International Journal of Psychology and Psychological Therapy, 2017, 17, 3, 291-303

Effects of Mixed Training Structures on Equivalence Class Formation

Joaquín Menéndez*

Universidad de Buenos Aires & Instituto de Biología y Medicina Experimental, Argentina

Federico José Sánchez, Matías Alejandro Avellaneda, Sebastián Ariel Idesis

Universidad de Buenos Aires, Argentina

Alberto Andrés Iorio

Universidad de Buenos Aires & Instituto de Biología y Medicina Experimental, Argentina

Abstract

The experimental literature reports differences in performance when participants are tested for the emergence of derived relations after stimulus equivalences class training, depending on which training structured is used. Comparison-as-node and sample-as-node structures have shown to be more effective in producing the emergence of derived relations than linear series, with inconclusive results about which of the first two structures is more effective. Intertrial correspondence was manipulated between the stimuli via the use of mixed training structures. 48 participants were divided in four groups: the first received equivalence-class training using a sample-as-node structure, the second following a comparison-as-node structure, and the other two following a mixed structure with the same nodal density of the central node as the first two. The four groups were taught two five-member equivalence classes with a nodal density of four. Both during training and testing, the performances were higher for the sample-as-node and the comparison-as-node structures, compared to the other two structures. Results are discussed from the lens of hypotheses based on simple-discriminations learning and the role of samples and comparisons.
Key words: equivalence relations, training structures, nodal distance, matching-to-sample.

How to cite this paper: Menéndez J, Sánchez FJ, Avellaneda MA, Idesis SA & Iorio AA (2017) Effects of Mixed Training Structures on Equivalence Class Formation. *International Journal of Psychology & Psychological Therapy, 17,* 291-303.

Novelty and Significance

What is already known about the topic?

- There is evidence about the influence of different training structures on learning and emergence of stimulus equivalence relationships.
- Three training structures are reported in the literature: Linear series, comparison-as-node and sample-as-node. In addition, the performance of the linear series structure decreases markedly with the increase of class members in comparison to the others structures.
- The lower performance is explained by a nodal distance effect. This phenomenon is only possible in linear training structures due to its configuration in the presentation of the stimuli.

What this paper adds?

- Provides a procedure to compare different training structures maintaining the number of nodal density.
- The learning and emergence of equivalence relations would be influenced by the function of the stimuli as a sample and/or comparison during training and not by the simple discriminations learned during training neither by the nodal density of the classes.

It has been argued that the stimulus-equivalence-class (SEC) paradigm provides the theoretical and experimental bases for the study of semantic and symbolic processes, including natural language (Sidman, 1986, 1994). As Fields and Verhave (1987) propose, a picture of a peach, the spoken word "peach", the written word "peach", and the

* Correspondence concerning this article: Joaquín Menéndez, Facultad de Psicología, Universidad de Buenos Aires, Hipólito Yrigoyen 3242, C1207ABR Ciudad Autónoma de Buenos Aires, Argentina. Email: joaquin14@gmail.com. *Acknowledgements*: The present experiment was supported by the University of Buenos Aires (UBACyT-Code: 20020130100861BA). The authors would like to thank Mateo Santillán and Nicolás Rodríguez Arriaga for their useful comments on the manuscript.

smell of a peach are stimuli that are not related in terms of physical properties. With exposure to the proper contingencies, these stimuli become interrelated and form an equivalence class. Given that the stimuli in the class share no physical properties, the relation among them is probably the result of training, both in the laboratory and in the natural environment.

We say that a group of stimuli belong to a SEC when, after training certain conditional relations among them (e.g., in presence of a specific stimulus, usually called A1, the choice of stimulus B1 is reinforced, and in presence of stimulus A2, the choice of B2 is reinforced), other, untrained relations are found to have emerged (Sidman, 1994). Traditionally, this training was carried out employing the procedure known as matching to sample (Sidman & Tailby, 1982), in which participants are exposed to a sample stimulus (e.g., A1), and are given the option to choose one of two or more comparison stimuli (e.g., B1 and B2). In this way, participants learn the conditional relation A1-B1, and other relations can be taught following the same procedure (e.g., A2-B2, B1-C1 and B2-C2). These stimuli share no physical relation; thus, the correct comparison stimulus in the presence of a certain sample is completely arbitrary and decided beforehand by the experimenter. Participants must learn the conditional relations through exposure to repeated trials of this produce, during which their choices are followed by feedback messages indicating whether their choice was correct or not.

After training resulting in at least two sets of three stimuli each (e.g., A1-B1-C1 and A2-B2-C2), a variation of this procedure without feedback is used to test for the emergence of relations that had not been directly trained, namely, reflexivity (in presence of a certain sample stimulus, choosing the comparison stimulus that is identical; e.g., in presence of sample stimulus A1, choosing comparison stimulus A1), symmetry (in presence of a sample stimulus that previously functioned as comparison, choosing the comparison stimulus which previously functioned as its sample; e.g., choosing comparison stimulus A1 in presence of sample stimulus B1), transitivity (in presence of a sample stimulus, choosing the comparison stimulus that belongs to the same class but was not related directly to it, but indirectly through a third stimulus; e.g., choosing comparison stimulus C1 in presence of sample stimulus A1) and a combination of symmetry and transitivity, also called equivalence relation (e.g., in presence of sample stimulus C1, choosing sample stimulus A1). The typical result is that verbal humans show the emergence of these relations, which are referred to as emergent relations.

Within the SEC paradigm, nodal stimuli (i.e., stimuli which are related to two or more other stimuli) mediating the transitivity and equivalence relations can have many singles related to them during training. Singles are stimuli that, through training, are related to only one stimulus. The number of singles related to a node is called nodal density, and is determined by the distribution of the singles (Fields & Verhave, 1987).

The literature reports differences in the performance of participants in testing for the emergence of derived relations after conditional discrimination training, depending on the order in which baseline relations are trained. The different possible orders receive the name of training structures, and manipulating them has proven to have differential effects on the outcome of the training. Namely, comparison-as-node (CaN) and sample-as-node (SaN) structures have proven to be more effective for the emergence of derived relations than linear series (LS), especially as class size increases, but no difference has been found consistently between CaN and SaN (Arntzen & Holth, 1997; Arntzen, Grondahl, & Eilifsen, 2010; Plazas & Villamil, 2016; Sánchez, Menéndez, Avellaneda, Idesis, & Iorio, 2016). This differential effect between LS and the other structures has been studied within the nodal distance effect; increasing the nodal distance results in a decrease in performance (Bentall, Jones, & Dickins, 1999). Since nodal distance is increased with class-size increase in the linear-series structure only, it could constitute an explanation as to why SaN and CaN structures tend to result in better performances. Nonetheless, differences are also found when nodal distance is the same for different structures (Kinloch, McEwan, & Foster, 2013). Fields, Hobbie, Adams, and Reeve (1999) found that there were no differences between SaN and CaN structures with 5-member classes, although CaN proved to be superior with 7-member classes. The LS structure was not tested in this work. Similarly, Drake and Saunders (1987, cited in Saunders & Green, 1999) found that the CaN structure was more effective for the formation of 4-member and 5-member classes than the LS structure. Testing the differences between structures containing more than three stimuli would prove useful by extending the literature related to big class comparisons.

Nonetheless, there has not been adequate control of nodal density among the different structures when using bigger classes (Arntzen & Holth, 2000). Some authors (Fields & Verhave, 1987) have suggested that nodal density could either facilitate or hinder acquisition. Given that CaN and SaN structures result in better performances, it could be expected that increasing nodal density would facilitate acquisition of trained relations and the emergence of derived relations in a training structure similar to LS. If this is not the case, differences in performance must be due to other factors. This task proves to have limitations, however, since incrementing the amount of stimuli in the LS structure increases nodal distance, and therefore it is hard to make comparison among classes with more than three stimuli.

Another possible explanation of the differences among structures may be the intertrial correspondence, which depends on the consistency of stimuli in their roles of sample during the baseline-relations learning task. Worsham (1975) and Roberts (1980) pointed out that proactive interference also occurs when the incorrect comparison in the present trial appeared as a sample in the previous trial. In other words, if a stimulus functions as a comparison in one trial and as a sample in the next (or *vice versa*), the accuracy of the responses decreases, as opposed to what happens when the stimulus maintains its sample or comparison function in both trials. This could explain why performances are inferior in LS when compared to other structures. Both in SaN and CaN, sample stimuli do not fluctuate in comparison to previous trials, contrary to what happens in the LS structure. That is, the more the sample changes from trial to trial, the worse the performance. However, it is important to note that different training protocols (e.g., simple to complex vs. simultaneous) would affect the degree of intertrial correspondence. If training occurs with a simultaneous protocol, such as that used by Fields *et alii* (1999), all possible training relations are presented in the same block of trials. This would decrease inter-trial correspondence of the sample. If training occurs with a simple-to-complex or sequential protocol, however, specific relations are trained to criterion before moving on to training of additional relations. In this case, the degree of intertrial correspondence now does become affected by the training structure used. The degree of intertrial correspondence would also be affected by the number of classes being trained. That is, if we are training the AB relation for three classes, then the intertrial correspondence of the sample would be reduced relative to training two classes.

The objective of this study was to find out what the influence of intertrial correspondence of sample stimuli is in the formation of equivalence-class formation through a matching-to-sample procedure. It was also our goal to find out whether matching the nodal density of the central node and the class size among the different groups, while also changing the intertrial correspondence among them, would eliminate the differences among structures. If Fields and Verhave (1987) are correct, the training structures that have the same nodal distance and density, should demonstrate no differences between them. In contrast, Saunders and Green (1999) propose that the difference in performance of the different structures during testing is due to differences in the amount of simple discriminations learned during training, more precisely simultaneous and successive discriminations. To achieve this, this study will examine the relations of transitivity and equivalence only. If Saunders and Green's (1999) hypothesis happened to be correct the training structures in which the required simple discriminations are presented during training should perform better, those structures being CaN and Mixed SaN-CaN. If, on the contrary, structures with less variability in the roles of the stimuli result in better performances, the hypothesis of the roles of the stimuli being determinant of the performance would receive support.

This study was devised in order to comprehend and identify whether the differences in structures are due to the differences in nodal density or to their roles as sample and comparison. To achieve this, two equivalence classes of five members each were trained. Participants were divided into four groups, following a different training structure in each one. The first performed an SEC training task following an SaN structure, the second following a CaN structure, and the other two following a mixed structure but matching the nodal density of the central node with the former two groups. Our hypothesis is that the structures with higher trial variation during training, that is, those in which sample and comparison stimuli were more variable in their sample-comparison function, will result in a worse performance (lower amount of correct responses) in later testing for the emergence of derived relations than structures with lower trial variation during training.

Method

Participants

Participants ranging from 18 to 40 years old. All of them were undergraduate psychology students of the Universidad de Buenos Aires (Argentina), and signed a statement of informed consent. They were properly informed about the goals and characteristics of the investigation. The following exclusion criteria were used: 1) history of sensory-motor or neuropsychiatric disorders, 2) consumption of psychiatric medication, 3) prior knowledge of the experimental paradigms employed. To assess the presence of any of these exclusion criteria, we relied on a self-administered questionnaire. 48 participants participated in the experiment and were divided randomly into four groups: SaN (seven women, five men, age 21.7 ±1.76), CaN (11 women, one man, age 22.8 ±3.48), mixed-LS (seven women, five men, age 22.2 ±2.39) and mixed SaN-CaN (eight women, four men, age 27.4 ±3.14). Participants signed a note of informed consent for their participation in the experiment. National and international recommendations for research with human beings were strictly followed (American Psychological Association, 2002).

Equipment

The study was carried out in a soundproof, windowless room. Each subject sat in front of a table on which sat a PC with an Intel ® Core (™) 2 Duo CPU E4700 2,6 GHz. Computerized tasks programmed in Python language were used. Task instructions were provided by successive messages presented on the PC screen, and verbally before beginning each task.

Procedure and Tasks

The protocol was authorized by the Ethics Committee from the *Instituto de Biología y Medicina Experimental* (Buenos Aires, Argentina). Participants signed a note of informed consent for their participation in the experiment. The national and international recommendations for human experimentation were strictly followed (APA, 2002).

First group received training following an SaN structure, the second following a CaN structure, the third following a structure similar to LS but matching the nodal density of the central node to the previous two groups (mixed-LS), and the fourth following a mixed structure combining aspects of SaN and CaN (mixed SaN-CaN, Figure 1). Nodal density and nodal distance were matched in the four groups based on what Fields and Verhave (1987) proposed.

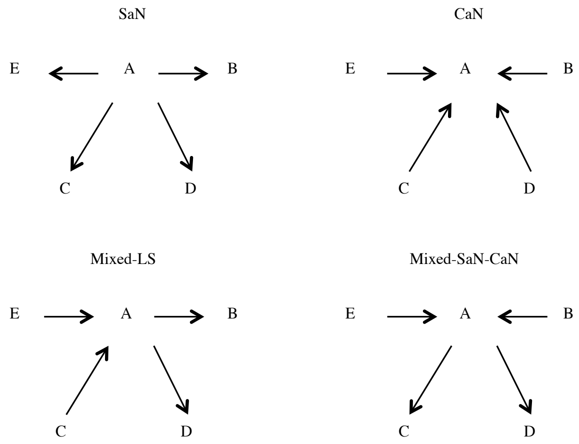

Figure 1. Diagram of the training structures for each group.

SEC Training. The stimuli used were pronounceable bisyllabic pseudowords, which possess no perceptual similarity nor previous semantic relations (Aguado Alonso, 2005). White letters over a black background were employed. The stimuli used are shown in Table 1. Participants were given the following instruction (in Spanish): "The task consists of learning associations between stimuli.

1) A stimulus will be presented at the top of the screen at the beginning of each trial. When you click on it, three stimuli will appear in the lower part of the screen.
2) Click on ONE of those three stimuli.
3) After each choice, you will be informed whether your choice was CORRECT or INCORRECT.

To continue, click on the center of the screen."

Each participant was instructed to select, using the PC mouse, the comparison stimulus corresponding to the sample stimulus (matching to sample). The stimuli were presented on the PC screen. In each trial, a sample stimulus appeared in the center of the screen.

Table 1. Pseudowords used for the formation of two five-member equivalence classes.

Stimulus	Class 1	Class 2
A	LAFU	COTE
B	TULE	MIDU
C	ESGA	FAPE
D	DOLA	SUNA
E	GABE	ROCU

After the subject clicked on the sample stimulus, it disappeared and immediately two comparison stimuli, in the lower area, left and right, were presented. The stimuli remained in the screen until the subject chose a comparison stimulus. During the training phase, four series of arbitrary relations among pseudowords were trained. During this phase only, the participants' responses were followed by "CORRECT" or "INCORRECT" messages written in the center of the screen, depending on whether their choice coincided with the relation arbitrarily established by the researchers. This message was presented during one second. The inter-trial was of one second. Trials were grouped in four simple sequential blocks: AE, AB, AC, and AD in the case of SaN; EA, BA, CA, and DA in the case of CaN; AB, CA, AD and EA in the case of mixed-LS; and EA, BA, AC and AD in the case of mixed SaN-CaN (Table 2). Every block consisted of 18 trials, during which each relation (baseline relation) was presented 9 times. These four blocks were followed by a fifth combined block, which contained trials of the four types. This block consisted in 32 trials during which all baseline relations were presented four times each. To avoid effects of order of trial presentation, all trials were counterbalanced. If the percentage of correct choice was lower than 90% (29 correct responses) in the fifth block, the block was restarted automatically (three times maximum). If the subject did not reach the learning criterion, the experiment would end and their data would not be used for the analysis of derived relations. The amount of simple discriminations seen in each learning structure is shown in Table 3.

Emergent Relations Testing. During the testing phase, two five-member equivalence classes were tested (A1-B1-C1-D1-E1, A2-B2-C2-D2-E2) through a task similar to the one used in training, but without feedback messages and persistent sample (zero-delayed matching to sample). Sample and comparison stimuli related by symmetry, transitivity and combined symmetry and transitivity (equivalence) were presented (see Table 2 for a more detailed explanation). The learning criterion was 87% correct choices (Sánchez *et alii*, 2016). This phase consisted of a single block of 48 trials without feedback.

Data analysis

Analyses were performed with the SPSS 15 statistical package. The alpha level of significance was .05 for every test. One-way ANOVAs was used for the percentage of correct responses in the total sum of the combined block during training among the four training structures and another one-way ANOVA for the amount of correct responses in the last training block, also among the four training structures. Tukey's HSD post hoc tests were used. A Kruskal-Wallis H test was used to test the differences in the amount of training blocks needed to reach the criterion. Mann-Whitney U tests were used to test differences among the groups. A one-way ANOVA was used for the variable percentage of correct responses during the testing phase among the different training structures. Tukey's HSD *post hoc* tests were used.

Table 2. Stimuli used in each of the two five-member classes.

Relation	Training structure							
	SaN		CaN		Mixed-LS		Mixed-SaN-CaN	
	Samples	Comparisions	Samples	Comparisions	Sample	Comparisions	Samples	Comparisions
Baseline	A1	B1* B2	B1	A1* A2	A1	B1* B2	E1	A1* A2
	A2	B1 B2*	B2	A1 A2*	A2	B1 B2*	E2	A1 A2*
	A1	C1* C2	C1	A1* A2	C1	A1* A2	B1	A1* A2
	A2	C1 C2*	C2	A1 A2*	C2	A1 A2*	B2	A1 A2*
	A1	D1* D2	D1	A1* A2	A1	D1* D2	A1	C1* C2
	A2	D1 D2*	D2	A1 A2*	A2	D1 D2*	A2	C1 C2*
	A1	E1* E2	E1	A1* A2	E1	A1* A2	A1	D1* D2
	A2	E1 E2*	E2	A1 A2*	E2	A1 A2*	A2	D1 D2*
Symmetry	B1	A1* A2	A1	B1* B2	B1	A1* A2	A1	E1* E2
	B2	A1 A2*	A2	B1 B2*	B2	A1 A2*	A2	E1 E2*
	C1	A1* A2	A1	C1* C2	A1	C1* C2	A1	B1* B2
	C2	A1 A2*	A2	C1 C2*	A2	C1 C2*	A2	B1 B2*
	D1	A1* A2	A1	D1* D2	D1	A1* A2	C1	A1* A2
	D2	A1 A2*	A2	D1 D2*	D2	A1 A2*	C2	A1 A2*
	E1	A1* A2	A1	E1* E2	A1	E1* E2	D1	A1* A2
	E2	A1 A2*	A2	E1 E2*	A2	E1 E2*	D2	A1 A2*
Equivalence	E1	B1* B2	E1	B1* B2	E1	B1* B2	E1	B1* B2
	E2	B1 B2*	E2	B1 B2*	E2	B1 B2*	E2	B1 B2*
	B1	E1* E2	B1	E1* E2	B1	E1* E2	B1	E1* E2
	B2	E1 E2*	B2	E1 E2*	B2	E1 E2*	B2	E1 E2*

Note: Asterisks indicate the correct comparison.

Table 3. Simple discriminations presented in each training structure.

Training structure	Simultaneous discriminations	Successive discriminations	Amount of simple discrimination
SaN	B1B2, A1B1, A1B2, A2B1, A2B2, E1E2, A1E1, A1E2, A2E1, A2E2, D1D2, A1D2, A1D1, A2D1, A2D2, C1C2, A1C1, A1C2, A2C1, A2C2,	A1A2	21
CaN	B1A1, B1A2, B2A1, B2A2, C1A1, C1A2, C2A1, C2A2, D1A1, D1A2, D2A1, D2A2, E1A1, E1A2, E2A1, E2A2, A1A2	B1B2, B1C1, B1C2, B1D1, B1D2, B1E1, B1E2, B2C1, B2C2, B2D1, B2D2, B2E1, B2E2, C1C2, C1D1, C1D2, C1E1, C1E2, C2D1, C2D2, C2E1, C2E2, D1D2, D1E1, D1E2, D2E1, D2E2, E1E2	45
Mixed-LS	E1A1, E1A2, E2A1, E2A2, C1A1, C1A2, C2A1, C2A2, A1D1, A1D2, A2D1, A2D2, A1B1, A1B2, A2B1, A2B2, D1D2, B1B2, A1A2	E1E2, C1C2, E2C1, E2C2, E1C1, E1C2	25
Mixed-SaN-CaN	E1A1, E1A2, E2A1, E2A2, B1A1, B1A2, B2A1, B2A2, A1C1, A1C2, A2C1, A2C2, A1D1, A1D2, A2D1, A2D2, D1D2, C1C2	E1E2, E1B1, E1B2, E2B1, E2B2, B2B1, A1A2	25

Results

The number of training trials and blocks needed to reach criterion, and the amount of correct responses during the test phase are shown for participants in all groups in Table 4. A total of 37 out of 48 participants reached the training criterion. 11 of 12 subjects of the SaN group reached the criterion in the training phase. Except for subject 9, who required two sessions, all of the participants completed the training phase in one block. 11 participants of the CaN group passed the training phase. Only two participants (5 and 8) required another block of training to reach the criterion. Nine participants of the Mixed-LS group passed the training phase. All participants completed the training phase in one block. Six participants of the Mixed-SaN-CaN reached the criterion in the training phase. Four of the six participants completed the training in one block, while the others (participants 1 and 11) required two blocks to overcome the task.

Table 4. Scores of the four training structures.

		Training		Test		
Participant	Structure	Correct response (last block)	Total blocks	Total correct response	Correct responses	Percentage of correct response
1	SaN	32*	1	32	38	79%
5		32*	1	32	48	100%*
6		32*	1	32	48	100%*
8		32*	1	32	48	100%*
10		32*	1	32	43	90%*
12		32*	1	32	47	98%*
4		31*	1	31	46	96%*
7		31*	1	31	45	94%*
3		30*	1	30	30	63%
11		30*	1	30	32	67%
9		29*	2	64	42	88%*
2		17	3	47	22	46%
2	CaN	32*	1	32	47	98%*
4		32*	1	32	48	100%*
7		32*	1	32	48	100%*
9		32*	1	32	48	100%*
10		32*	1	32	45	94%*
1		31*	1	31	48	100%*
3		31*	1	31	42	88%*
12		31*	1	31	45	94%*
8		31*	2	56	47	98%*
6		30*	1	30	48	100%*
5		29*	2	54	45	94%*
11		26	3	73	23	48%
2	Mixed-LS	32*	1	32	36	75%
3		32*	1	32	38	79%
4		32*	1	32	39	81%
1		31*	1	31	35	73%
9		31*	1	31	43	90%*
5		30*	1	30	35	73%
10		30*	1	30	47	98%*
11		30*	1	30	31	65%
8		29*	1	29	43	90%*
7		25	3	75	31	65%
6		22	3	67	23	48%
12		17	3	48	27	56%
5	Mixed-SaN-CaN	32*	1	32	34	71%
10		32*	1	32	40	83%
12		32*	1	32	46	96%*
11		32*	2	58	47	98%*
1		30*	2	52	37	77%
3		29*	1	29	33	69%
4		27	3	83	23	48%
8		26	3	72	38	79%
2		24	3	64	31	65%
6		20	3	63	35	73%
9		16	3	55	23	48%
7		16	3	51	23	48%

Note: *= Participants that reached criterion in each case.

Twenty-four out of 37 participants passed the test criterion. All participants who did not reach the criterion in the training phase did not show evidence of derived relations. Performance in the test phase is described below only for participants that passed the training phase.

All 11 participants in the CaN group scored above 87% in the derived relations test. Eight participants out if 11 in the SaN group reached the criterion in the test phase. In the case of the Mixed-LS and Mixed-SaN-CaN only three and two participants reached the criterion in the test phase, respectively (Figure 2).

A main effect of group was found for the amount of correct responses in the combined blocks, $F(3,47)= 4.53$, $p= .007$, and amount of correct responses in the last block of training, $F(3,47)= 4.13$, $p= .011$.

In the amount of correct choices in the last combined training block, a difference can only be observed between SaN and mixed SaN-CaN, with SaN having a higher amount of correct responses ($p= .01$), with no differences with the other structures (*p*s >.1).

A main effect of group was found for the variable amount of blocks, $\chi^2(3)= 8.57$, $p= .036$. The amount of training blocks needed to proceed to the testing phase was

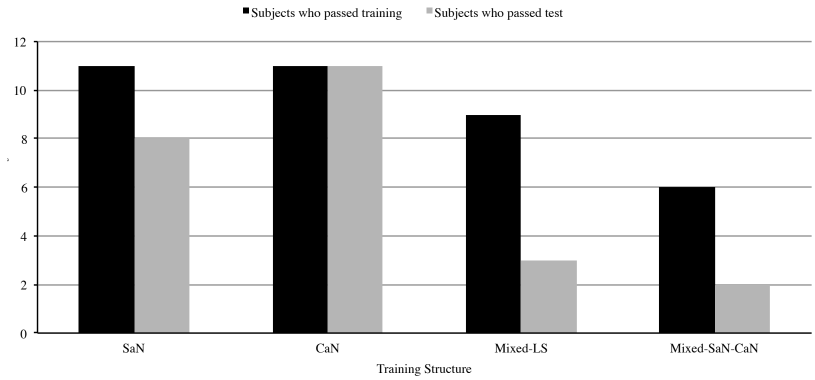

Figure 2. Participants who reached the training (black) and testing (grey) criteria for each training structure.

significantly higher for the mixed SaN-CaN structure than for SaN, Z= -2.49, p= .013 and CaN, Z= -2.24, p= .025. It also required more blocks than mixed-LS, although the significance was marginal, Z= -1.78, p= .079. No differences were found between the SaN and CaN structures, Z= -0.45, p=.65

Participants who did not meet training criteria were excluded from the analysis. A main effect was found, $F(3,36)= 4.53$, $p <.001$. Post hoc tests revealed that the percentage of correct responses is lower for mixed-LS than for CaN ($p <.05$), but not for SaN (p= .38) or mixed SaN-CaN (p= .96). The mixed SaN-CaN structure, on the other hand, had a lower percentage of correct responses than CaN ($p <.05$), but not lower than SaN (p= .17).

DISCUSSION

If we take the difference in the proportion of participants that reached the 90% correct responses criterion during training, the proportion in SaN and CaN is identical (92%). On the other hand, in the mixed-LS group 75% of the participants reached this criterion, while only 50% did so in the mixed SaN-CaN group.

Among the participants that reached the training criterion, the SaN and CaN groups required a lower amount of blocks than the mixed SaN-CaN group to proceed to the testing phase. A clear learning deficiency can be observed when the mixed SaN-CaN structure was employed. Even though various authors propose that the CaN structure is more difficult to learn (since all simple discriminations must be learned; Saunders & Green, 1999; Arntzen *et alii*, 2010) when we analyze the amount of simple discriminations learned in each structures and the subsequent performances, it does not seem that the amount of discriminations is related to a higher difficulty, since there were no differences between SaN and CaN (although the amount of simple discriminations to be learned in the latter is significantly higher than in the former). Furthermore, the amount of blocks necessary was significantly lower for SaN than for mixed SaN-CaN, even though the same amount of simple discriminations was presented during training. If Saunders and Green's proposal was correct, given that nodal density was the same among structures, the amount of blocks needed to reach the training criterion should have been higher in the CaN group.

On the other hand, Fields and Verhave (1987) propose that nodal density could affect the performance, either by promoting or hindering learning. In spite of having kept the nodal density constant among the different structures, differences were found between them. This allows us to infer that differences in performance between classes with more than three members, in which nodal density was not controlled, do not depend on the differences in nodal density among groups, but on the roles of stimuli as sample or comparison.

An alternative and possible source of interference during learning could be the intertrial correspondence. Adamson, Foster, and McEwan (2000) found that the performance of participants in a zero-delay matching-to-sample paradigm worsened from one trial to the next if the sample stimulus was different than that of the previous trial. Furthermore, Worsham (1975) and Roberts (1980) pointed out that proactive interference occurs when the incorrect comparison in the present trial appeared as a sample in the previous trial. In the mixed-LS and mixed SaN-CaN groups, the roles of the stimuli (i.e., functioning as sample or comparison) varied during training (differently to the case of SaN and CaN, where the function of each stimulus was kept constant). In other words, if a stimulus functions as comparison in one trial and as sample in the next (or vice versa), the accuracy of the responses decreases, as opposed to what happens when the stimulus maintains its sample or comparison function in both trials. This explains the better performance in the SaN and CaN groups, in which samples and comparisons maintain their roles in all trials.

Other research has evaluated the role of the comparison stimuli functions in the formation of equivalence classes (Plazas & Peña, 2016). In Experiment 3, two groups were trained. One by an altered matching-to-sample, which was aimed to increase positive control and reduce negative control. The other group was trained with a new matching-to-sample procedure called semistandard because, on each three-choice training trial, one of the incorrect comparisons was correct for another sample, but the other comparison stimulus was never correct on any trial. Altered-group participants achieved a better performance in the test for derived relations compared to the group with negative comparisons. The authors suggests that learning that a stimulus may have different functions (positive and negative) is a more difficult task than learning that a stimulus has one and only one function (positive or negative). Although intra-class negative relationships seem to be necessary for the emergence of equivalence classes (but see Hinojo, Pérez Fernández, & García García, 2017), the fact that the function of the comparison stimulus does not remain constant seems to be a factor that influences learning.

A phenomenon that may have exerted an influence, not necessarily incompatible with the previous one, is backward conditioning between sample and comparison stimuli. Siegel and Domjan (1971) showed that backward associations tend to retard the development of a conditioned response when the unconditioned stimulus (in this case, the comparison) is later used in forward conditioning, that is, functioning as a sample. This is consistent with the literature proposing that respondent conditioning is the process underlying the formation of equivalence classes (Avellaneda *et alii*, 2016; Delgado & Hayes, 2013, 2014; Hayes, 1992; Tonneau, 2001a, b, 2002). Rehfeldt and Hayes (1998), following this idea, proposed that the formation of equivalence classes using a matching-to-sample procedure is a product of stimulus pairings occurring between samples and comparisons, and that the role of feedback is limited to reinforce attending responses to the correct comparison. Future studies should investigate the influence of backward conditioning in SEC learning.

Of the participants that reached the training criterion, all reached the testing criterion in the CaN group. 72.7% did so from the SaN group, while 33.3% from the mixed-LS group and 33.3% from the mixed SaN-CaN group did so, with the percentage of correct responses being significantly higher in the CaN group compared to the mixed-LS and mixed SaN and CaN groups (Figure 2.).

Saunders and Green (1999) propose that the difference in performance of the different structures during testing is due to differences in the amount of simple discriminations learning during training. If we analyze the results in regards to this hypothesis, we can observe that 21 simple discriminations were learned in the SaN group, 45 in the CaN group, and 25 both in the mixed-LS and the mixed SaN-CaN groups (Table 3). As can be seen, more discriminations were learned in the mixed-LS and mixed SaN-CaN groups than in the SaN groups, and results show clearly that this did not result in better performance (either during training). Furthermore, during training in the CaN and mixed SaN-CaN groups, the E-B and B-E simple discriminations were presented, which are the ones tested afterwards. If this hypothesis is correct, learning these simple discriminations should have resulted in a better performance compared to the SaN and mixed-LS groups, which was not the case. We recommend to employ in future experiments a methodology that allows to compare an even amount of participants that reached the training criterion to see if these differences among structures are preserved.

The variability in the functions of the stimuli during training could result in an improvement in learning by simplifying the formulation of verbal rules during training or promoting more “adequate” rules to be extrapolated to the testing phase. Even though no explicit instructions were given to the participants to develop strategies or name the stimuli, McIlvane and Dube (1996) propose that these differential responses can happen independently of instructions given by the experimenters. Furthermore, in structures where the function of the stimuli vary between sample and comparison, the emergence of rules would become even harder, as is the case in LS or mixed structures (such as those used in the present case). Even though rule formulation by the participants seems to promote equivalence class formation, there is not conclusive evidence to determine if it is a sufficient or necessary condition (see Pérez Fernández, 2015, for a review).

Results of this study provide an alternative interpretation of the phenomena known as nodal distance (Fields & Verhave, 1987) and training structure (Saunders & Green, 1999). One possibility is that these effects may be caused by intertrial incongruence product of the higher amount of nodes, characteristic of the LS structure, and neither necessarily of the amount of nodes that link the stimuli nor the simple discriminations learned during training. Also respondent processes and verbal behavior may develop a role in the learning of baseline relations and later emergence of derived relations. Despite that these factors were not approached in the present research, is strongly suggested to study their influence in future investigations.

The need for more strict methodological control procedures that specifically evaluate the emergence and incidence of verbal rules in SEC learning is proposed. Alternatively, a design which hinders rule formulation (as that used in Delgado, Medina, & Soto, 2011) could be used. On the other hand, the idea that respondent conditioning during training of baseline relations could be influencing their learning should be tested. A respondent-type training procedure is proposed as a suitable medium to investigate this. Results could be useful in better understanding the differences among training structures. These aforementioned facilitating effects are not necessarily incompatible with each other. It is possible that some (or all) exert an effect when the SEC are being trained.

The aim of this study was to increase the knowledge of the method of SEC concerning the procedures used. Furthermore, it is helpful to develop more effective learning methods due to the fact that this paradigm has proven to be useful in several areas like: cognitive rehabilitation, education and psychotherapy (Fiorentini *et alii*, 2012).

References

Adamson C, Foster TM, & McEwan, JS (2000). Delayed matching to sample: the effects of sample-set size on human performance. *Behavioural Processes*, *49*, 149-161. Doi: 10.1016/S0376-6357(00)00087-5

Aguado Alonso G (2005). Contribuciones al diagnóstico del trastorno específico del lenguaje por medio de la repetición de pseudo-palabras. *Revista Extremeña de Atención Temprana, 3*, 17-28.

American Psychological Association (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, *57*, 1060-1073. Doi: 10.1037/0003-066X.57.12.1060

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

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

Arntzen 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. *The Psychological Record*, *60*, 437-462.

Avellaneda M, Menéndez J, Santillán M, Sánchez F, Idesis S, Papagna V, & Iorio A (2016). Equivalence class formation is influenced by stimulus contingency. *The Psychological Record, 66,* 477-487. Doi: 10.1007/s40732-016-0187-y

Bentall RP, Jones RM, & Dickins DW (1999). Errors and response latencies as a function of nodal distance in 5-member equivalence classes. *The Psychological Record*, 49, 93–115.

Delgado D & Hayes LJ (2013). The integration of learning paradigms by way of a non-causal analysis of behavioral events. *Conductual*, *1*, 39-54.

Delgado D & Hayes LJ (2014). An integrative approach to learning processes: revisiting substitution of functions. *The Psychological Record*, *64*, 625-637. Doi: 10.1007/s40732-014-0071-6

Delgado D, Medina IF, & Soto JS (2011). El lenguaje como mediador de la transferencia de funciones: ¿es necesario nominar para relacionar? *Revista Mexicana de Análisis de la Conducta, 37,* 31-52

Fields L, Hobbie SA, Adams BJ, & Reeve KF (1999). Effects of training directionality and class size on equivalence class formation by adults. *The Psychological Record*, *49*, 703.

Fields L & Verhave T (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior, 48*, 317-332.

Fiorentini L (2012). Una revisión de las aplicaciones del paradigma de equivalencia de estímulos. *International Journal of Psychology and Psychological Therapy*, *12*, 261-275.

Hayes LJ (1992). Equivalence as process. In SC Hayes & LJ Hayes (Eds.), *Understanding Verbal Relations* (pp. 97-108). Reno, NV: Context Press.

Hinojo Z, Pérez Fernández V, & García García A (2017). The formation of equivalence classes in adults without training in negative relations between members of different classes. *International Journal of Psychology and Psychological Therapy, 17*, 107-118.

Kinloch JM, McEwan JSA, & Foster TM (2013). Matching-to-sample and stimulus-pairing-observation procedures in stimulus equivalence: The effects of number of trials and stimulus arrangement. *The Psychological Record*, *63*, 157-173.

McIlvane WJ & Dube W (1996). Naming as a facilitator of discrimination. *Journal of the Experimental Analysis of Behavior*, *65*, 267-272.

Moise SL (1976). Proactive effects of stimuli, delays, and response position during delayed matching from sample. *Animal Learning & Behavior*, *4*, 37-40.

Pérez-Fernández V (2015). Clases de equivalencia y conducta verbal. *Conductual, 3*, 26-44.

Plazas EA & Villamil CW (2016). Efecto del entrenamiento de relaciones entre-clases y estructuras de entrenamiento en la formación de relaciones de equivalencia [Effect of Negative Between-class Training and Structure

Training on Formation of Equivalence Relations.]. *International Journal of Psychology and Psychological Therapy, 16*, 295-314.

Rehfeldt RA & Hayes LJ (1998). The operant-respondent distinction revisited: Toward an understanding of stimulus equivalence. *The Psychological Record*, *48*, 187-210.

Roberts WA (1980). Distribution of trials and intertrial retention in delayed matching to sample with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *6*, 217-237. Doi: 10.1037/0097-7403.6.3.217

Sánchez FJ, Menéndez J, Avellaneda MA, Idesis SA, & Iorio AA (2016). Training structures of equivalence classes and their influence on the priming effect. *Revista Argentina de Ciencias del Comportamiento*, *8*, 8-19.

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

Sidman M (1986). Functional analysis of emergent verbal classes. In T Thompson & MD Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213-245). Hillsdale, NJ: Erlbaum.

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

Siegel S & Domjan M (1971). Backward conditoning as an inhibitory procedure. *Learning and Motivation*, *2*, 1-11.

Tonneau F (2001a). Equivalence relations: A critical analysis. *European Journal of Behavior Analysis*, *2*, 1-33.

Tonneau F (2001b). Equivalence relations: A reply. *European Journal of Behavior Analysis*, *2*, 99-128.

Tonneau F (2002). Who can understand relational frame theory? A reply to Barnes-Holmes and Hayes. *European Journal of Behavior Analysis*, *3*, 95-102.

Worsham RW (1975). Temporal discrimination factors in the delayed matching-to-sample task in monkeys. *Animal Learning & Behavior*, *3*, 93-97.

Received, May 12, 2017
Final Acceptance, July 18, 2017