al., 2015; Nartey et al., 2014; Nartey, Arntzen, & Fields, 2015a, b; Nedelcu, Fields, & Arntzen, 2015; Travis, Fields, & Arntzen, 2014). One study focused on the effect of structural location of meaningful stimuli on equivalence class formation (Nartey et al., 2015b). The results from the experiment showed that 60% of participants formed classes when the A stimulus was meaningful and the B to E stimuli were abstract, 40% of participants formed classes

when the B stimulus was meaningful and the A, C, D, and E stimuli were abstract, and 70% of participants formed classes when the C stimulus was meaningful and the A, B, D, and E stimuli were abstract. Also, 40% of participants formed classes when the D stimulus was meaningful and the A, B, C, and E stimuli were abstract, and 20% of participants formed classes when the E stimulus was meaningful and the A to D stimuli were abstract. A meaningful stimulus serving as the middle node results in higher yields relative to other locations in the class structure. Thus, would similar effects be found with three, two, or one classes with C stimuli as pictorial stimuli?

Percentage of correct responding has been one of the most important measurements within the research on emergent relations. However, Dymond and Rehfeldt (2001) have suggested some additional measures in research on emergent relations, like sorting or categorization of stimuli and reaction time to comparison stimuli. Lately, a number of studies within research on emergent relations have included sorting tests (Arntzen, 2004; Arntzen, Granmo, & Fields, 2016; Arntzen, Norbom, & Fields, 2015; Cowley, Green, & Braunling-McMorrow, 1992; Dickins, 2011, 2015; Dymond & Rehfeldt, 2001; Eilifsen & Arntzen, 2009, 2011; Fields, Arntzen, & Moksness, 2014; Fields et al., 2012; Fienup & Dixon, 2006; Green, 1990; Hove, 2003; Lowe, Horne, Harris, & Randle, 2002; Mackay, Wilkinson, Farrell, & Serna, 2011; Pilgrim & Galizio, 1996; Sigurdardottir, Mackay, & Green, 2012; Smeets, Dymond, & Barnes-Holmes, 2000). In general, studies have found that sorting tests have been a quick and easy measurement to administer to assess class partitioning. Furthermore, that all the participants who responded in accordance with equivalence in the matching-to-sample (MTS) based test sorted the stimuli in accordance with the experimenter-defined classes; however, some of the participants who sorted correctly did not respond in accordance with stimulus equivalence on the MTS test (e.g., Arntzen et al., 2016; Arntzen, Norbom, et al., 2015). The sorting test in this experiment is used to ascertain if the sorting of stimuli after the MTS test gave a different pattern depending on the number of classes with pictorial C stimuli.

Studies on reaction time or speed (inversed reaction time) have showed faster responding to baseline trials compared to test trials for emergent relations (e.g., Arntzen & Hansen, 2011; Bentall, Dickins, & Fox, 1993; Spencer & Chase, 1996; Wulfert & Hayes, 1988) and also slower responding to equivalence trials compared symmetry trials (Arntzen & Hansen, 2011; Wulfert & Hayes, 1988). Previous studies have shown that speed is faster for correct responses than incorrect responses (Arntzen, Nartey, et al., 2015) and also for participants forming equivalence classes and those who were failing on the test (Arntzen, Braaten, Lian, & Eilifsen, 2011).

Based on the findings of previous studies that have included familiar pictures as C stimuli when trying to establish three 5-member classes (A→B→C→D→E), our experiment extends the literature by examining the effect of the number of C stimuli as familiar pictures. Thus, we ask if the formation of equivalence classes will vary as a function of having three, two or one classes with C stimuli as pictorial stimuli. Furthermore, our experiment asks if stimuli sorting after MTS tests will indicate consistent patterns with the results of the MTS test of emergent relations on the number of classes with C stimuli as pictorial stimuli. We used a modified simultaneous training and testing protocol in which the baseline trials were introduced on a serialized basis with mix blocks at the end of training and during testing. The LS training structure was used in this experiment, in which participants attempted to form three 5-member equivalence classes by training AB, BC, CD, and DE relations. Participants were assigned to four experimental groups: ABS, 1PIC, 2PIC, and 3PIC.

Method

Participants

Sixty university students (35 males and 25 females) voluntarily participated in this experiment. The participants were between age 18 and 24 years ($M = 21.33$, $SD = 1.46$). None of the participants had any prior knowledge of stimulus equivalence. An informed consent form was given to each participant upon entering the experimental room. The informed consent form entailed details of the experiment, the rights of the participants during the experiment, and the approximated duration of the experiment (90 minutes). They were also informed that they could quit at any time without any negative consequences for them. The participants were fully debriefed after they had completed the experimental session.

Design

A between-groups experimental design was used for this experiment. The 60 participants were randomly and equally assigned to one of the following four (4) experimental groups: (1) C1-as-PIC (1PIC), (2) C1 and C2-as-PIC (2PIC), (3) C-as-PIC (3PIC), and (4) Abstract (ABS).

Apparatus and Setting

The experiment was conducted in a lab room at the University of Ghana, Legon. The lab room measured approximately 7 m² and furnished with tables and chairs. The experiment was conducted on a 17-in. screen computer laptop. The computer laptop used 1828 MHz Intel Centrino® processor. The computer laptop had a mouse that was used by participants to click on the stimuli displayed on the laptop. The training and testing

sessions for conditional discriminations for all the participants were conducted with a customized matching-to-sample (MTS) software. The software controlled how stimuli were presented, recorded the response data of participants including the trial number, the number of training trials, whether participants chose the correct responses or not, as well as the consequences of the responses made by participants. The software also recorded a summary of directly trained trials or baseline trials, symmetry trials, transitivity trials, and equivalence trials as well as the exact duration of the experiment.

Stimuli

The experiment used abstract and meaningful stimuli as members of equivalence classes. The top section of Fig. 1 displays the 15 abstract stimuli used while the bottom section shows three meaningful pictorial stimuli used to replace the C abstract stimuli. For the abstract group all stimuli, A–E, were abstract shapes. The 1PIC group had the C1 member as the picture stimulus, while the A, B, D, and E stimuli, as well as the C2 and C3, were abstract stimuli. The 2PIC group had the C1 and C2 as pictures, while the A, B, D, and E stimuli, as well as the C3, were abstract stimuli. The 3PIC group had all C stimuli set as pictures, while the A, B, D, and E stimuli were abstract stimuli. The abstract stimuli were displayed in black, and the meaningful stimuli were displayed in colors. All stimuli were displayed on a white background. The size of each stimulus displayed on the computer monitor was 9.4 cm × 3.4 cm.

Procedure

Card Sorting Participants were given 15 plastic-laminated cards that corresponded to the stimuli to be used in their respective assigned groups. They were told to "put them into groups." This was done to ensure that participants did not partition the stimuli into experimenter-defined classes prior to the experiment. Participants were given the same cards after the experiment with the same instruction. The purpose of post-sorting class formation was to provide a quick evaluation of equivalence class formation after the MTS training and testing of emergent relations.

Instruction After the pre-class formation sorting task, participants remained seated in the experimental cubicle behind the computer and were presented with the following instruction on the computer screen:

In a moment, a stimulus will appear in the middle of the screen. Click on this by using the computer mouse. Three stimuli will then appear in the three corners of the screen. Choose one of them by clicking on it with the mouse. If you choose the stimulus we have defined as correct, words like "very good," "excellent," and so on will appear on the screen. If you press a wrong stimulus, the word "wrong" will appear on the screen. At the bottom of the screen, the number of correct responses you have made will be counted. 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 so far, you can get all of the tasks

Fig. 1 Stimuli used as members of the equivalence classes. *Top section* shows the 15 abstract stimuli; *bottom section* shows the meaningful stimuli that replaced C abstract stimuli in different groups

correct. Please do your best to get everything right. Thank you and good luck!

No further instructions were given after the experiment started.

Trial Structure and Contingencies Each trial began with the presentation of the sample stimulus at the middle of the computer screen. Responding to the stimulus by clicking on it with the computer mouse is followed by the presentation of three comparison stimuli displayed at three of the four corners of the computer screen, while the sample stimulus was still displayed at the middle of the computer screen. The location of the comparison stimuli throughout the trials was randomized.

A comparison was chosen or selected by clicking on it with the computer mouse. Upon clicking one on of the comparisons, both sample stimulus and comparison stimuli disappear from the screen. A click on the correct comparison was followed by the presentation of the words *correct, very good, super,* or *excellent* on the screen. A click on one of the incorrect comparisons was followed by the presentation of the word *wrong* on the screen. The programmed consequence was displayed in the middle of the computer screen for 1,000 ms and followed with a 500-ms intertrial interval. Between the trials, the mouse cursor was reset to the middle of the screen.

Acquisition of Baseline Relations Baseline relations were presented sequentially. All baseline relations were trained in seven phases, with programmed consequences for correct and incorrect comparison choices (see Table 1 for a full overview of each of the experimental phases). Phase 1 was for the training of AB relations in a block containing nine trials; three each of the trial types. A mastery criterion of 100% correct comparison selection was required for the training of each relation. Phase 2 was the same as Phase 1 except that BC relations were trained. Phase 3 was for the training of ABC relations in a block containing 18 trials. A mastery criterion of at least 90% correct comparison selection was required for the training of each relation. Phase 4 was the same as Phase 1, except that CD relations were trained. Phase 5 was for the training of ABCD relations in a block containing 36 trials. A mastery criterion of at least 90% correct comparison selection was required for the training of each relation. Phase 6 was the same as Phase 1, except that DE relations were trained. Participants repeated block trials for each phase until the criterion was met before proceeding to the next phase. An equalization feature was used to ensure that each of the baseline relations during training were presented the same number of times.

Phase 7 was a mixed training block of all relations trained AB/BC/CD/DE. The trials presented in this block were A1/B1B2B3, A2/B1B2B3, A3/B1B2B3, B1/C1C2C3, B2/C1C2C3, B3/C1C2C3, C1/D1D2D3, C2/D1D2D3, C3/ D1D2D3, D1/E1E2E3, D2/E1E2E3, D3/E1E2E3. For each trial representation, the first stimulus is the sample, and the other three are the comparison stimuli. The underlined comparison stimulus is the correct comparison. Phase 7 contained 45 trials, and the block was repeated until a mastery criterion of at least 90% correct was met.

Maintenance of Baseline Trials Participants continued with training blocks but with reduced programmed consequences after acquisition. The probability of programmed consequences after the last acquisition of baseline trials was reduced to 75%, 50%, 25%, and 0%, depending on at least 90% correct responding for each block. The presentation of trials was randomized in each block. The maintenance phase was completed with the mastery level of responding in the last block of 36 baseline trials with no programmed consequences.

Testing for Emergent Relations The last block with no programmed consequences was followed by a test block for emergent relations that contained 180 trials. Of the 180 trials, there were 36 baseline trials, 36 symmetry trials, 54 transitivity, and 54 equivalence trials. The 180 trials in the last block were presented without any programmed consequences. The formation of equivalence class was defined by the selection of at least 90% correct comparison consistent with the experimenter-defined classes in the test block. Failure to form classes was defined by the selection of class-indicative comparisons of less than 90% of the trials in the test block.

Results

Acquisition and Maintenance of Baseline Relations

The median number of trials to acquire baseline relations were 486, 513, 495, and 441 for the ABS, 1PIC, 2PIC, and 3PIC groups, respectively. An H test showed no significant difference for the speed of baseline acquisition across groups ($H = 2.60$, $p = .46$). However, the 24 participants who formed classes required a median of 450 trials to acquire baseline relations relative to a median of 517.5 trials for the 36 participants who did not form classes (see Fig. 2). A U test showed a significant difference between participants who responded in accordance to equivalence and those who did not respond in accordance to equivalence ($U = 2.02$, $p = .04$).

Formation of Equivalence Classes

Across the groups, 24 participants responded in accordance with equivalence, and 36 participants failed to form equivalence classes (as can be seen in Fig. 3). Twelve of 15 participants responded in accordance with equivalence in the 3PIC group, eight of 15 participants responded in accordance with

Table 1 Sequence of training and testing

| Experimental phases	Trial types	% program consequences	Number of trials
Acquisition of baseline relations (all trial types presented randomly)			
1. Serialized trials	A1B1, A2B2, A3B3	100	9
2. Serialized trials	B1C1, B2C2, B3C3	100	9
3. Serialized trials	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3	100	18
4. Serialized trials	C1D1, C2D2, C3D3	100	9
5. Serialized trials	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3	100	36
6. Serialized trials	D1E1, D2E2, D3E3	100	9
7. Serialized trials	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	100	45
8. Mixed trials (trials presented randomly)	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	75	36
9. Mixed trials (trials presented randomly)	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	50	36
10. Mixed trials (trials presented randomly)	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	25	36
11. Mixed trials (trials presented randomly)	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	0	36
Test for emerged relations (trials presented randomly intermixed)			
Baseline trials	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3, C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	0	36
Symmetry trials	B1A1, B2A2, B3A3, C1B1, C2B2, C3B3, D1C1, D2C2, D3C3, E1D1, E2D2, E3D3	0	36
Transitivity trials	A1C1, A2C2, A3C3, A1D1, B1D1, A2D2, B2D2, A3D3, B3D3, A1E1, B1E1, C1E1, A2E2, B2E2, C2E2, A3E3, B3E3, C3E3	0	54
Equivalence trials	C1A1, D1A1, E1A1, C2A2, D2A2, E2A2, C3A3, D3A3, E3A3, D1B1, E1B1, D2B2, E2B2, D3B3, E3B3, E1C1, E2C2, E3C3	0	54

Note. For all the training phases, there are three trials of each relation of the three classes except for Phases 5 and 7, which have nine more additional CD and EF relations trials, respectively, due to a programming error

equivalence in the 2PIC group, two of 15 participants responded in accordance with equivalence in the 1PIC group, and two of 15 participants responded in accordance with equivalence in the ABS group.

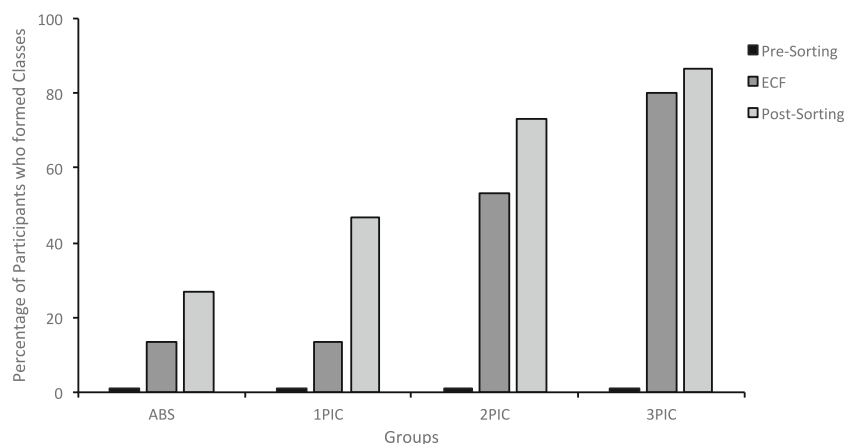A significant difference in terms of responding in accordance with equivalence was found between the four experimental groups, $F(3, 56) = 8.66$, $p = .00$. A post hoc analysis using the Least Significant Difference was used to test for the significant differences among the four experimental groups. The analysis showed a significant difference exists between no picture and all the picture groups, 1PIC ($p = -.18$), 2PIC ($p = -.27$), and 3PIC ($p = -.35$). Furthermore,

Fig. 2 Median number of trials needed to acquire baseline relations for participants who formed classes and those who did not form classes regardless of experimental condition. Error bars show standard error of the mean
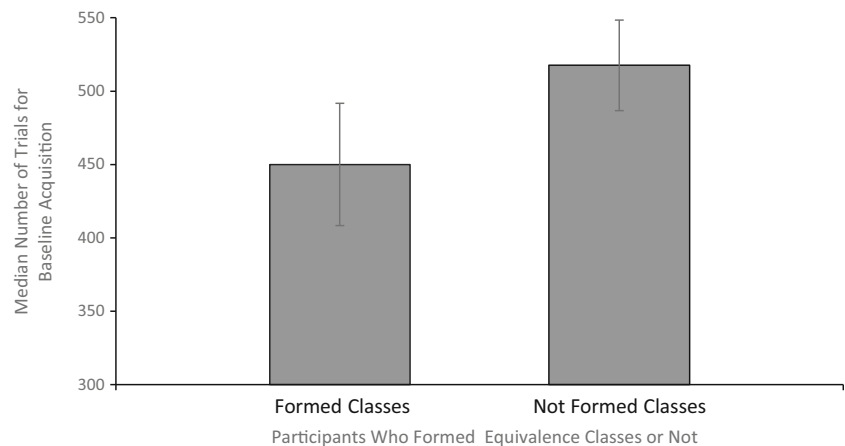
Fig. 3 Equivalence classes in pre-class formation sorting tests (Pre-Sorting), derived relations test and post-class formation sorting tests (Post-Sorting). ECF = equivalence class formation; ABS = abstract; 1PIC = C1 as picture; 2PIC = C1 and C2 as pictures; 3PIC = C as pictures

a significant difference exists between 1PIC and 3PIC ($p = -.17$). There was no significant difference between 1PIC and 2PIC ($p = -.87$) as well as 2PIC and 3PIC ($p = -.08$).

Card Sorting

Figure 3 further shows how each group responded in accordance with equivalence in terms of sorting the cards prior to and after the exposure to the MTS training and testing. Each cluster of bars in the panel in Fig. 3 represents data in percentages for a particular experimental group. The leftmost bars in each cluster shows pre-class formation sorting, the middle bars in each cluster shows performance on the derived relations test, and the rightmost bars in each cluster shows post-class formation sorting. Hence, 0%, 13.3%, and 26.7% of the total participants in ABS responded according to equivalence in pre-sorting class formation test, derived relations test, and post-sorting class formation test, respectively. Also, 0%, 13.3%, and 46.7% of the participants in 1PIC responded in accordance with equivalence in pre-sorting class formation test, derived relations test, and post-sorting class formation test, respectively. In addition, 0%, 53.3%, and 73.3% of the participants in 2PIC responded in accordance with equivalence in pre-sorting class formation test, derived relations test, and post-sorting class formation test, respectively. Finally, 0%, 80%, and 86.7% of the participants in 3PIC responded in accordance with equivalence in pre-sorting class formation test, derived relations test, and post-sorting class formation test, respectively.

The pre- and post-class formation sorting tests of all the participants are displayed in Tables 2 and 3. If the participants sorted the cards in accordance with the experimenter-defined classes (three 5-member classes), such sorting is indicated by clusters designated as 500, 050 or 005 strings. Hence, the first cluster with five Class-1 stimuli, the second cluster with five Class-2 stimuli, and the third cluster with five Class-3 stimuli. Hence, before class formation training, no participant sorted the stimuli into the experimenter-defined classes. For example, P1034 sorted the cards in three clusters. The first cluster with five Class-1 stimuli, the second cluster with five

two Class-1 stimuli and one Class-2 stimulus. The second cluster with one Class-1 stimulus, one Class-2 stimulus and two Class-3 stimuli. The third cluster with two Class-1 stimuli, three Class-2 stimuli, and three Class-3 stimuli. After the MTS training and test, 35 of 60 participants sorted the cards according to the experimenter-defined classes. All the participants who formed equivalence classes sorted the cards to experimenter-defined classes. Eleven participants who did not form classes also sorted the cards according to experimenter-defined classes. Furthermore, 13 of 15 members in the 1PIC sorted the stimuli set for Class 1 to the experimenter-defined stimuli set for Class 1.

Error Analysis

For the 13 participants who failed to form classes in 1PIC, there were fewer errors for Class 1 (class with C stimulus as a picture) than for Class 2 and Class 3 (classes without C stimuli as pictures). Specifically, an average of 10.92 errors was made responding in Class 1 relative to an average of 19.46 and 17.62 for Class 2 and Class 3, respectively (see Fig. 4, upper panel). Furthermore, for the seven participants who failed to form classes in 2PIC, there were more errors in Class 3 (class without C stimulus as a picture) than for Class 1 and Class 2 (classes with C stimuli as pictures). Specifically, an average of 21 errors was made responding in Class 3 relative to an average of 14.14 and 17.14 for Class 1 and Class 2, respectively (see Fig. 4, lower panel).

Different Types of Response Patterns

An analysis was conducted to examine the response patterns of participants who did not form equivalence classes. The purpose of this analysis was to find out whether participants who failed to form classes responded in accordance with the participant-defined classes or showed an indeterminate pattern of responding. Experimenter-defined classes include relations occurring four or five times in a test block. Participant-defined

Table 2 Pre-class formation sorting tests and post-class formation sorting tests outcomes for ABS and 1PIC

Group	Participant	Pre Cls-1/2/3	Pre Cls-1/2/3	Pre Cls-1/2/3	Pre Cls-1/2/3	Pre Cls-1/2/3	Pre Cls-1/2/3	Post Cls-1/2/3	Post Cls-1/2/3	Post Cls-1/2/3	Post Cls-1/2/3	Post Cls-1/2/3	ECF	Sort
ABS	1001	1 0 2	2 0 1	1 1 1	1 2 0	0 2 1		5 0 0	0 5 0	0 0 5			Yes	Yes
	1036	2 0 2	1 1 1	1 1 0	0 3 0	0 1 1	1 0 1	5 0 0	0 5 0	0 0 5			Yes	Yes
	1003	4 0 0	1 2 2	0 1 2	0 2 1			5 0 0	0 5 0	0 0 5			No	Yes
	1068	1 1 1	3 0 1	1 2 1	0 2 2			5 0 0	0 5 0	0 0 5			No	Yes
	1004	1 2 1	2 1 2	2 1 1	0 1 1			2 1 1	1 0 2	2 0 0	0 3 0	0 1 2	No	No
	1002	2 0 1	1 0 1	1 1 0	1 2 1	0 2 2		1 0 3	2 0 1	1 2 1	1 3 0		No	No
	1014	3 1 0	1 0 2	1 1 1	0 1 2	0 2 0		2 1 1	0 2 1	1 1 2	2 0 2		No	No
	1025	1 0 1	2 1 0	1 2 1	1 1 1	1 1 1		2 2 1	1 2 2	2 1 2			No	No
	1032	3 1 1	1 0 2	1 2 1	0 2 1			1 4 0	4 1 0	0 0 5			No	No
	1021	1 1 2	1 2 1	2 1 1	1 1 1			4 0 1	0 4 1	1 1 3			No	No
	1034	2 1 0	1 1 2	2 3 3				2 3 1	1 2 3	2 2 0 1			No	No
	1045	2 0 1	2 1 1	1 1 1	0 3 1			4 1 0	1 4 2	2 0 3			No	No
	1050	2 2 1	1 0 2	1 1 1	1 2 1			3 1 1	2 1 1	0 3 0	0 0 3		No	No
	1061	3 0 0	0 3 1	1 0 3	0 2 2			2 2 1	1 2 2	2 1 2			No	No
	1069	2 3 2	3 2 0	0 0 3				3 2 0	2 3 0	0 0 5			No	No
1PIC	1017	3 1 0	1 1 1	1 2 1	0 1 3			5 0 0	0 5 0	0 0 5			Yes	Yes
	1006	1 3 1	3 1 1	1 1 3				5 0 0	0 5 0	0 0 5			Yes	Yes
	1008	2 0 2	2 1 0	1 3 2	0 1 1			5 0 0	0 5 0	0 0 5			No	Yes
	1012	2 1 2	1 2 2	1 0 0	1 2 1			5 0 0	0 5 0	0 0 5			No	Yes
	1051	2 1 1	2 4 0	1 0 4				5 0 0	0 5 0	0 0 5			No	Yes
	1030	3 1 1	1 2 1	1 2 3				5 0 0	0 5 0	0 0 5			No	Yes
	1018	2 1 1	1 1 1	1 0 0	1 2 1	1 0 2		5 0 0	0 5 0	0 0 5			No	Yes
	1022	3 1 0	1 0 1	1 2 1	0 2 3			5 0 0	0 1 4	0 4 1			No	No
	1038	3 1 1	1 0 0	1 1 2	0 3 2			5 0 0	0 4 1	0 1 4			No	No
	1005	2 1 3	2 2 1	1 2 1				5 0 0	0 2 3	0 3 2			No	No
	1048	2 2 2	2 3 3	1 0 0				2 3 1	2 2 4	1 0 0			No	No
	1020	2 1 1	1 1 2	1 0 0	1 1 0	0 2 2		5 1 2	0 4 3				No	No
	1060	2 1 2	1 1 1	1 3 2	1 0 0			5 0 0	0 2 3	0 3 2			No	No
	1065	2 3 2	1 0 0	2 0 3	0 2 0			5 0 0	0 1 4	0 4 1			No	No
	1072	2 0 3	1 0 0	2 4 0	0 1 2			5 0 0	0 3 1	0 2 4			No	No

Note. Shaded cells indicate sorting according with the experimenter-defined classes. ABS = abstract; 1PIC = C1 as picture

classes and indeterminate patterns of responding consist of stimulus relations drawn from different experimenter-defined classes such as B1-D2 or C1-E3. However, for a relation to qualify as a participant-defined relation, it must occur three

Table 3 Pre-class formation sorting tests and post-class formation sorting tests outcomes for 2PIC and 3PIC

Group	Participant	Pre Cls-1/2/3	Pre Cls-1/2/3	Pre Cls-1/2/3	Pre Cls-1/2/3	Pre Cls-1/2/3	Pre Cls-1/2/3	Post Cls-1/2/3	Post Cls-1/2/3	Post Cls-1/2/3	Post Cls-1/2/3	Post Cls-1/2/3	ECF	Sort
2PIC	1019	1 2 0	2 1 1	1 1 1	1 1 3			5 0 0	0 5 0	0 0 5			Yes	Yes
	1040	3 3 1	1 1 3	1 1 1				5 0 0	0 5 0	0 0 5			Yes	Yes
	1023	2 4 1	3 1 4					5 0 0	0 5 0	0 0 5			Yes	Yes
	1011	3 1 1	1 0 1	1 4 3				5 0 0	0 5 0	0 0 5			Yes	Yes
	1031	3 1 3	2 4 2					5 0 0	0 5 0	0 0 5			Yes	Yes
	1015	2 3 0	2 1 2	1 1 3				5 0 0	0 5 0	0 0 5			Yes	Yes
	1066	1 1 1	3 3 1	1 1 3				5 0 0	0 5 0	0 0 5			Yes	Yes
	1071	3 1 0	1 0 1	1 3 0	0 1 4			5 0 0	0 5 0	0 0 5			Yes	Yes
	1026	2 0 2	1 2 0	1 1 1	1 2 2			5 0 0	0 5 0	0 0 5			No	Yes
	1049	4 2 1	1 0 1	1 2 3				5 0 0	0 5 0	0 0 5			No	Yes
	1062	3 0 0	1 1 1	1 3 0	0 1 3			5 0 0	0 5 0	0 0 5			No	Yes
	1010	3 1 0	1 0 1	1 1 3	0 3 1			4 1 0	1 4 0	0 0 5			No	No
	1007	1 1 0	2 0 0	1 0 1	1 1 3	0 3 1		5 0 2	0 3 5				No	No
	1013	1 1 1	2 0 1	1 1 1	1 2 0	0 1 2		1 1 2	1 2 1	2 1 0	1 1 2		No	No
	1047	2 1 0	2 2 2	1 0 1	0 2 2			1 3 0	4 0 0	0 2 0	0 0 5		No	No
3PIC	1024	1 1 1	2 0 1	1 1 1	1 2 0	0 2 1		5 0 0	0 5 0	0 0 5			Yes	Yes
	1027	3 0 0	1 1 1	1 0 2	0 3 0	0 1 2		5 0 0	0 5 0	0 0 5			Yes	Yes
	1029	2 2 0	1 1 1	1 0 3	0 2 2			5 0 0	0 5 0	0 0 5			Yes	Yes
	1052	1 1 1	3 0 1	1 2 1	0 2 2			5 0 0	0 5 0	0 0 5			Yes	Yes
	1009	2 1 1	1 1 1	2 2 1	0 1 2			5 0 0	0 5 0	0 0 5			Yes	Yes
	1046	2 2 0	2 1 1	1 1 1	0 1 3			5 0 0	0 5 0	0 0 5			Yes	Yes
	1037	1 1 2	2 1 1	1 1 1	0 2 0			5 0 0	0 5 0	0 0 5			Yes	Yes
	1039	1 1 2	3 2 1	1 2 2				5 0 0	0 5 0	0 0 5			Yes	Yes
	1033	2 1 0	1 1 1	1 1 0	1 2 2	0 0 2		5 0 0	0 5 0	0 0 5			Yes	Yes
	1063	2 3 1	1 1 1	2 1 3				5 0 0	0 5 0	0 0 5			Yes	Yes
	1067	2 2 1	1 1 1	1 2 1	1 0 2			5 0 0	0 5 0	0 0 5			Yes	Yes
	1070	2 1 0	1 1 1	1 2 0	1 1 4			5 0 0	0 5 0	0 0 5			Yes	Yes
	1028	3 1 0	1 1 1	1 2 2	0 2 1			5 0 0	0 5 0	0 0 5			No	Yes
	1035	2 3 1	1 1 2	2 2 1				2 0 0	1 1 0	1 1 0	0 2 0	0 0 5	No	No
	1016	2 1 2	2 2 1	1 2 2				4 0 1	1 2 0	0 0 3	0 3 1		No	No

Note. Shaded cells indicate sorting according with the experimenter-defined classes. 2PIC = C1 and C2 as pictures; 3PIC = C stimuli as pictures

Fig. 4 Errors by participants who did not form classes in the 1PIC (C1-as-PIC) and 2PIC (C1 and C2 as pictures) condition. *Upper panel* shows errors made by participants in 1PIC; *lower panel* shows errors made by participants in 2PIC. Error bars show standard error of the mean
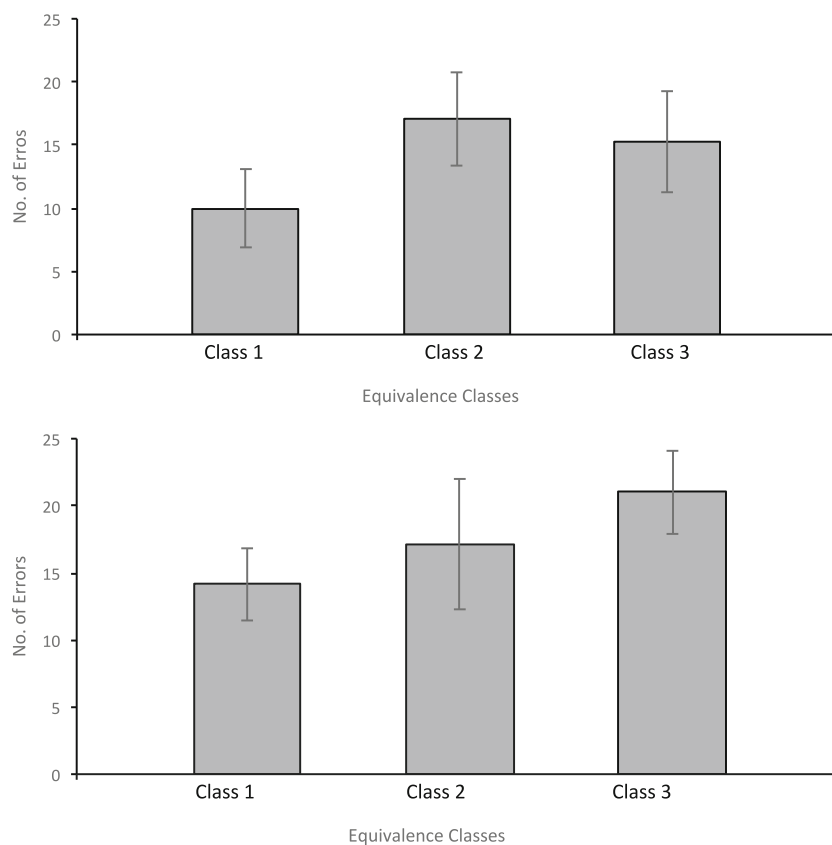

times in the test block, while indeterminate responding occurs when the trial type occurred less than three times in the test block. As shown in Fig. 5, for Classes 1, 2, and 3, an average of 69.07%, 64.27%, and 61.98% probes produced responses in accordance with experimenter-defined Classes of 1, 2, and 3, respectively. A total of 9.77 probe trials produced responses in accordance with participant-defined classes, and a total of 27.56 probe trials produced indeterminate responses. As shown in Fig. 5, 50.94%, 69.17%, 70.2%, and 65.92% of probe trials produced responses in accordance with experimenter-defined classes for the ABS, 1PIC, 2PIC, and 3PIC groups respectively. Also, 11.11%, 10.83%, 5.28%, and 11.67% of probe trials produced responses in accordance with participant-defined classes for the ABS, 1PIC, 2PIC, and 3PIC groups respectively. Finally, 37.95%, 20%, 24.52%, and 22.41% of probe trials produced indeterminate responses for the ABS, 1PIC, 2PIC, and 3PIC groups, respectively.

Response Speed

The response speed or inversed reaction time (time from presentation of the comparisons to a choice was made) was analyzed using an average of the median of the speed of responding for the last five baseline trials for correct and incorrect responses, first five and last five of

correct test trials, and first five and last five of incorrect test trials. These data were further divided among participants who formed classes (passers) and participants who did not form classes (failers) regardless of the experimental group. As shown in Fig. 6 (upper panel/ passers), at the end of baseline relation training, the response speed for the first five emergent relations trials was greater on correct selections relative to incorrect selections. Participants had no incorrect selections for the last five emergent relations trials. Finally, the response speed for the last five emergent relations trials
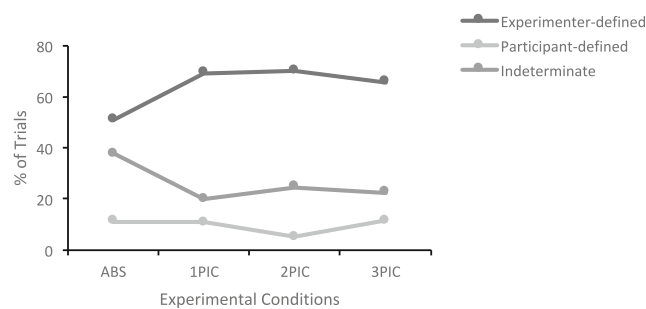
Fig. 5 Proportion of trials indicative of experimenter-defined relations, participant-defined relations, and indeterminate relations during emergence relations test. ABS = abstract; 1PIC = C1 as pictures; 2PIC = C1 and C2 as pictures; 3PIC = C1, C2, and C3 as pictures

Fig. 6 Average speed for all participants regardless of experimental groups. *Upper panel* shows response speed of participants who responded according to equivalence (*passers*); *lower panel* shows response speed of participants who did not respond according to equivalence (*failers*). BSL = baseline; SYM = symmetry; TR = transitivity; EQ = equivalence
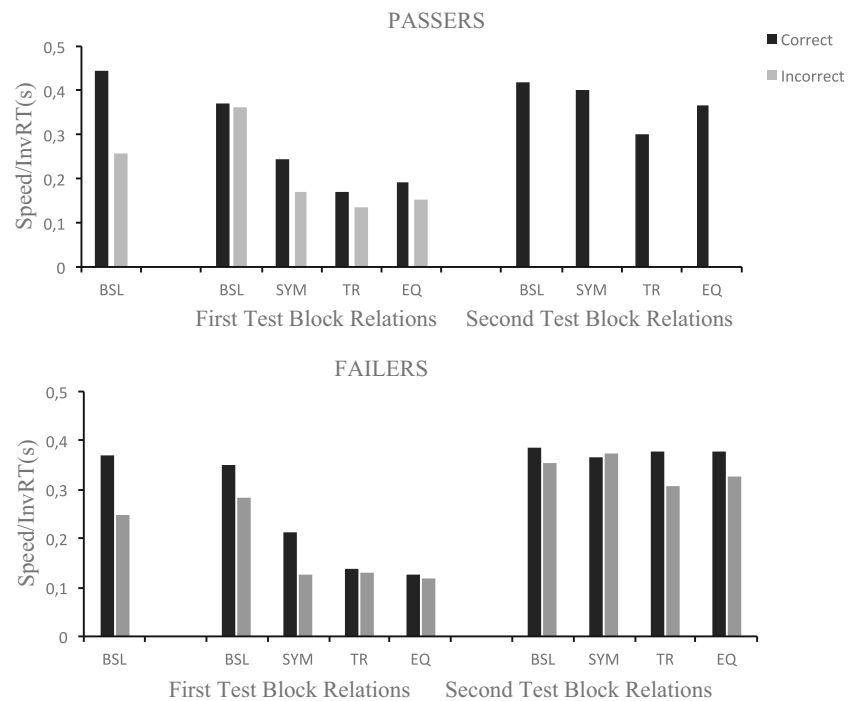
was faster relative to the first five emergent relations trials. The lower panel (failers) of Fig. 6 shows that, at the end of baseline relation training, the response speed for the first five and last five emergent relations trials was faster on trials for correct selections relative to incorrect selections except for the last five symmetry trials. Finally, the response speed for the last five emergent relations trials was faster relative to the first five emergent relations trials.

Discussion

The present experiment investigated the differential effects of the number of C stimuli as meaningful stimuli in three 5-member equivalence classes using the linear series training structure with the administration of training and testing under the simultaneous protocol. The findings from the experiment showed that two of 15 participants in the ABS group formed classes. In addition, two of 15 participants in the 1PIC group formed classes, eight of 15 participants in the 2PIC group formed classes, and 12 of 15 members in the 3PIC group formed classes. There was a significant difference in equivalence class formation between the abstract group (ABS) and all the picture groups (1PIC, 2PIC, and 3PIC). A significant difference was also found between the 1PIC group and the 3PIC group, but not for the 2PIC group.

Equivalence Class Formation

As shown by the results, the likelihood of participants forming equivalence classes is increased by the introduction of meaningful stimuli (as the nodal stimuli) in a potential class of abstract stimuli. These results are consistent with the findings reported by Arntzen, Nartey. et al. (2015) and Fields et al. (2012), which showed an increased yield when C stimuli were pictures. Furthermore, for participants who did not form classes, the number of probes that produced experimenter-defined classes was greater for the 1PIC, 2PIC, and 3PIC groups relative to the ABS group. Fields et al. (2012), and Travis et al. (2014) posit that the meaningful stimulus enhances class formation because it is most likely a member of some established equivalence class prior to the experiment and serves as a discriminative stimulus for the formation of equivalence classes in a set of abstract stimuli. Therefore, the formation of a five-member equivalence class consisting of one meaningful stimulus and four abstract stimuli reflects an expansion of a preexisting class by the addition of four abstract stimuli to the already existing class of which the meaningful stimulus is a member (e.g., Arntzen, Nartey, et al., 2015; Fields et al., 2012). However, it could also be argued that the possibility of class formation enhancement was due to the color of the meaningful stimuli and not necessarily the stimuli themselves. It may well be that the enhancement was facilitated by the compound (color + stimuli).

Meaningful Stimuli Contained in Different Numbers of Classes

The yields of equivalence class formation as a result of the different numbers of meaningful included in 1PIC, 2PIC, and 3PIC give rise to two questions; first, why is there a lack of significant difference in equivalence class formation between 2PIC and 3PIC? Second, why is there a significant difference in equivalence class formation between 1PIC and 3PIC? The lack of significant difference between 2PIC and 3PIC may be attributed to rejection control in the experiment and not necessarily due to strong stimulus control set up by the experimental contingencies. Thus, the performance of the 2PIC in relation to the 3PIC may be spurious or an artifact resulting from false positives (Sidman, 1987; Stikeleather & Sidman, 1990). Also, the lack of significant difference between 2PIC and 3PIC may be attributed to learning by exclusion promoted by the experimental design or operations (Ferrari, de Rose, & McIlvane, 2008; Wilkinson, Dube, & McIllvane, 1998; Wilkinson, Rosenquist, & McIlvane, 2009). Participants, therefore, come to exclude the Class 3 members by rejecting to assign them to either Class 1 or Class 2. The significant difference between the 1PIC and 3PIC can be attributed to the inclusion of a picture in only one class of the 1PIC relative to the inclusion of pictures in all the classes of the 3PIC. This is because the classes in the 1PIC and 2PIC without pictures lead to more errors of responding relative to the classes with pictures. The finding that the inclusion of pictures facilitates correct responding has been discussed by Bentall et al. (1993), and Dickins, Bentall, and Smith (1993). These authors have argued that the inclusion of nameable stimuli in a class of abstract stimuli enhances class formation. This is because the introduction of nameable stimuli in a class allows for participants to assign same names to the stimuli belonging to the same class. On the other hand, findings from Nartey et al. (2014) indicate that performance on equivalence class formation varied in accordance with the order of the introduction of the meaningful stimuli. Therefore, if naming was that critical to equivalence class formation, then the order of presentation of the meaningful stimuli should not have any effect on equivalence class formation. Sidman (1994) buttresses this point by indicating naming can only be facilitative and not a prerequisite for equivalence class formation.

Speed of Acquisition and Response Speed

In this experiment, the number of trials to criterion or speed of acquisition seem to be a good predictor of forming equivalence classes or not. This finding is in accordance with previous studies (e.g., Nartey et al., 2015a). Hence, more studies need to confirm this finding.

This study replicated the general finding of previous studies that speed of responding decreases from training to testing

(Arntzen et al., 2011; Arntzen & Hansen, 2011; Arntzen, Nartey, et al., 2015; Spencer & Chase, 1996). Furthermore, the speed of responding (inversed reaction time) to comparison stimuli decreased from baseline test trials to equivalence test trials for all participants. Specifically, response speed was an inverse function of emergent relations test difficulty. This finding confirms to a large extent the assertions by Bentall et al. (1993) and Spencer and Chase (1996) that test difficulty arising from nodal distance leads to a decrease in response speed.

The results showed that the speed of responding was higher for the last five testing trials compared to the first five testing trials. The increase in the speed of response in the later part of the testing trials could be due to gained familiarity of the testing trials. Also, the speed of responding was faster for correct responses compared to incorrect responses (Arntzen Nartey et al., 2015; Spencer & Chase, 1996). In addition, speed of responding was much faster for participants who formed classes (passers) compared to participants who did not form classes (failers), which is in accordance with Arntzen, Nartey, et al. (2015).

This variation in response speed may be attributed to the effect of one type of stimulus control topography for the correct test trials compared to a variety of stimulus control topographies for the incorrect test trials. Thus, the slow response speed for failers could be a result of a conflict arising from many controlling stimulus control topographies that are not consistent with the experimenter-defined topography (see McIlvane & Dube, 2003, for a discussion). The finding also emphasizes that participants who do not form equivalence classes do not necessarily respond fast just to finish the test.

Sorting

The present experiment replicated previous studies on sorting as a measure of equivalence class formation (Arntzen et al., 2016; Arntzen, Norbom, et al., 2015; Fields et al., 2014; Fields et al., 2012). In this study, all participants who showed the formation of the three experimenter-defined equivalence classes in the MTS-based emergent relations test performance showed maintenance of these three classes in the post-sorting class formation test. Furthermore, some participants who did not show the formation of the three experimenter-defined equivalence classes produced the three experimenter-defined equivalence classes in the post-sorting class formation test. This finding implies that the post-sorting class formation test may have provided a more sensitive measure of equivalence class formation than the one provided in the MTS-based emergent relations test. This finding also confirms and extends the knowledge about the correlation between MTS test and sorting test (Fields et al., 2014). Hence, for the ABS group, two participants who did not form equivalence classes did sort the stimuli in accordance with the experimenter-defined classes. More important, for the 1PIC and 2PIC groups,

five and three participants who did not form equivalence classes did sort the stimuli in accordance with the experimenter-defined classes, respectively.

Future Studies

In this study, one could argue that the pictorial stimuli are different from the abstract shapes in other dimensions than they are meaningful or familiar. There is a possibility that the class formation enhancement was due to color in the meaningful stimuli and not the stimuli themselves. A future study should control for this potential confound and use all black and white stimuli as a control. Another experiment could involve an assessment of the abstract shapes used because it could be argued that some of the stimuli could have some degree of meaningfulness. This experiment contributes to the knowledge about how classes of stimuli are partitioned and the most efficient way to do this by having all C stimuli as pictures. Thus, an important extension would be to arrange an experiment with three pictures located as C1, B2, and D3 in a three 5-member class (A→B→C→D→E). Such an experiment will clarify if it is the ratio of pictures or the pictures within the same set that is the critical variable. Finally, because there already is a substantial amount of reports on classes with five members or fewer that have shown the effect of meaningful stimuli, we need to find out if the effect also holds for larger equivalence classes.

Summary

This study extends our knowledge of the influence of the inclusion of meaningful stimuli in equivalence research. It strengthens the view that placing meaningful stimuli in a set of abstract stimuli helps in converting that set into an equivalence class. Also, the findings on the inclusion of meaningful pictures have a great impact on teaching and the arrangement of effective conditional discrimination procedures in behavioral programs. Thus, the findings bring to bear that in a potential three 5-member equivalence class, one picture is not enough to enhance the formation of equivalence classes. The study also strengthens the view that there is a correlation between MTS testing and sorting.

Our experiment explored equivalence class formation as a function of the number of meaningful stimuli as C stimuli in a three 5-member equivalence class (A→B→C→D→E). The main findings were that, when all stimuli were abstract, 13% of the participants formed classes, whereas 80% of participants formed classes when the C members were all meaningful and the other stimuli were abstract. When the C1 and C2 members were meaningful stimuli, and all the other stimuli were abstract, 53% of the participants formed classes. Finally, when only C1 was a meaningful stimulus and all the other stimuli were abstract, 13% of the participants formed classes. For all four groups, more

participants formed classes in the post-sorting class formation test than in the MTS-based emergent relations test.

References

Arntzen, E. (2004). Probability of equivalence formation: Familiar stimuli and training sequence. *The Psychological Record, 54*, 275–291. Retrieved from http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1235&context=tpr.

Arntzen, E., Braaten, L. F., Lian, T., & Eilifsen, C. (2011). Response-to-sample requirements in conditional discrimination procedures. *European Journal of Behavior Analysis, 12*, 505–522. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/15021149.2011.11434398.

Arntzen, E., Granmo, S., & Fields, L. (2016). The relation between sorting tests and matching-to-sample tests in the formation of equivalence classes. *The Psychological Record*. doi:10.1007/s40732-016-0209-9.

Arntzen, E., & Hansen, S. (2011). Training structures and the formation of equivalence classes. *European Journal of Behavior Analysis, 12*, 483–503. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/15021149.2011.11434397.

Arntzen, E., & Lian, T. (2010). Trained and derived relations with pictures as nodes. *The Psychological Record, 60*, 659–677. Retrieved from http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1291&context=tpr.

Arntzen, E., Nartey, R. K., & Fields, L. (2014). Identity and delay functions of meaningful stimuli: Enhanced equivalence class formation. *The Psychological Record, 64*, 349–360. doi:10.1007/s40732-014-0066-3.

Arntzen, E., Nartey, R. K., & Fields, L. (2015). Enhancing responding in accordance with stimulus equivalence by the delayed and relational properties of meaningful stimuli. *Journal of the Experimental Analysis of Behavior, 103*, 524–541. doi:10.1002/jeab.152.

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

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. doi:10.1080/14640749308401085.

Cowley, B. J., Green, G., & Braunling-McMorrow, D. (1992). Using stimulus equivalence procedures to teach name-face matching to

adults with brain injuries. *Journal of Applied Behavior Analysis, 25*, 461–475. doi:10.1901/jaba.1992.25-461.

Dickins, D. W. (2011). Transitive inference in stimulus equivalence and serial learning. *European Journal of Behavior Analysis, 12*, 523–555. Retrieved from www.tandfonline.com/doi/abs/10.1080/15021149.2011.11434399.

Dickins, D. W. (2015). A simpler route to stimulus equivalence? A replication and further exploration of a "simple discrimination training procedure" (Canovas, Debert and Pilgrim 2014). *The Psychological Record*, 1–11. doi:10.1007/s40732-015-0134-3.

Dickins, D. W., Bentall, R. P., & Smith, A. B. (1993). The role of individual stimulus names in the emergence of equivalence relations: The effects of interpolated paired-associates training of discordant associates between names. *The Psychological Record, 43*, 713–724.

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/1/wp-content/uploads/2014/03/Dymond-Rehfeldt.pdf.

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.tandfonline.com/doi/abs/10.1080/15021149.2009.11434318.

Eilifsen, C., & Arntzen, E. (2011). Single-subject withdrawal designs in delayed matching-to-sample procedures. *European Journal of Behavior Analysis, 12*, 152–172. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/15021149.2011.11434361.

Ferrari, C., de Rose, J. C., & McIlvane, W. J. (2008). A comparison of exclusion and trial-and-error procedures: Primary and secondary effects. *Experimental Analysis of Human Behavior Bulletin, 29*, 9–16. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2853183/.

Fields, L., Arntzen, E., Nartey, R. K., & Eilifsen, C. (2012). Effects of a meaningful, a discriminative, and a meaningless stimulus on equivalence class formation. *Journal of the Experimental Analysis of Behavior, 97*, 163–181. doi:10.1901/jeab.2012.97-163.

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

Fienup, D. M., & Dixon, M. R. (2006). Acquisition and maintenance of visual-visual and visual-olfactory equivalence classes. *European Journal of Behavior Analysis, 6*, 87–98. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/15021149.2006.11434266.

Green, G. (1990). Differences in development of visual and auditory-visual equivalence relations. *American Journal of Mental Deficiency, 95*, 260–270.

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://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1542&context=tpr.

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://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1501&context=tpr.

Lowe, C. F., Horne, P. J., Harris, F. D. A., & Randle, V. R. L. (2002). Naming and categorization in young children: Vocal tact training. *Journal of the Experimental Analysis of Behavior, 78*, 527–549. doi:10.1901/jeab.2005.31-04.

Mackay, H. A., Wilkinson, K. M., Farrell, C., & Serna, R. W. (2011). Evaluating merger and intersection of equivalence classes with one member in common. *Journal of the Experimental Analysis of Behavior, 96*, 87–105. doi:10.1901/jeab.2011.96-87.

McIlvane, W. J., & Dube, W. V. (2003). Stimulus control topography coherence theory: Foundations and extensions. *The Behavior Analyst, 26*, 195–213. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22478506.

Nartey, R. K., Arntzen, E., & Fields, L. (2014). Two discriminative functions of meaningful stimuli that enhance equivalence class formation. *The Psychological Record, 64*, 777–789. doi:10.1007/s40732-014-0072-5.

Nartey, R. K., Arntzen, E., & Fields, L. (2015a). Enhancement of equivalence class formation by pretraining discriminative functions. *Learning & Behavior, 43*, 20–31. doi:10.3758/s13420-014-0158-6.

Nartey, R. K., Arntzen, E., & Fields, L. (2015b). Training order and structural location of meaningful stimuli: Effects on equivalence class formation. *Learning & Behavior, 43*, 342–353. doi:10.3758/s13420-015-0183-0.

Nedelcu, R. I., Fields, L., & Arntzen, E. (2015). Conditional discriminative functions of meaningful stimuli and enhanced equivalence class formation. *Journal of the Experimental Analysis of Behavior, 103*, 349–360. doi:10.1002/jeab.141.

Pilgrim, C., & Galizio, M. (1996). Stimulus equivalence: A class of correlations or a relation of classes. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 173–195). Amsterdam, The Netherlands: Elsevier.

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

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

Sigurdardottir, Z. G., Mackay, H. A., & Green, G. (2012). Stimulus equivalence, generalization, and contextual stimulus control in verbal classes. *The Analysis of Verbal Behavior, 28*, 3–29. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3363409/.

Smeets, P. M., Dymond, S., & Barnes-Holmes, D. (2000). Instructions, stimulus equivalence, and stimulus sorting: Effects of sequential testing arrangements and a default option. *The Psychological Record, 50*, 339–354. Retrieved from http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1401&context=tpr.

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.

Stikeleather, G., & Sidman, M. (1990). An instance of spurious equivalence relations. *The Analysis of Verbal Behavior, 8*, 1–11. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2748619/pdf/anverbbehav00037-0003.pdf.

Travis, R. W., Fields, L., & Arntzen, E. (2014). Discriminative functions and over-training as class-enhancing determinants of meaningful stimuli. *Journal of the Experimental Analysis of Behavior, 102*, 47–65. doi:10.1002/jeab.91.

Wilkinson, K. M., Dube, W. V., & McIllvane, W. J. (1998). Fast mapping and exclusion (emergent matching) in developmental language, behavior analysis, and animal cognition research. *The Psychological Record, 48*, 407–422. Retrieved from http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1517&context=tpr.

Wilkinson, K. M., Rosenquist, C., & McIlvane, W. J. (2009). Exclusion learning and emergent symbolic category formation in individuals with severe language impairments and intellectual disabilities. *The Psychological Record, 59*, 187–206. Retrieved from http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1014&context=tpr.

Wulfert, E., & Hayes, S. C. (1988). Transfer of conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50*, 125–144. doi:10.1901/jeab.1988.50-125.