

Training order and structural location of meaningful stimuli: Effects on equivalence class formation

Richard K. Nartey¹ · Erik Arntzen¹ · Lanny Fields²

Published online: 16 June 2015
© Psychonomic Society, Inc. 2015

Abstract In the present study, equivalence class formation was influenced by the temporal point of inclusion of a meaningful stimulus when baseline relations were serially or sequentially trained, and much less so by the location of the meaningful stimulus in the nodal structure of the class. In Experiment 1, participants attempted to form three 3-node, 5-member classes ($A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$) under the simultaneous protocol. After serially training the baseline relations AB, BC, CD, and DE, in that order, the emergence of all emergent relations was tested concurrently. In the A-as-PIC condition, A was meaningful stimulus and B to E were meaningless stimulus, and 60 % of the participants formed classes. In addition, classes were formed by 40 %, 70 %, 40 %, and 20 % of the participants in the B-as-PIC, C-as-PIC, D-as-PIC, and E-as-PIC groups, respectively. Thus, the likelihood of class formation could have been influenced by the location of a meaningful stimulus in the class structure and/or by its order of introduction during training. In Experiment 2, we controlled for any effect of order of introduction by the concurrent training of all of the baseline relations. Regardless of the location of the meaningful stimulus, 0–20 % of participants formed classes. Thus, the temporal order of introducing a meaningful stimulus was the primary modulator of the class-enhancing property of meaningful stimuli, and not the location of the meaningful stimulus in the class structure.

Keywords Meaningful stimulus · Stimulus equivalence · Simultaneous protocol · Serialized · Order of training

When a stimulus that controls a response can be replaced with other perceptually disparate stimuli without altering the probability of responding, the stimuli are functionally interchangeable with each other. These stimuli, which were initially unrelated to each other, are said to be acting as members of an equivalence class, and also to show the single properties of reflexivity, symmetry, and transitivity, as well as the combined properties of symmetry and transitivity (Sidman & Tailby, 1982). For instance, if a child who only speaks and understands one of the local languages from Ghana is then taught to match *kraman* (the Akan word for dog) to the English word *dog* and is also trained to match *dog* to *hund* (the Norwegian word for dog), the emergence of untrained relations that show reflexivity ($kraman \rightarrow kraman$, $dog \rightarrow dog$, $hund \rightarrow hund$), symmetry ($dog \rightarrow kraman$, $hund \rightarrow dog$), transitivity ($kraman \rightarrow hund$), and equivalence—the combined properties of symmetry and transitivity ($hund \rightarrow kraman$)—will demonstrate that all of the words function as members of one equivalence class. Thus, stimulus equivalence is synonymous with stimulus *substitutability* (Green & Saunders, 1998).

The formation of equivalence classes has been documented with a diverse range of participants, including typically developing children (e.g., Arntzen & Lian, 2010; Pilgrim, Chambers, & Galizio, 1995), autistic children (e.g., Arntzen, Halstadro, Bjerke, & Halstadro, 2010; LeBlanc, Miguel, Cummings, Goldsmith, & Carr, 2003), and typically developing adults (e.g., Bentall, Dickins, & Fox, 1993; Pilgrim & Galizio, 1995). The classes, which had different training structures, varied in terms of the number of nodes and directionality of training (e.g., Arntzen & Holth, 1997; Arntzen, Grondahl, & Eilifsen, 2010; Arntzen & Hansen, 2011) and were formed

✉ Richard K. Nartey
rknartey@yahoo.co.uk

✉ Erik Arntzen
erik.arntzen@equivalence.net

¹ Department for Behavioral Science, Oslo and Akershus University College, PO Box 4, St. Olavs Plass, 0130 Oslo, Norway

² Queens College and The Graduate School, City University of New York, New York, NY, USA

using a variety of training protocols (Imam, 2006). The sensory modality of the stimuli used as class members—for instance, visual (e.g., Fields, Arntzen, Nartey, & Eilifsen, 2012), tactile (e.g., Belanich & Fields, 1999), and gustatory (e.g., Hayes, Tilley, & Hayes, 1988)—has also been shown to influence the formation of equivalence classes. When visual stimuli were used as the class members, they consisted of abstract shapes (e.g., Sidman & Tailby, 1982) or meaningful pictures (e.g., Arntzen, 2004; Fields et al., 2012). In addition, the inclusion of a meaningful stimulus at different temporal points in training influenced the likelihood of equivalence class formation (Arntzen, 2004; Fields et al., 2012; Nartey, Arntzen, & Fields, 2014). In Nartey et al. (2014), participants attempted to form three 3-node, 5-member equivalence classes that consisted of four abstract stimuli and one meaningful stimulus. After the training of AB, BC, CD, and DE relations, in that serial order, classes were formed by 70 % of participants when the A stimuli were meaningful pictures and were members of the first-trained AB relations, while the B through E stimuli were abstract stimuli. When the E stimuli were meaningful pictures and were members of the last-trained DE relations, while A through D were abstract stimuli, classes were formed by only 40 % of participants. Thus, the formation of large multinodal classes could have been influenced by the location of a meaningful stimulus in the structure of a class (i.e., A or E) and/or by their order of introduction during training (first or last). In that study, however, the effects of the pictures used as the B, C, or D stimuli were not explored. Thus, that study did not fully disclose how the class-enhancing effects of a meaningful stimulus is modulated by its order of introduction during training and/or by its structural location in a potential equivalence class.

Fields et al. (2012) explored the use of the meaningful stimulus as the C member of the class. None of the participants formed classes when all of the stimuli were abstract. However, when the A, B, D, and E stimuli were the same abstract stimuli but the C stimuli, which served as nodes, were meaningful pictures, 80 % of participants formed classes. In systematic explorations, Travis, Fields, and Arntzen (2014) and Nedelcu, Fields, and Arntzen (2015) also found that the use of the meaningful stimulus as the middle node had a substantial class enhancement effect. Thus, the formation of large multinodal classes was also enhanced by the inclusion of a meaningful stimulus as the middle node.

Experiment 1

When the results of these experiments are considered, the use of a meaningful stimulus as the A or the C member of the class appears to produce a similar high yield, whereas the use of the meaningful stimulus as the E member of the class produces a much lower yield. This conclusion, however, must be qualified, because the A and E data were obtained in one

experiment and the C data in another; thus, the relative effects of the A and C stimuli could have reflected the influences of other variables confounded with the effects attributed to the A and C stimuli. That interpretive problem could be obviated by conducting a single experiment that used the meaningful stimuli as the A, C, or E members of equivalence classes. In addition, comprehensive information could be obtained by using the meaningful stimuli as the B or D members of the equivalence classes. Conducting such an experiment would determine whether the class enhancement effect of a meaningful stimulus declines systematically as its location in a class varies from the C to the D to the E stimulus, or whether it drops precipitously, such that the same low yield would be produced when the meaningful stimulus was used in the D or the E location in the class structure. Either outcome would inform the development of a plausible theoretical account of the variables responsible for modulating the class-enhancing effects of a meaningful stimulus that is a member of a to-be-formed equivalence class.

Experiment 1 consisted of five groups. In each, the participants attempted to form 3-node, 5-member equivalence classes with an $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$ training structure, in which four of the five stimuli were abstract and one was a meaningful stimulus. The simultaneous protocol was used for training and testing (Buffington, Fields, & Adams, 1997; Fields et al., 1997; Imam, 2006). Because this protocol typically produces relatively poor yields, it is a preparation that is sensitive to variables that influence the likelihood of equivalence class formation (Arntzen, 2012; Fields et al., 2012; Fields, Hobbie-Reeve, Adams, & Reeve, 1999). Baseline relations were established in a serial order as AB, BC, CD, and DE, respectively. The acquisition of all of the baseline relations was followed by the administration of two test blocks for all emergent relations. In the different groups, the meaningful stimulus was included in the set of other meaningless stimuli as the A, B, C, D, or E member of a to-be-formed class. This experiment would then document how the likelihood of equivalence class formation was influenced by the inclusion of a meaningful stimulus as each member of a to-be-formed class. The results of the experiment would identify how the order of introducing the meaningful stimulus and its location in the nodal structure of the class influenced the class enhancement by meaningful stimuli.

Method

Participants

Fifty university students (27 males and 23 females) between the ages of 19 to 25 years voluntarily participated in this study. The average age was 21 years ($SD = 1.4$). None of the participants had any prior knowledge of stimulus equivalence research and methodology. Participants were assigned on a

block-randomized basis to one of the five experimental groups.

Apparatus

Setting The experiments were conducted in the graduate seminar room of the Department of Psychology at the University of Ghana, Legon. The seminar room measured approximately 5 m square and was furnished with tables and chairs.

Hardware The experiments were conducted on an HP Compaq nc6320 laptop computer that used an 1828-MHz Intel Centrino processor and had a screen with a 16.8-in. diagonal length and a 16 × 9 horizontal-to-vertical ratio. An external mouse was used by participants to control the position of the cursor throughout the experiment.

Software All sessions for training and testing of conditional discriminations for all of the participants were conducted with an MTS software program, version 3.12, made by Psych Fusion Software in collaboration with the second author. The software controlled the presentation of all stimuli and also made recordings of data including the trial number, number of training trials, reaction times to the sample and comparison stimuli, whether or not participants made the correct/incorrect comparison choice, and whether or not programmed consequences were delivered. Finally, the software provided a summary of the symmetry, transitivity, and equivalence tests, as well as the duration of the experiment.

Stimuli The stimulus sets used in the experiment are shown in Fig. 1. The top section of the figure shows the 15 abstract stimuli used as members of the equivalence classes, whilst the bottom section shows the three meaningful pictures—the stimuli used to replace the A, B, C, D, or E abstract stimuli in the A-as-PIC, B-as-PIC, C-as-PIC, D-as-PIC, or E-as-PIC condition, respectively. The abstract stimuli were displayed in black and the picture stimuli in color, both on a white background. The size of the stimuli as they were presented on the computer monitor was 9.4 × 3.4 cm, which also corresponded to the touch-sensitive area of the screen. Although no explicit measures were taken to determine whether any of the stimuli had some meaning or significance to the participants, none of the participants commented on the meaningfulness or familiarity of the stimuli.

Procedure

Design The participants were randomly assigned to one of the following five groups: (1) A-as-PIC, (2) B-as-PIC, (3) C-as-PIC, (4) D-as-PIC, and (5) E-as-PIC. In the A-as-PIC groups, the A stimuli were the meaningful pictures shown in the bottom section of Fig. 1, whereas the B–E stimuli were

the same abstract stimuli in the top section of the figure. Likewise, the B, C, D and E stimuli in the B-as-PIC, C-as-PIC, D-as-PIC, and E-as-PIC groups, respectively, were the meaningful pictures, whereas the other stimuli were abstract.

Preliminary information Upon arrival at the experimental setting, participants were made to take a seat and given an informed consent document to read. The document informed participants that they were about to participate in an experiment in the field of behavior analysis that would last approximately one-and-a-half hours, that they were required to respond to certain stimuli on the screen of a computer with mouse clicks, and that there were no known harmful effects of participating in the study. They were also told that they could choose to withdraw from the experiment at any time without any negative consequences. After reading this information on a form, those who agreed to be participants signed the form and began the experiment.

Card sorting The participants in the experiment remained seated in the experimental cubicle and were given the 15 plastic-laminated cards that corresponded to the stimuli to be used in the group to which the participant was assigned and were told to “put them into groups.” All participants were presented with 12 abstract stimulus cards and the three picture stimuli. After the experiment, participants were asked to “put them into groups” again.

Instruction After the pre-class-formation sorting task, the participants remained seated in the experimental cubicle behind the computer and were presented with the following instructions 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 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 correct. Please do your best to get everything right. Thank you and good luck!

No further instructions were given before or after the experiment, started.

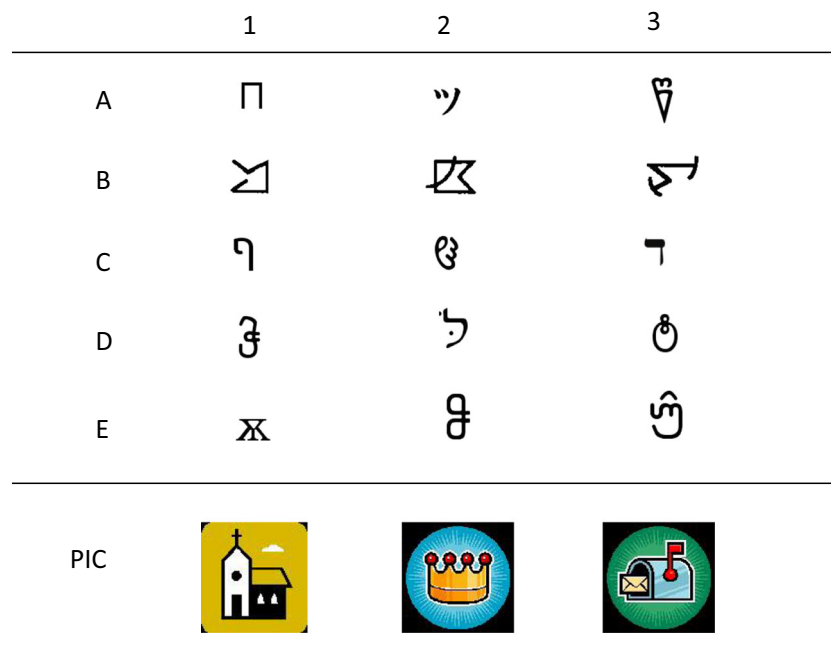


Fig. 1 Stimuli used as members of the equivalence classes. The top section shows the 15 abstract stimuli, whereas the bottom section shows the meaningful stimuli that replaced the abstract stimuli in the different groups

Trial structure and contingencies All participants were exposed to the simultaneous protocol in order to form equivalence classes, with training and test trials being presented in a matching-to-sample format. Each trial began with the presentation of the sample stimulus in the middle of the screen. Responding to the sample stimulus by a mouse click on it was immediately followed by the presentation of the three comparison stimuli at three of the four corners of the screen, while the sample stimulus still remained on the screen. The locations of the corners used to present the comparisons were randomized across trials.

A comparison was selected by moving the mouse cursor to it and then left-clicking the mouse. The selection of the correct comparison stimulus on a trial resulted in the removal of the sample and comparison stimuli and the presentation of the words *correct*, *very good*, *super*, or *excellent* on the screen. Clicking on one of the incorrect comparison stimuli also resulted in the removal of the stimuli and the presentation of the word *wrong* on the screen. When a programmed consequence was presented after the selection of a comparison, it was displayed in the middle of the screen for 1,000 ms. Termination of the programmed consequences message was followed with a 500-ms intertrial interval. Between trials, the mouse cursor was returned to the center of the screen.

Equivalence class formation The simultaneous protocol was used to establish the equivalence classes. First, all the baseline relations for the equivalence classes were trained in a serialized manner until the achievement of a mastery criterion, and these were then maintained with decreasing proportions of

programmed consequences. All of the baseline relations, together with all emergent relations, were then presented randomly in an emergent-relations test block.

Acquisition of baseline relations All baseline relations were trained in five serialized phases with programmed consequences provided for the selection of comparisons for each trial. Phase 1 was for the training of AB relations in a block containing nine trials, three for each of the three classes. A mastery criterion of at least 90 % correct was required for the training of each relation. When participants satisfied the mastery criterion (selecting the correct comparisons on at least 90 % of the trials in a block), they progressed to the next stage of the experiment. Participants repeated each block until the mastery criterion was met. Phases 2, 3, and 4 were the same as Phase 1, except that the BC, CD, and DE relations were trained in each phase, respectively. An equalization block was then used to ensure that each of the baseline relations was presented the same number of times.

Phase 5 involved the inclusion of all of the relations trained in the first four blocks: AB + BC + CD + DE. The trials presented in the Phase 5 block were A1/B1B2B3, A2/B1B2B3, A3/B1B2B3, B1/C1C2C3, B2/C1C2C3, B3/C1C2C3, C1/D1D2D3, C2/D1D2D3, C3/D1D2D3, D1/E1E2E3, D2/E1E2E3, and D3/E1E2E3. In each trial representation, the first stimulus is the sample and the other three are the comparison stimuli, where the underlined comparison is the correct comparison. Phase 5 contained 36 trials (three presentations of each of the 12 trial types listed above). The block was repeated until correct comparisons were selected on at

least 90 % of the trials of each baseline relation in a block (the mastery criterion).

Maintenance of baseline trials Participants continued with training blocks with reduced programmed consequences. The percentages of trials in a block that produced programmed consequences after the last acquisition-of-baseline-trials phase were reduced to 75 %, 50 %, 25 %, and 0 %, in that order. This cross-phase programmed reduction of reinforcement density was necessary to ensure that the baseline relations would still be intact even without reinforcement. For each level of programmed consequences, the trials that produced programmed consequences were randomized in a block. The maintenance phase was completed once participants had achieved the mastery level of responding in the last block of 36 baseline trials, with no programmed consequences. Table 1 shows a full overview of each of the experimental phases.

Emergent-relations test The last maintenance block, which was administered with no programmed consequences, was followed by an emergent-relations test block that contained 180 trials, consisting of 36 baseline trials, 36 symmetry trials, 54 one-node trials, 36 two-node trials, and 18 three-node trials. All of the trials were randomly presented without programmed consequences.

The emergent-relations test block was divided into two halves of 90 trials each, to assess either the immediate or the delayed emergence of equivalence classes. The first and second sets were referred to as *Test Blocks 1 and 2*, respectively. Equivalence classes were said to be formed if at least 90 % of the trials in a block resulted in the selection of comparisons that were consistent with the experimenter-defined classes. Immediate emergence occurred when this criterion was achieved in the first test block. In those instances, the second test block measured the maintenance of the classes. On the

Table 1 Sequence of training and testing

Experimental Phases	Trial Types	Programmed Consequences	Min # of Trials	Criterion
Acquisition of baseline trials				
Trial types presented in a random order				
1. Serialized trials	A1B1, A2B2, A3B3	100 %	9	9
2. Serialized trials	B1C1, B2C2, B3C3	100 %	9	9
3. Serialized trials	C1D1, C2D2, C3D3	100 %	9	9
4. Serialized trials	D1E1, D2E2, D3E3	100 %	9	9
5. Mixed trials (trials presented randomly)	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	100 %	36	34
Maintenance: Fading of programmed feedback				
6. Mixed trials (trials presented randomly)	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	75 %	36	34
7. Mixed trials (trials presented randomly)	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	50 %	36	34
8. Mixed trials (trials presented randomly)	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	25 %	36	34
9. Mixed trials (trials presented randomly)	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	0 %	36	34
Test for derived relations				
All trial types randomly intermixed				
	Baseline Trials			
	A1B1, A2B2, A3B3, B1C1, B2C2, B3C3 C1D1, C2D2, C3D3, D1E1, D2E2, D3E3	0 %	36	34
	Symmetry Trials			
	B1A1, B2A2, B3A3, C1B1, C2B2, C3B3 D1C1, D2C2, D3C3, E1D1, E2D2, E3D3	0 %	36	34
	1-Node Trials			
	A1C1, A2C2, A3C3, C1A1, C2A2, C3A3, B1D1, B2D2, B3D3, D1B1, D2B2, D3B3, C1E1, C2E2, C3E3, E1C1, E2C2, E3C3	0 %	54	49
	2-Node Trials			
	A1D1, A2D2, A3D3, D1A1, D2A2, D3A3, B1E1, B2E2, B3E3, E1B1, E2B2, E3B3	0 %	36	34
	3-Node Trials			
	A1E1, A2E2, A3E3, E1A1, E2A2, E3A3	0 %	18	17

other hand, delayed emergence of the classes occurred when the criterion was not reached in the first test block but was then reached in the second. A participant failed to form classes if submastery performances were obtained in both test blocks.

Results and discussion

Acquisition of baseline relations

Figure 2 shows the median number of trials needed to acquire the baseline relations in each condition. Median, rather than mean, trials were computed because a few participants in some groups required many more trials than the rest in their group. Acquisition of the baseline relations was fastest for the participants in the C-as-PIC group and slowest for the participants in the E-as-PIC group, but the difference across all of the groups was not statistically significant, Kruskal–Wallis $H(4) = 2.278$, $p = .685$. Thus, the speed of acquiring the baseline relations was not influenced by the ordinal introduction and/or the structural location of the meaningful stimuli in the training structure. In addition, we observed no systematic or significant differences in the acquisition speeds for participants who did and did not subsequently form equivalence classes. Finally, when errors were considered, they occurred on an average of 19 % of trials and did not vary systematically across conditions or between participants who did or did not subsequently form the equivalence classes.

Equivalence class formation

Table 2 shows the performances in emergent-relations Test Blocks 1 and 2, the immediate or delayed emergence of the equivalence classes, and the outcomes of the

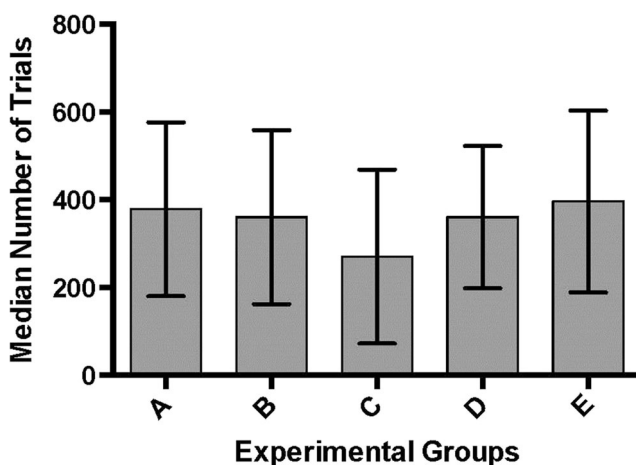


Fig. 2 Median numbers of trials to acquire baseline relations during acquisition of the baseline relations in all groups of Experiment 1. Pictures were used in the different groups, labeled A, B, C, D, and E (i.e., Groups A-as-PIC through E-as-PIC). Error bars indicate interquartile ranges for the medians

sorting tests for each participant in each group. Across all conditions, 27 participants responded substantially below mastery in both test blocks, and thus failed to form equivalence classes. An additional 16 participants responded at the mastery criterion in both test blocks, indicating the immediate emergence of the classes in Test Block 1 and their maintenance in Test Block 2. Another seven participants did not respond at criterion in the first test block but did so in the second test block, which documented the delayed emergence of the equivalence classes. When the delayed emergence of the classes was considered, each group produced essentially the same small increment in yield: an addition of one or two participants per group.

For six of these seven participants (4720, 4725, 4706, 4729, 4716, and 4710) who showed delayed emergence, the first test block produced rather high levels of accuracy that approximated the mastery level. This observation suggests that equivalence classes may have emerged on a delayed basis in the latter portion of the first test and was then maintained in the second test block.

Figure 3 shows how the positioning of the picture stimuli influenced the overall likelihood of equivalence class formation. The overall emergence was defined as the percentage of participants in a group who formed classes, regardless of the rate of emergence. The figure also indicates how participants sorted the stimuli prior to and after exposure to the simultaneous protocol. Each cluster of bars in this panel represents the data for a specific group, with the leftmost bar showing the pre-class-formation sorting, the rightmost representing post-class-formation sorting, and the middle bars representing performance in the emergent-relations test, (i.e., overall emergence of the classes).

Very similar high yields were obtained when the meaningful stimuli served as the A or C member of a class. Similar and somewhat lower yields were obtained when the meaningful stimulus was used as the B or D member in a class, and a very low yield was obtained when the meaningful stimulus served as the E member of the class. The overall yield in the A-as-PIC condition was significantly greater than that obtained in the E-as-PIC condition, $\chi^2(1) = 4.85$, $p = .028$. Likewise, the overall yield in the C-as-PIC condition was significantly greater than that obtained in the E-as-PIC condition, $\chi^2(1) = 5.74$, $p = .017$. In addition, yield declined in a linear manner when the meaningful stimulus served as the C, D, and E member of a class, respectively.

Card sorting

Table 2 and Fig. 3 show the results of the sorting tests. When the sorting performances were analyzed, none of the participants sorted the stimuli into any of the experimenter-defined classes prior to the class formation training. After exposure to

Table 2 Accurate selection of class-indicative comparisons in the test blocks for each participant in each group of Experiment 1

Group	Participant	% Correct in Test Block 1	% Correct in Test Block 2	ECF	Post-ECF Sorting
A-as-PIC	4732	100	100	YES	YES
	4711	100	100	YES	YES
	4720	89	99	YES	YES
	4703	97	98	YES	YES
	4727	94	98	YES	YES
	4741	96	92	YES	YES
	4738	82	88	NO	YES
	4745	76	80	NO	NO
	4715	60	71	NO	NO
	4733	48	48	NO	NO
B-as-PIC	4707	100	99	YES	YES
	4750	98	98	YES	YES
	4706	89	98	YES	YES
	4725	84	96	YES	YES
	4734	78	79	NO	NO
	4748	83	74	NO	YES
	4709	70	64	NO	NO
	4723	60	63	NO	NO
	4743	63	56	NO	NO
	4718	58	42	NO	NO
C-as-PIC	4721	98	100	YES	YES
	4701	98	100	YES	YES
	4716	83	100	YES	YES
	4739	99	99	YES	YES
	4722	97	99	YES	YES
	4729	87	99	YES	YES
	4744	94	97	YES	YES
	4719	79	88	NO	NO
	4749	62	60	NO	NO
	4713	64	49	NO	NO
D-as-PIC	4728	100	100	YES	YES
	4730	99	100	YES	YES
	4735	99	100	YES	YES
	4710	89	99	YES	YES
	4724	83	81	NO	NO
	4702	62	57	NO	NO
	4742	53	56	NO	NO
	4705	64	44	NO	NO
	4747	57	43	NO	NO
	4737	51	38	NO	NO
E-as-PIC	4708	99	100	YES	YES
	4714	64	97	YES	NO
	4726	74	80	NO	NO
	4712	64	66	NO	NO
	4704	71	64	NO	NO
	4717	62	61	NO	NO
4731	68	60	NO	NO	

Table 2 (continued)

Group	Participant	% Correct in Test Block 1	% Correct in Test Block 2	ECF	Post-ECF Sorting
	4736	58	58	NO	NO
	4740	80	52	NO	NO
	4746	53	47	NO	NO

Performances of 90 % correct or more indicate formation of classes and are written in bold in the table

ECF Equivalence class formation

the simultaneous protocol, 24 of the 50 participants across all of the groups sorted the stimuli into the three experimenter-defined classes. The remaining participants sorted the stimuli into clusters that did not correspond to the experimenter-defined classes.

All but one of the participants who formed classes during the emergent-relations tests also showed maintenance of those classes during the sorting test. Only Participant 4714 formed classes in the emergent-relations test but failed to do so in the sorting test. Also, two participants (4748 and 4738) who did not form classes during the emergent-relations test did sort the cards into the experimenter-defined classes afterward. However, their performances in the emergent-relations test approximated mastery level. All other participants who did not form classes also did not sort the cards into the three experimenter-defined classes. Therefore, for 94 % of the participants (47 out of 50), the sorting tasks yielded outcomes that corresponded to their performance in the emergent-relations test.

These results replicate previous findings (e.g., Fields, Arntzen, & Moksness, 2014; Nartey, Arntzen, & Fields, 2015) that showed the generalization of class-based behavior between two trial formats: matching-to-sample trials during class formation, and sorting during post-class-formation testing. The data, thus, show that the inclusion of a meaningful

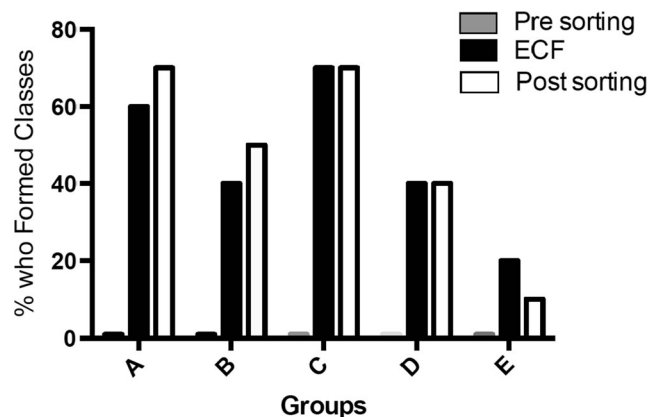


Fig. 3 Percentages of participants who showed class formation in the pre- and post-class-formation sorting tests, as well as those who showed overall equivalence class formation in the derived-relations test in Experiment 1. Pictures were used in the different groups, labeled A, B, C, D, and E

stimulus and its introduction in the training of baseline relations did not influence the seeming agreement in performance between the two test formats.

Summary

The inclusion of a meaningful stimulus in a set of otherwise meaningless stimuli can enhance the formation of an equivalence class by the stimuli in the set. That class-enhancing function, however, is modulated by its order of introduction during the serial training of the baseline relations and/or by its location in the structure of the class.

Experiment 2

The enhancement of class formation by the inclusion of a meaningful stimulus in Experiment 1 could have been influenced by its order of introduction during training and/or by its location in the structure of a class. These two possible determinants of class enhancement could not be isolated in Experiment 1 because they covaried with each other.

In the serial training, relatively high yields were obtained when the meaningful stimuli were introduced early in training, and yields declined when the stimuli were introduced later in training. If the order of introduction was the determinant of class enhancement, the concurrent training of all baseline relations would eliminate any advantage provided by early introduction of the meaningful stimulus. As compared to serialized training, if order was the determinant of class enhancement, the concurrent training of the baseline relations should suppress the class enhancement seen when the meaningful stimulus was introduced early in training to a level that would approximate the low yield observed when the meaningful stimuli were introduced last in training. Alternatively, if concurrent training were to result in the same pattern of yields seen in Experiment 1, this would indicate that order of introduction was not the factor that modulated the enhancing effect of a meaningful stimulus on equivalence class formation.

Experiment 2, then, replicated Experiment 1, with one exception: All of the baseline relations were trained concurrently, instead of serially. Yields were compared across the five conditions in Experiment 2 and were also compared across the two experiments for each of the five conditions: A-as-PIC, B-as-PIC, C-as-PIC, D-as-PIC, and E-as-PIC.

Method

Participants

Fifty students, also from the University of Ghana, 28 males and 22 females between 18 and 25 with an average age of 20 years ($SD = 1.71$), participated in this experiment. All of the

participants were naïve in terms of knowledge of stimulus equivalence research and methodology and were randomly assigned to five groups of ten participants each.

Apparatus

This was the same as in Experiment 1.

Procedure

With the exception that the baseline relations were trained concurrently, everything about the procedure used in Experiment 2 was the same as in Experiment 1. Thus, the AB, BC, CD, and DE trials for all three potential classes were presented randomly in each training block of 36 trials, with the block being repeated until all relations produced the mastery level of responding. Five groups were studied: A-as-PIC, B-as-PIC, C-as-PIC, D-as-PIC, and E-as-PIC. The stimuli used in each class were the same ones used in Experiment 1, shown in Fig. 1.

Results and discussion

Acquisition of baseline relations

Figure 4 shows the median number of trials needed to form the baseline relations in each group in Experiment 2, along with the same information for Experiment 1. When the data for Experiment 2 are considered, acquisition of the baseline relations was quite similar for the participants in all of the groups other than the C-as-PIC group: The participants in the latter group took many more trials to acquire the baseline relations than the other four groups. A Kruskal–Wallis H test showed a statistically significant difference in the speeds of acquisition among the different groups, $H(4) = 13.044$, $p < .05$.

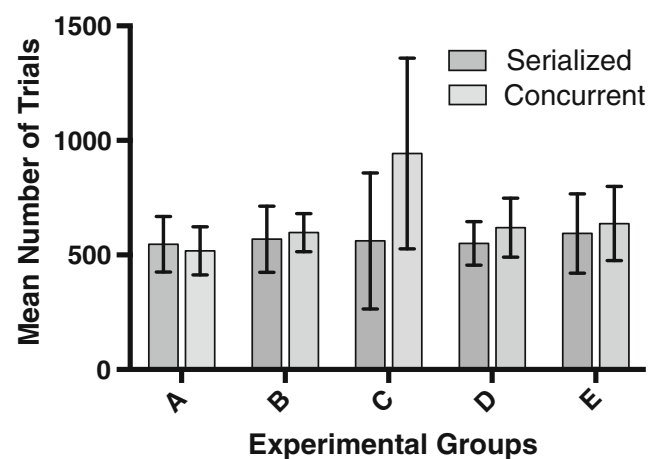


Fig. 4 Comparison of baseline acquisition in Experiment 1 (serialized arrangement) and Experiment 2 (concurrent arrangement). Dark gray bars are for the serialized condition, and light gray bars are for the concurrent condition. Pictures were used in the different groups, labeled A, B, C, D, and E. Error bars indicate standard errors of the means

Figure 4 also provides a comparison of the acquisition speeds when the baseline relations were trained serially or concurrently. The acquisition speed was significantly faster during serialized than during concurrent training, $U = 900$, $p = .015$, $r = .24$. Across experiments, similar numbers of trials were needed to acquire the baseline relations in the A, B, D, and E groups. In the C group, however, acquisition was much slower during concurrent training than serial training.

When the absolute numbers of errors were considered across experiments, significantly fewer errors were needed to acquire the baseline relations during serialized training than during concurrent training, $U = 210.5$, $p < .001$, $r = .72$. On a proportional basis, 19 % and 42 % of the acquisition trials produced errors across all groups in Experiments 1 and 2, respectively.

Equivalence class formation

Figure 5 depicts the overall yields obtained in each condition of Experiments 1 and 2. After the concurrent training of baseline relations, none of the participants formed classes in the B-, D-, or E-as-PIC groups; only 10 % of the participants formed classes in the A-as-PIC group; and 20 % formed classes in the C-as-PIC group. An overall statistical analysis showed no significant differences among the yields in the five groups in Experiment 2, $\chi^2(4) = 3.804$, $p = .433$.

When the yields obtained across experiments were compared, a significantly larger percentage of participants formed classes when baseline relations were established in a serialized order than when they were established concurrently, $\chi^2(1) = 18.32$, $p < .0001$. For each condition in the two experiments, the yields were significantly greater after serial training than after concurrent training in the A-as-PIC condition [$\chi^2(1) =$

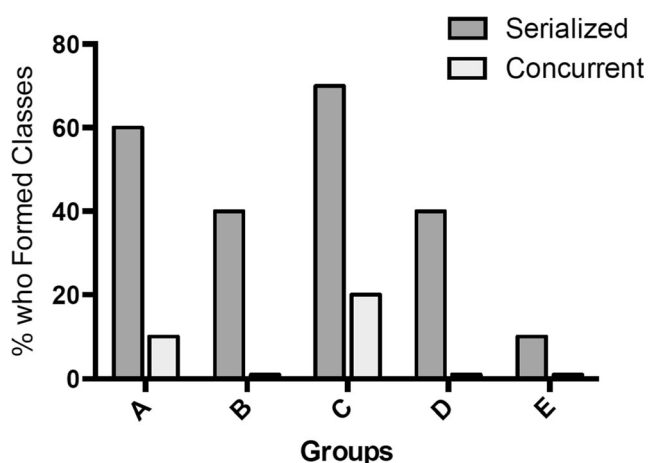


Fig. 5 Yields for each condition during emergent-relations testing in Experiments 1 and 2. Pictures were used in the different groups, labeled A, B, C, D, and E. Dark gray bars are for the serialized condition, and light gray bars are for the concurrent condition

5.495 , $p = .019$], the C-as-PIC condition [$\chi^2(1) = 5.015$, $p = .025$], and the D-as-PIC condition [$\chi^2(1) = 5.00$, $p = .025$]. In contrast, no significant difference in yields was found in the B-as-PIC condition [$\chi^2(1) = 2.400$, $p = .121$] or the E-as-PIC condition [$\chi^2(1) = 1.053$, $p = .305$]. From visual inspection, the pattern of yields across conditions was the same as in Experiment 1, but with greatly attenuated levels in the A, B, C, and D conditions in Experiment 2, and the same very low levels in the E condition.

When compared to the yields observed in Experiment 1, the concurrent training of baseline relations in four of the five conditions (A, B, C, and D) suppressed the class-enhancing effect of the inclusion of a meaningful stimulus as a member of a to-be-formed equivalence class. Also, when concurrent training was conducted, the likelihood of class formation was not influenced by the location of the meaningful stimulus in the structure of the class. Thus, the class-enhancing effect of a meaningful stimulus was substantially modulated by its temporal order of introduction during training. Specifically, the introduction of a meaningful stimulus in the first or the middle trained baseline relation in a series produced a maximal enhancement of equivalence class formation. In contrast, the likelihood of class formation was not influenced by the static location of the meaningful stimulus in the nodal structure of an equivalence class.

General discussion

Modulation of class enhancement by meaningful stimuli

In Experiment 1, relatively high and similar percentages of participants formed classes when a meaningful stimulus was used as the A or the C member in a class that had an A→B→C→D→E training structure. Lower and equal yields were produced when the meaningful stimulus was used as the B or the D member in the class, and a very low percentage of participants formed classes when the meaningful stimulus was the E member of the class. Previously, separate articles have reported the enhancement effects produced by some of these conditions: meaningful stimuli as C stimuli (Arntzen, Narthey, & Fields, 2014, 2015; Fields et al., 2012; Narthey et al., 2015; Travis et al., 2014) and meaningful stimuli as A and E stimuli (Narthey et al., 2014). In the present study, we explored the effects of all of the conditions in a single experiment and replicated the findings obtained in the prior experiments. In addition, the present experiment allowed us to explore the effect of including a meaningful stimulus as the B or the D member of a class on the enhancement of class formation; indeed, the yields obtained under these conditions could not have been predicted from the results of the previously referenced research. Thus, the present experiments provide a

comprehensive measure of the effects of including a meaningful stimulus on the enhancement of equivalence class formation.

The results of Experiment 2 showed the static location of a meaningful stimulus in a training structure did not influence the likelihood of equivalence class formation. As was seen in Experiment 1, the temporal point of introducing the meaningful stimulus in training had rather sizable effects on class enhancement. The highest yields were obtained when the meaningful stimulus was included in the first trained baseline relation (A in AB) and when it was used as a member of the middle trained BC and CD relations. The following are some factors that might and might not be responsible for each of these class-enhancing effects.

When it was included in the first trained relation, the meaningful stimulus served many functions: It was a “single” (a stimulus linked by training to only one other stimulus in a set, as was noted by Fields & Verhave, 1987), a sample stimulus, the first stimulus to be introduced in training, and a member of the relation that defined one of the ends of the training structure: AB in an $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$ class. The class-enhancing potential of each of these factors will be considered next.

If its functionality as a single was the determinant of class enhancement, the use of a meaningful stimulus as a single in another location in the class should have had the same effect. The only other stimulus in the class to serve as a single was the E stimulus, which was the last presented stimulus in training. When the meaningful stimulus was used as the E member of the classes, the likelihood of class formation was very low. Therefore, the use of a meaningful stimulus as a single, per se, did not account for its class enhancement effect when it was used as the A stimulus in the class.

When the meaningful picture was used as the A member of the class, it also functioned as a sample stimulus in the AB relation and was correlated with a high yield. If the sample functionality of the single was a critical class-enhancing factor, another single that did not serve the sample function should be correlated with a lower likelihood of equivalence class formation. In our preparation, the only other single was the E stimuli, which served as a comparison in the last-trained, DE relation. That stimulus was correlated with a very low yield. Thus, it is possible that the sample function of the single, A stimulus was a possible factor that was a determinant of class enhancement. The meaningful stimulus used as the sample in the AB relation was also the first stimulus presented in training, and by dint of its meaningfulness, it served many connotative and denotative functions. Thus, it is safe to assume that it was much more salient than any of the other stimuli used in training. When a stimulus has these properties, they transfer to other stimuli that are in the same class (e.g., Bortoloti, Pimentel, & de Rose, 2014). In the present circumstance, when the meaningful stimulus was used as the first stimulus in the first-trained baseline relation, its properties

could generalize or transfer to the B stimulus, the other element in the first-trained AB relation, which could then transfer to C during BC training, thence to D during CD training, and finally to E during DE training. This enriched network of connotative and denotative relations by the other abstract stimuli used in training would not have occurred if a meaningful stimulus had been introduced later in training, and certainly not when it was introduced last in training. This may be the way in which the introduction of the meaningful stimulus as the sample in the first-trained relation could have enhanced the likelihood of forming the equivalence classes. Indeed, the proposed transfer of connotative properties could be evaluated by conducting semantic differentials at the completion of the training of each baseline conditional relation.

Finally, the meaningful A stimulus was a member of the relation that defined the starting point in the class structure: A in $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$. Logically, the meaningful stimulus could be used in the first-trained relation even if was not AB. Specifically, the first-trained relation could be BC, CD, or DE, and a meaningful stimulus could be used as one of the members in each of these relations. If being used first is critical, all should produce the same high yields. If being the lead stimulus in the class structure (i.e., A in $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$) is vital, the yields should differ from the first-trained relation. Thus, the highest yield should occur when the meaningful stimulus was A, with decreasing yields when the meaningful stimulus was used first in BC training, CD training, and DE training. Clearly, this is a speculation that is subject to disconfirmation in new research.

During Experiment 1, a higher yield was obtained when the meaningful stimulus served as the C member of the equivalence class than when it served as the B or E member of the classes (70 % vs. 40 %). All three stimuli served as nodes during training: The C stimulus was linked by training to B and D, the B stimulus was linked by training to A and C, and the D stimulus was linked by training to C and E. If serving as a nodal stimulus was the determinant of class enhancement, using the meaningful picture as the C, B, or D stimulus should have produced equivalent levels of class enhancement. In fact, using the meaningful picture as the C stimulus produced greater enhancement than when it was used as the B or D stimulus. Therefore, having a meaningful stimulus as a node, per se, was not predictive of the degree of class enhancement.

On the other hand, class enhancement was correlated with the types of nodes that served as the B, D, and C stimuli. Specifically, the C stimulus was a node that was linked by training to two other nodal stimuli. In contrast, the B and D stimuli were nodes that were linked by training to one nodal stimulus and one single. Therefore, the likelihood of class formation was correlated with the types of stimuli that were linked to the meaningful stimulus when it acted as a node. The efficacy of this factor could be evaluated using large multi-modal classes that would permit the manipulation of the types

of stimuli that were linked to the nodes in the training structure.

The information presented above describes how the structural parameters of an equivalence class may modulate the class-enhancing effects of a meaningful stimulus that is included as a member of the potential class. It does not address a mechanism by which the inclusion of such a stimulus enhances class formation. One possibility is that the inclusion of the meaningful stimulus results in class expansion rather than class formation (Fields et al., 2012; Travis et al., 2014). Specifically, assuming that a meaningful stimulus is a member of a number of different categories, it is related to a very wide network of other stimuli. Thus, when it is embedded with other meaningless stimuli, when they become class members they also become related to all of the stimuli that are related to the meaningful stimulus. As such, when a set of meaningless stimuli and one meaningful stimulus become members of an equivalence class, we may be witnessing the expansion of an existing class rather than the *de novo* formation of a new and isolated class. This process, however, did not result in the formation of new classes when the meaningful stimulus was introduced at different times during training or when all relations were trained concurrently. Thus, even if this is the mechanism of class enhancement by a meaningful stimulus, it is still subject to modulation by the protocol used for the training of the baseline relations and by the position of the meaningful stimulus in the training structure of a potential class.

Absence of a null condition

The present experiments did not include null groups in which all of the potential class members were abstract stimuli. On the other hand, the present experiment involved the inclusion of a meaningful stimulus in each location in a training structure, it was possible to evaluate effects on the likelihood of class formation. However, as we noted in the introduction, many other experiments have included a reference condition that did not have a meaningful stimulus as a class member, which would be a null condition. All of them produced yields in a narrow range that varied from 0 % and 15 %. These outcomes were similar to those reported in the E condition of Experiment 1 and to virtually all of the conditions in Experiment 2. Given these similarities, it is reasonable to assume that in Experiment 1, the inclusion of a meaningful stimulus as the A, B, C, or D member of a class produced absolute enhancements of class formation. In addition, it is reasonable to conclude that the inclusion of the meaningful stimulus as the E member of the class did not enhance class formation.

Implications for applications

The results of the present experiments have a number of implications for the establishment of equivalence classes that are of real-world import in applied settings. First, the mode of training and testing that occurs in many standard educational settings can be characterized as “train all and test all”—a procedure not unlike the simultaneous protocol, which includes the concurrent training of all baseline relations for equivalence classes. The first recommendation that flows from the results of the present experiments is that if training and testing are to be conducted using the simultaneous protocol, training of the baseline relations should be done in a serialized instead of a concurrent manner (see Amtzen, Halstadro, Bjerke, Wittner, & Kristiansen, (2014)). Second, if real-world equivalence classes are to be established through the inclusion of a meaningful stimulus to facilitate class formation, it should be included as the sample stimulus in the first-trained relation or in the relations that are trained in the middle of the sequence, so that it functions as a node that is linked to two nodal stimuli (e.g., C in BC and CD). Furthermore, it should never be used in the last-trained relation. Following these guidelines should optimize the establishment of equivalence classes that are of real-world significance when the classes are established using pedagogical procedures such as the simultaneous protocol.

Conclusion

The results of these experiments replicated and extended the finding that the inclusion of a meaningful stimulus in a set of meaningless stimuli can enhance the formation of equivalence classes. However, the absolute enhancement property of a meaningful stimulus operates not only through its mere inclusion in the class structure. Rather, class enhancement by the inclusion of a meaningful stimulus is modulated by the order of introduction of such a stimulus during the training of the baseline relations. It is probably commonly assumed that the inclusion of a positively valenced meaningful stimulus should enhance the likelihood of equivalence class formation. The results of the present experiments show that that assumption is unwarranted. Rather, the type of protocol used for training and testing and the nodal structure of the to-be-formed class both modulate the class-enhancing effects of a meaningful stimulus, and in some cases essentially neutralize that presumed effect. The information provided in the present experiments then might also account for presumed failures of class enhancement when meaningful stimuli have been included in sets of new stimuli that were supposed to become members of equivalence classes.

Author note The three authors have no conflicts of interest to declare.

References

- Amtzen, E. (2004). Probability of equivalence formation: Familiar stimuli and training sequence. *Psychological Record*, *54*, 275–291. Retrieved from <http://opensiuc.lib.siu.edu/tpr/vol54/iss2/7/>
- Amtzen, E. (2012). Training and testing parameters in formation of stimulus equivalence: Methodological issues. *European Journal of Behavior Analysis*, *13*, 123–135. Retrieved from www.ejoba.org/cgi-bin/search.cgi?volume=13&issue=1
- Amtzen, E., & Hansen, S. (2011). Training structures and the formation of equivalence classes. *European Journal of Behavior Analysis*, *12*, 483–503. Retrieved from <http://www.ejoba.org/cgi-bin/search.cgi?volume=12&issue=2>
- Amtzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record*, *47*, 309–320. Retrieved from <http://the-psychological-record.siu.edu/index.html>
- Amtzen, E., & Lian, T. (2010). Trained and derived relations with pictures versus abstract stimuli as nodes. *Psychological Record*, *60*, 659–677. Retrieved from <http://opensiuc.lib.siu.edu/tpr/vol60/iss4/8/>
- Amtzen, E., Grondahl, T., & Eilifsen, C. (2010). The effects of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *Psychological Record*, *60*, 437–462. Retrieved from <http://opensiuc.lib.siu.edu/tpr/vol60/iss3/4/>
- Amtzen, E., Halstadro, L.-B., Bjerke, E., & Halstadro, M. (2010). Training and testing music skills in a boy with autism using a matching-to-sample format. *Behavioral Interventions*, *25*, 129–143. doi:10.1002/Bin.301
- Amtzen, E., Halstadro, L. B., Bjerke, E., Wittner, K. J., & Kristiansen, A. (2014). On the sequential and concurrent presentation of trials establishing prerequisites for emergent relations. *The Behavior Analyst Today*, *14*, 23–30.
- Amtzen, E., Nartey, R. K., & Fields, L. (2014). Identity and delay functions of meaningful stimuli and enhanced equivalence class formation. *Psychological Record*, *64*, 349–360. doi:10.1007/s40732-014-0066-3
- Amtzen, E., Nartey, R. K., & Fields, L. (2015). Enhanced equivalence class formation by the delay and relational functions of meaningful stimuli. *Journal of the Experimental Analysis of Behavior*, *103*, 524–541. doi:10.1002/jeab.152
- Belanich, J., & Fields, L. (1999). Tactual equivalence class formation and tactual-to-visual cross-modal transfer. *Psychological Record*, *49*, 75–91. Retrieved from <http://opensiuc.lib.siu.edu/tpr/vol49/iss1/6/>
- 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
- Buffington, D. M., Fields, L., & Adams, B. J. (1997). Enhancing equivalence class formation by pretraining of other equivalence classes. *Psychological Record*, *47*, 69–96. Retrieved from <http://opensiuc.lib.siu.edu/tpr/vol47/iss1/5/>
- 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
- Fields, L., Amtzen, 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., Amtzen, E., & Moksness, M. (2014). Stimulus sorting: A quick and sensitive index of equivalence class formation. *Psychological Record*, *64*, 487–498. doi:10.1007/s40732-014-0034-y
- 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. *Psychological Record*, *49*, 703–724.
- Fields, L., Reeve, K. F., Rosen, D., Varelas, A., Adams, B. J., Belanich, J., & Hobbie, S. A. (1997). Using the simultaneous protocol to study equivalence class formation: The facilitating effects of nodal number and size of previously established equivalence classes. *Journal of the Experimental Analysis of Behavior*, *67*, 367–389. doi:10.1901/jeab.1997.67-367
- Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229–262). New York, NY: Plenum Press.
- Hayes, L. J., Tilley, K. L., & Hayes, S. C. (1988). Extending equivalence and membership to gustatory stimuli. *Psychological Record*, *38*, 473–482.
- 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
- LeBlanc, L. A., Miguel, C. F., Cummings, A. R., Goldsmith, T. R., & Carr, J. E. (2003). The effects of three stimulus-equivalence testing conditions on emergent us geography relations of children diagnosed with autism. *Behavioral Interventions*, *18*, 279–289. doi:10.1002/bin.144
- Nartey, R. K., Amtzen, E., & Fields, L. (2014). Two discriminative functions of meaningful stimuli that enhance equivalence class formation. *Psychological Record*, *64*, 777–789. doi:10.1007/s40732-014-0072-5
- Nartey, R. K., Amtzen, E., & Fields, L. (2015). Enhancement of equivalence class formation by pretraining discriminative functions. *Learning & Behavior*, *43*, 20–31. doi:10.3758/s13420-014-0158-6
- Nedelcu, R. I., Fields, L., & Amtzen, 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. (1995). Reversal of baseline relations and stimulus equivalence: I. Adults. *Journal of the Experimental Analysis of Behavior*, *63*, 225–238. doi:10.1901/jeab.1995.63-225
- Pilgrim, C., Chambers, L., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: II. Children. *Journal of the Experimental Analysis of Behavior*, *63*, 225–238. doi:10.1901/jeab.1995.63-239
- 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
- Travis, R. W., Fields, L., & Amtzen, 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