The Psychological Record, 2008, 58, 597–622

EQUIVALENCE-EQUIVALENCE RESPONDING: TRAINING CONDITIONS INVOLVED IN OBTAINING A STABLE BASELINE PERFORMANCE

Andrés García, Cristóbal Bohórquez, Vicente Pérez,
María Teresa Gutiérrez, and Jesús Gómez
Universidad Nacional de Educación a Distancia, Madrid, Spain

Recent research has focused on the variables associated with equivalence-equivalence responding, in which participants match pairs of equivalent or nonequivalent stimuli. One such variable is the presence of response competition from nonarbitrary (physical) relational response options. In the current analysis, the experimenters examined the effect of training conditions on the likelihood of equivalence-equivalence responding in the presence of these competing response options. The conditional discrimination training conditions prompting equivalence relations were systematically manipulated across five experiments. The study included the variable reflexivity training, an equivalence test, the type of stimuli used, and the number of training trials. Results revealed that the manipulations gradually increased the percentage of participants responding in accordance with equivalence-equivalence. The implications of these findings are discussed.

If conditional discrimination training is used to generate the unidirectional relational responses "match B with A" and "match C with B", where A, B, and C are arbitrary stimuli, then in an additional appropriate testing context, a series of further, nonexplicitly trained relational responses will often appear. Such derived relational responses might include matching A with A, B with B, and C with C (reflexivity or identity matching); matching A with B and B with C (reversal of trained relations or symmetry); matching C with A (transitivity); and matching A with C (the combination of symmetry and transitivity). This phenomenon is referred to as stimulus equivalence (Sidman, 1971; Sidman & Tailby, 1982). Since it is characterized by both generativity and bi-directionality, it has provided a useful avenue for the empirical investigation of certain complex processes. Examples are creativity (Gómez, García, Pérez, Gutiérrez, & Bohórquez, 2004), symbolic behavior (Carr & Blackman, 1996; Horne & Lowe, 1996), language (Fields, Verhave & Fath, 1984; Sidman, 1986a, 1990b), and concept formation (Zentall, Galizio, & Critchfield, 2002).

Correspondence concerning this article should be addressed to Andrés García García, Departamento de Psicología Experimental, Universidad de Sevilla, Campus Ramón y Cajal, C/ Camilo José Cela s/n Sevilla-41018 (Spain). E-mail: andresgarcia@us.es

Evidence from previous research has shown stimulus equivalence to be a highly robust phenomenon, appearing in studies that include psychologically-impaired populations (Escuer, García, Bohórquez, & Gutiérrez, 2006; Hall, DeBernardis, & Reiss, 2006; O'Donnell & Saunders, 2003; Saunders & Spradlin, 1993; Sidman, 1971; Sidman, Cresson, & Wilson-Morris, 1974; Stromer & Osborne, 1982), normally-developed children of different ages (Denavy, Hayes, & Nelson, 1986; Lipkens, Hayes, & Hayes, 1993; Pilgrim, Chambers, & Galizio, 1995), adults from different cultures and with differing levels of education (Bush, Sidman, & de Rose, 1989; Lazar, 1977; Wulfert & Hayes, 1988), and elderly people (Pérez-González & Moreno-Sierra, 1999). Likewise, the variable that best shows the importance of studying equivalence relations is its capacity to improve the learning process (e.g., Cowley, Green, & Braunling-McMorrow, 1992; de Rose, Souza, Rossito, & de Rose, 1992; García, Gutiérrez, Gómez, & Puche, 2001; Lynch & Cuvo, 1995; Maydak, Stromer, Mackay, & Stoddard, 1995; Stromer, Mackay, & Stoddard, 1992; for an in-depth review, see García & Benjumea, 2002).

For the most part, studies on equivalence have employed single-element stimuli as sample and comparison stimuli. However, several research projects on derived relations have used compound or multiple-element stimuli (Benigno-Alonso & Pérez-González, 2006; Carpentier, Smeets, & Barnes-Holmes, 2000, 2002a; Maguire, Stromer, Mackay, & Demis, 1994; Markhan & Dougher, 1993; Markhan, Dougher, & Augustson, 2002; Pérez-González, 1994; Schenk, 1993; Smeets, Schenk, & Barnes, 1995; Stromer & Stromer, 1990a, 1990b). Furthermore, recent studies have shown that human subjects readily match compound stimuli containing single elements included in an equivalence relation with other compound stimuli containing equivalent elements. Additionally, they are able to relate compound stimuli containing nonequivalent elements to other nonequivalent compounds. This phenomenon is referred to as equivalence-equivalence (Barnes, Hegarty, & Smeets, 1997; Carpentier, Smeets, & Barnes-Holmes, 2002b, 2003a, 2003b; Carpentier, Smeets, Barnes-Holmes, & Stewart, 2004; Stewart, Barnes-Holmes, & Roche, 2004; Stewart, Barnes-Holmes, Roche, & Smeets, 2001).

Barnes et al. (1997) conducted the first study on equivalence-equivalence relations. These authors trained four three-member equivalence relations (A1B1C1, A2B2C2, A3B3C3, A4B4C4) and tested for several BC-BC (equivalence-equivalence) derived relations. One of the findings of this study was that subjects were more likely to choose a comparison with elements pertaining to an equivalence class (e.g., B1C1). This occurred when the sample was made up of two elements that joined the same equivalence relation. On the other hand, when the sample consisted of two elements pertaining to different equivalence classes (i.e., B2C1), the subjects tended to choose a comparison whose elements were formed by two elements from different equivalence relations. They called this phenomenon equivalence-equivalence responding, and it has been regularly used as a model in which the elements concerned are comparable with those involving analogical reasoning (Stewart, Barnes-Holmes, Hayes & Lipkens, 2001).

The aforementioned authors (Barnes et al., 1997) have already asked one question that should be addressed: The possible interference from the reflexive property (physical similarity between stimuli) in responses based on the equivalence-equivalence criterion.

Concerning equivalence between single-element stimuli, whenever one of the incorrect comparisons maintains a nonarbitrary relation of physical similarity with the sample (e.g., the same color, but different from that of the correct comparison), subjects tend to respond in a way that is consistent with the physical likeness between the two stimuli. Nonetheless, this choice is correct if the nonarbitrary criterion of physical similarity between stimuli is evaluated (Stewart, Barnes-Holmes, Roche, & Smeets, 2002). Responses based on nonarbitrary relations provide a context that is more relevant in the subject's reinforcement history. A child seldom shows the properties of symmetry and transitivity without having previously demonstrated nonarbitrary responses based on similarity (Hayes, 1991). However, it could be argued that the ability of the subjects to respond to arbitrary (i. e., equivalence) or nonarbitrary (i. e., identity) properties of the stimuli in a matching-to-sample task depends on the nature of the learning procedure and the programmed consequences. A design employing an adequate number of comparisons, effective reinforcement programming, counterbalanced comparisons, and other contingencies would lead a subject to learn the conditional discriminations of the training, and derived equivalence relations would appear as a result. Instead, a nonstructured and weak contingency programmed task (i. e., when no feedback is given) would lead a subject to select the more conservative option, which is the physical likeness between stimuli.

Evidence from previous experiments suggests the absence of interference from a nonarbitrary criterion of physical similarity in equivalence-equivalence responding tasks (Barnes et al., 1997; Carpentier, Smeets, & Barnes-Holmes, 2002b). In those studies, an equivalence-equivalence test, including a competition criterion based on the nonarbitrary relation of physical similarity, was carried out with the subjects having previously performed an equivalence-equivalence test without any competitive element. Just as Barnes et al. (1997) pointed out about their investigation, a blocking effect (Kamin, 1968, 1969) may be a competing response during test trials. This phenomenon of competition has been reported when both criteria are introduced at the outset of the equivalence-equivalence tests. A subject is more likely to select a physical nonarbitrary criterion when it is available from the beginning. (García, Gutiérrez, Bohórquez, Gómez, & Pérez, 2002). It has also been shown that overshadowing (García, Bohórquez, Gómez, Gutiérrez, & Pérez, 2001) and blocking effects arise in these types of relations (García, Gómez, Pérez, Bohórquez, & Gutiérrez, 2003). Obtained data have pointed to the intensity of training as being an essential element to take into account when attempting to establish equivalence-equivalence relations while avoiding interference from a nonarbitrary criterion (Bohórquez, García, Gutiérrez, Gómez, & Pérez, 2002).

The objective of the present study was to test the influence of a series of variables on establishing a baseline of responses based on an arbitrary equivalence-equivalence criterion. First, the explicit training of reflexive relation was accomplished before the necessary relations for equivalence relations were trained to appear. Second, the decision was made to include an equivalence test after the training, or, where appropriate, to omit it. Furthermore, an assessment was carried out as to whether the type of stimuli may have had an influence. For this purpose, either figures or meaningless syllables were used. Finally, the effect of the intensity of the training on the prerequisites of the equivalence relations was evaluated, based on the number of trials composing each block (A-A, A-B, and A-C). Five experiments were conducted to evaluate how such variables influence the choice of a stable criterion, based either on arbitrary equivalence-equivalence relations or on nonarbitrary relations of similarity.

Experiment 1

The purpose of the first experiment was to check how subjects performed in a competition test in which two possible response criteria were presented. The first criterion involved the possibility of responding according to the arbitrary relations established through a conditional discriminations training. The learned relations met the necessary requisites to form three equivalence relations (A1B1C1, A2B2C2, and A3B3C3). The second criterion was based on the physical similarity between stimuli, where the sample and one of the comparisons always shared a common element (e.g., A1B1 as the sample, B2C2 as a comparison based on the trained equivalence relations, and A1B3 as a comparison based on the criterion of physical similarity between stimuli).

Method

Subjects

Fifteen subjects voluntarily participated in this experiment. All had a university education and were between 19 and 45 years of age. The subjects were quasi-randomly assigned to three groups, according to the stimuli that formed the equivalence relations, to counterbalance each possibility (e.g., a stimulus that was A1 for one subject would be A2 for another subject, and so on).

Apparatus and Procedure

The stimuli comprising the trained conditional discriminations were nine figures, designed during a previous study (see Figure 1). Each element was assigned an alphanumeric code (A1, B1, C1, etc.) for reference purposes during the investigation, but the subjects were never aware of this code. The stimuli were displayed on a standard 14-in. monitor using software specially designed to administer the series of trials that made up the task the subjects would perform.

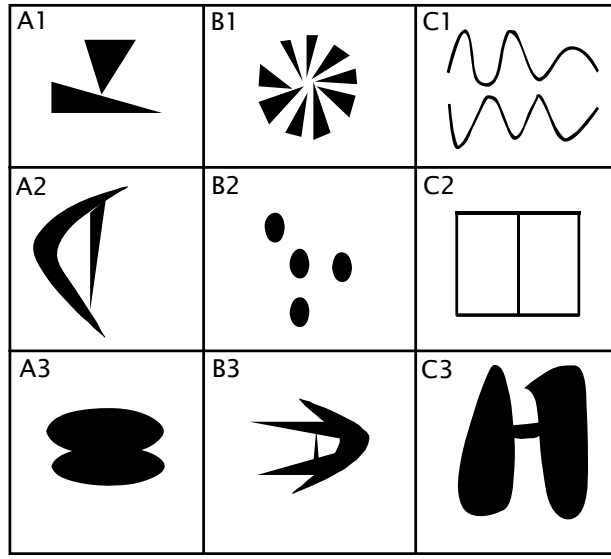

Figure 1. Stimuli used with the assigned codes A1, B1, C1, etc., corresponding to the three equivalence classes trained (A1B1C1, A2B2C2, and A3B3C3).

At the start of the task, the subjects sat in front of the computer, and an identification form appeared on screen, which they had to complete with certain personal details (name, age, course, current date and time, and the teaching institution to which they belonged). Once this was completed, the following instructions, identical for all participants, were then displayed.

> First, we would like to thank you for participating in this study.
>
> We would also like to remind you of the following:
>
> This is not an intelligence test.
>
> This is not a personality test.
>
> Nor is it a test of speed, you may take as long as you need.
>
> Use only the left mouse button, do not use the keyboard or the right mouse button.
>
> A series of stimuli will now appear on the monitor. A sample will always appear first, which you must click on. Some possible response options will then appear at the top of the screen. You must click on whichever you think is correct.

Having read the instructions, the subjects then carried out the series of training trials, each displayed on the computer monitor upon a white background.

Task Characteristics

The training procedure was one of simultaneous matching-to-sample, with an observational response. First, a sample stimulus appeared in the center of the screen, remaining there until the subject clicked the figure. The figure became smaller and moved to the bottom of the screen, and three comparison stimuli then appeared at the top: one on the left, one in the center, and one on the right. To go on to the next stage, the subjects had to click whichever of the three comparisons they deemed to be correct. If they chose correctly, a green background appeared displaying the word "RIGHT." If they chose incorrectly, a red background appeared saying "WRONG, TRY AGAIN," giving the subjects the chance to respond again in the same trial until they gave a correct response. This procedure was the same for each training phase.

Training and Test Sequence

Phase 1: Training A-B and A-C. The training trials began with the relation A-A (A1-A1, A2-A2, A3-A3) made up of 15 trials. If the subjects made no more than two mistakes, they would move on to training A-B (A1-B1, A2-B2, A3-B3). If at least 13 trials were correctly performed, they went on to training A-C (A1-C1, A2-C2, A3-C3); and having passed it on the basis of the same criterion established for the previous series, they were given a series of 36 trials with the A-A, A-B, and A-C discriminations mixed. They would pass this block if they had made no more than four errors. This series of training blocks provided the subjects with sufficient training to demonstrate three equivalence classes, with three members each (A1B1C1, A2B2C2, and A3B3C3), by using a "one-to-many" procedure.

Phase 2: Choice between equivalence-equivalence and similarity criteria. If the subjects managed to overcome the training phase, they moved on to compound-stimuli test blocks (the term "compound" is used when the relation between two elements acts as a stimulus). They were presented in the following format: First, a stimulus composed of two figures appeared in the center of the screen. Once the subject had responded to this stimulus (sample), it became smaller and moved to the bottom of the screen. At the same time, two compound-stimuli (comparisons), similar to the one described above, appeared at the top of the screen. Each was comprised of two individual stimuli from those used in training during the previous phase. None of the tests with compound stimuli was subsequently reinforced.

During this initial compound-stimuli trials phase, the subject's task consisted of making a choice; i.e., both comparison stimuli were considered correct, each through a different criterion. The response criterion for one comparison (left or right, positions were counterbalanced), consisted of an equivalence-equivalence relation shared with the sample. In other words, if the sample was composed of two figures belonging to the same equivalence class, that comparison would be in the same equivalence as well. In case the sample was made up of two stimuli of different equivalence classes, the same relation would exist in the comparison. This comparison was considered correct by means of the nonequivalence-nonequivalence criterion.

The second comparison shared a common element with the sample (the position was also counterbalanced throughout the tests); i.e., one of the two elements that made up both the sample and the comparison was identical to the other (Figure 2).

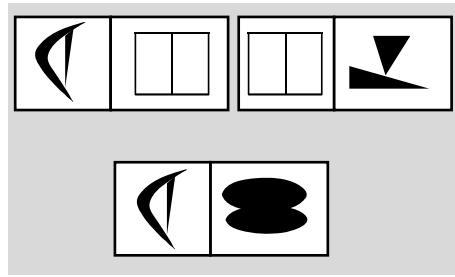

Figure 2. Example of equivalence-equivalence choice test trial, with the response option based on the nonarbitrary relation of similarity (left comparison) or in the arbitrary equivalence-equivalence relation (right comparison).

The participant was assumed to have made a stable choice when, during 20 consecutive trials, his or her responses were based on the same response criterion, whether equivalence-equivalence or physical similarity between sample and comparison. This test stage, consisting of a 54-trial block, stopped only when the computer had recorded20 successive responses based on the same criterion.The phase was repeated if no stable performance was registered after the end of the block until 20 responses were given on the same criterion.

Phase 3: Nonchosen criterion test. Once the subjects had made what was considered a stable choice, their performance was evaluated in the criterion that they had not previously been chosen. In other words, if one subject had chosen the equivalence-equivalence criterion during the selection phase, the similarity criterion would then be evaluated, and vice-versa. This procedure was carried out through 18 test trials similar to those described for the previous phase, but with one difference: This time, only one correct response criterion, either equivalence-equivalence or similarity (Figure 3), was available. Given the sample A1C1, for example, in the test of the similarity criterion, only one answer was correct; this was the comparison that shared a common element with the sample, A1A3. Further, given the same sample, the only correct comparison (Figure 3, on the right side) was the one based on the equivalence-equivalence criterion, B3C3. During this series, the number of correct responses was recorded.

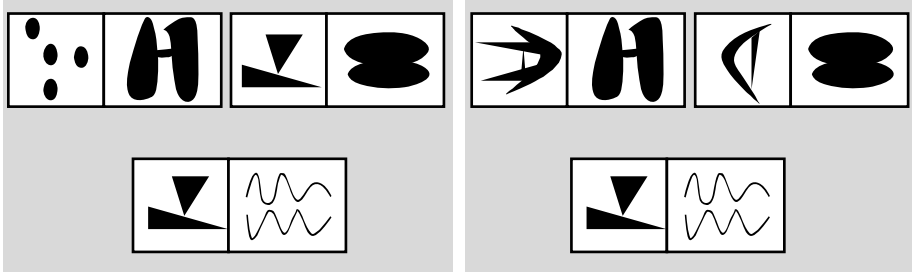

Figure 3. Example of compound-stimuli test trials, similarity (left) and equivalence-equivalence (right) as single response criteria.

Table 1
Experiment 1: Number of Correct Responses per Block of Trials, Criterion Chosen in the Selection Phase and Percentage of Correct Responses in the Nonchosen Criterion Test

	Subjects														
Tasks	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AA	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15
AB	15/15	12/15 13/15	2/15 13/15	10/15 13/15	10/15 13/15	5/15 14/15	12/15 15/15	15/15	1/15 4/15 13/15	9/15 13/15	12/15 14/15	8/15 15/15	6/15 12/15 14/15	14/15	13/15
AC	12/15 15/15	9/15 15/15	12/15 15/15	14/15	12/15 15/15	12/15 15/15	15/15	10/15 14/15	12/15 15/15	12/15 13/15	12/15 15/15	11/15 15/15	12/15 15/15	14/15	13/15 15/15
ABC	36/36	35/36	35/36	36/36	34/36	36/36	35/36	36/36	36/36	36/36	35/36	36/36	35/36	35/36	35/36
EQ-EQ / SIM	SIM	SIM	SIM	EQ-EQ	EQ-EQ	EQ-EQ	SIM	SIM	SIM	SIM	SIM	SIM	SIM	SIM	EQ-EQ
Correct Responses in Nonchosen Criterion	13/18 72%	13/18 72%	10/18 56%	11/18 61%	0/18 0%	11/18 61%	13/18 72%	14/18 78%	13/18 72%	10/18 56%	5/18 28%	13/18 72%	2/18 11%	10.18 56%	15/18 83%

The developed software recorded all the data that was gathered, the data relating both to training and to the different tests carried out. A detailed result sheet was then produced, displaying all the choices the subject made on a trial-by-trial basis.

Results and Discussion

The number of training trials the subjects needed to achieve the learning criterion ranged between 81 and 126 (with an average of 96 trials). For the results obtained during selection tests (arbitrary equivalence-equivalence criterion or nonarbitrary similarity criterion), all subjects performed consistently. Four subjects (27%) chose the equivalence-equivalence criterion, and the majority (11 subjects, 73%) based their choice on the physical similarity criterion between sample and comparison stimuli.

Concerning the percentage of correct choices, since subjects were given a test based on the criterion opposite to the one they had chosen during the selection phase, only two of the subjects achieved a score of over 75%. Only one subject (Subject 8) had chosen the similarity criterion; this subject's performance was then tested in a task where the only available criterion was the one based on equivalence-equivalence. Subject 8 showed a learning trend almost reaching criterion at this stage (78%). The second subject (Subject 15) opted for the arbitrary relations between stimuli, subsequently achieving 83% of correct responses in a test of physical similarity.

The current experiment showed how most subjects chose a criterion based on the identity between a stimulus shared by both the sample and one of the comparisons. Three of the four subjects who based their performance on the arbitrary relations of equivalence-equivalence were subsequently unable to respond on the basis of relations of physical similarity between stimuli. Therefore, there was competition between both criteria. Previous studies (Barnes et al., 1997; Bohórquez et al., 2002) have referred to the possible influence from an evaluation of equivalence carried out prior to the establishment of equivalence-equivalence responses. The question arises as to whether the subjects would change their choices if the aforementioned test were included. This variable was introduced in the next experiment of the series.

Experiment 2

For this second experiment, a similar approach to that employed in Experiment 1 was used except that an equivalence test was given before the compound-stimuli tests of criterion selection were conducted. The objective was to check how this test would influence the choices the subjects made. Testing and clearly defining the arbitrary relations of equivalence between stimuli might lead the subjects to make a more likely choice based on equivalence-equivalence relations.

Method

Subjects

Fifteen subjects took part voluntarily in the experiment. All were university graduates and between 19 and 45 years old. They were assigned to different counterbalance groups in a similar manner to that of Experiment 1.

Apparatus and Procedure

The stimuli were the same nine figures used in the previous experiment (Figure 1). The procedure was identical to the one used previously except for the inclusion of an equivalence test after the block of combined trials A-A, A-B, A-C. This test consisted of 15 C-B trials, with no programmed consequences after the completion of each one. If the subject made two or fewer errors, he or she moved on to the compound-stimuli test trials, which were performed in the same way as in the previous experiment. In case the subject failed to pass the equivalence test, a new block of combined trials (A-A, A-B, A-C) was administered, followed by a new equivalence test and the compound test trials phase. The equivalence test administered (C-B) combined symmetry and transitivity evaluations. It was therefore considered enough to confirm the appearance of derived relations from the trained conditional discriminations.

Results and Discussion

Including the equivalence test did not cause considerable alterations—in this case, in the task of learning about A-A, A-B, and A-C relations. All the subjects passed the test without incident except for two of them (Subjects16 and 29), who failed the first administration of the test. However, they passed the test after the block of combined training trials A-A, A-B, A-C was repeated. The average number of trials for this group was 111, with a range between 81 and 177 trials. During the choice-tests trials, 10 of the subjects (67%) chose the equivalence-equivalence criterion. The number of subjects who chose the criterion based on arbitrary relationships increased compared with the previous experiment. Only one modification was made: An equivalence test was included at the end of the training trials.

The result was that eight subjects who were assessed (Subjects 20, 21, 23, 24, 25, 26, 27, and 29) obtained good scores in the criterion they they had not chosen, with all of them scoring above 75%. Of these, one subject assessed in equivalence-equivalence scored 100%, having previously chosen a criterion based on the physical similarity between stimuli.

The performance of an equivalence test after the training therefore resulted in a greater number of subjects choosing the criterion based on the arbitrary relations of equivalence-equivalence. Until now, two different results have been found when the same type of stimuli was used. Figure 1 shows that (a) the subjects for the most part chose a nonarbitrary criterion when it competed with a criterion based on trained equivalence relations; and (b) the majority of the subjects responded according to trained arbitrary relations. In the following experiment, the main question that arose was related to the type of stimuli, particularly to what extent the change from figures to meaningless syllables would influence the subjects' choices.

Experiment 3

This time, a new modification was included, which in this case was based on the change of the type of stimuli used during the equivalence classes training. Syllables replaced the figures shown previously.

Table 2
Experiment 2: Number of Correct Responses per Block of Trials, Percentage of Correct Responses in the Equivalence Test, Criterion Chosen in the Selection Phase and Percentage of Correct Responses in the Nonchosen Criterion Test

	Subjects														
Tasks	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
AA	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15
AB	12/15 15/15	5/15 14/15	4/15 9/15 14/15	12/15 15/15	7/15 8/15 12/15 14/15	13/15	12/15 14/15	14/15	12/15 15/15	10/15 15/15	8/15 15/15	14/15	15/15	14/15	13/15
AC	3/15 14/15	12/15 15/15	13/15	14/15	11/15 14/15	11/15 14/15	12/15 15/15	14/15	3/15 14/15	12/15 15/15	13/15	12/15 15/15	11/15 15/15	11/15 15/15	12/15 15/15
ABC	35/36 35/36	36/36	36/36	35/36	32/36 35/36	35/36	35/36	36/36	35/36	36/36	36/36	36/36	36/36	35/36 36/36	36/36
CB Equivalence Test	13/15 87% 15/15 100%	15/15 100%	15/15 100%	15/15 100%	14/15 93%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	14/15 93%	14/15 93%	15/15 100%	15/15 100%	4/15 27% 15/15 100%	14/15 93%
EQ-EQ / SIM	SIM	EQ-EQ	EQ-EQ	EQ-EQ	SIM	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	SIM	SIM	EQ-EQ	EQ-EQ	EQ-EQ	SIM
Correct Responses in Nonchosen Criterion	11/15 61%	12/18 67%	13/18 72%	13/18 72%	17/18 94%	17/18 94%	12/18 67%	14/18 78%	11/18 94%	18/18 100%	14/18 78%	15/18 83%	12/18 67%	17/18 94%	8/18 44%

Method

Subjects

Fifteen subjects, of characteristics similar to those of the participants in the two previous studies, formed part of this third experiment. The assignment to counterbalance groups was carried out under the same conditions.

Apparatus and Procedure

The trained stimuli, which would form part of the training in three equivalence classes of three members each, were a set of nine meaningless syllables, created specially for this experiment (see Figure 4). The syllables were always designed with the same logic, with the composition being consonant-vowel-consonant. An alphanumeric code (A1, B1, C1, etc.) was assigned to each element for reasons of reference during the investigation, but the subjects never saw this code.

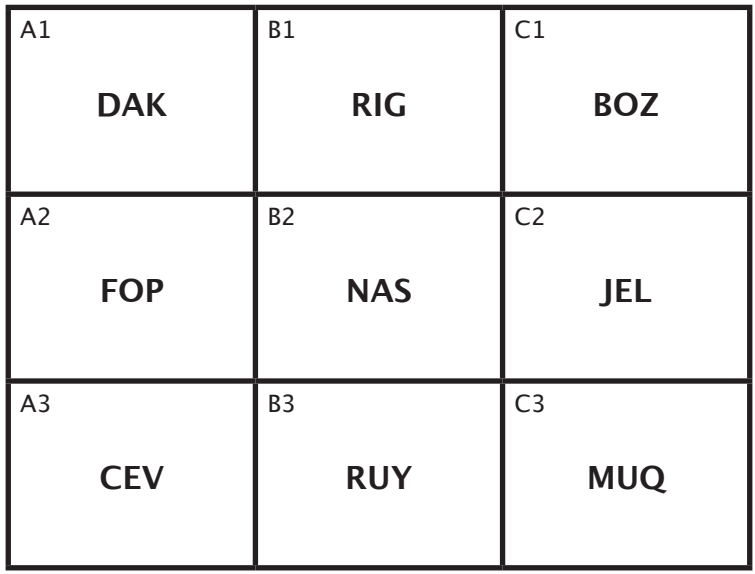

Figure 4. Stimuli used during Experiments 3, 4, and 5, the assigned codes A1, B1, C1, etc, corresponding to the three classes of equivalence trained.

The procedure used was similar to that in Experiment 2: After training of the A-A, A-B, and A-C relations, an equivalence test (C-B) was administered. After that, the compound-stimuli test trials were carried out in the same way as described previously.

Results and Discussion

During A-A initial training, Subject 36 made two mistakes. The remaining subjects showed a faultless performance. The block of trials A-B was the one that took the highest number of trials to pass, compared with the previous experiments; and two of the participants (Subjects 33 and 42) had to repeat

Table 3
Experiment 3: Number of Correct Responses per Block of Trials, Percentage of Correct Responses in the Equivalence Test, Criterion Chosen in the Selection Phase and Percentage of Correct Responses in the Nonchosen Criterion Test

	Subjects														
Tasks	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
AA	15/15	15/15	15/15	15/15	15/15	13/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15
AB	9/15 12/15 15/15	11/15 15/15	5/15 8/15 11/15 14/15	2/15 15/15	10/15 9/15 14/15	12/15 15/15	3/15 12/15 15/15	13/15	10/15 15/15	13/15	8/15 10/15 15/15	6/15 9/15 8/15 15/15	10/15 15/15	11/15 15/15	11/15 13/15
AC	14/15	12/15 15/15	14/15	11/15 12/15 15/15	11/15 15/15	11/15 15/15	12/15 15/15	13/15	14/15	11/15 14/15	11/15 15/15	15/15	10/15 15/15	12/15 15/15	13/15
ABC	36/36	36/36	34/36 35/36	29/36 29/36 35/36	34/36	36/36	36/36	36/36	36/36	36/36	34/36 35/36	35/36 36/36	34/36 34/36	35/36	35/36 35/36
CB Equivalence Test	15/15 100%	15/15 100%	13/15 87% 14/15 93%	15/15 100%	14/15 93%	15/15 100%	15/15 100%	14/15 93%	15/15 100%	14/15 93%	11/15 73% 15/15 100%	10/15 67% 15/15 100%	6/15 40% 15/15 100%	15/15 100%	10/15 67% 15/15 100%
EQ-EQ / SIM	SIM	EQ-EQ	SIM	SIM	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	SIM	EQ-EQ	SIM	EQ-EQ	SIM
Correct Responses in Nonchosen Criterion	9/18 50%	17/18 94%	13/18 72%	9/18 50%	5/18 28%	10/18 56%	15/18 83%	0/18 0%	16/18 89%	4/18 22%	14/18 78%	1/18 6%	10/18 56%	18/18 100%	16/18 89%

it up to four times. The situation normalized on reaching block A-C, where familiarity with the task enabled the subjects to perform it more accurately. Five subjects had to repeat the equivalence test (Subjects 33, 41, 42, 43, and 45), passing again through the mixed-trials block A-A, A-B, A-C, and performing it without difficulties on the second attempt. The average for the training trials carried out was 131 (ranging between 81 and 198 trials).

Considering the choices made during the block of compound-stimuli test trials, six subjects (40%) chose the option of physical similarity between stimuli. The same number of subjects showed a good performance in the test that measured the nonchosen criterion.

From these results, a change in the type of stimuli (syllables vs. figures) apparently didn't lead the subjects to choose an arbitrary criterion with greater probability over a criterion based on physical similarity between stimuli. Aiming to continue identifying the variables that might affect a stable response tendency based on an equivalence-equivalence criterion, the next step was to pose a question already hinted in previous research (Bohórquez et al., 2002): the intensity of conditional discrimination training (A-A, A-B, A-C).

Experiment 4

To carry out this experiment and to ascertain the influence of training on the subjects' choices, the number of trials that made up each of the blocks A-A, A-B, and A-C was increased to 21 (compared with the 15 trials used in previous blocks).

Method

Subjects

To form part of this experiment, fifteen subjects were chosen, with characteristics similar to those of the participants in previous experiments. They were assigned to groups arranged in the same way as in previous procedures.

Apparatus and Procedure

In this case, the procedure used was exactly the same as that used in Experiment 3 except for a modification related to training. Each training A-A, A-B, and A-C block was formed of 21 trials. In the same way, the subjects then took part in a block of mixed trials A-A, A-B, A-C. After an equivalence test (C-B), the subjects moved on to the assessment based on equivalence-equivalence, or similarity.

Results and Discussion

The number of training trials the subjects in this group performed was high, because of the increase in the size of the blocks; The number of repetitions carried out was similar to those the participants in the previous experiment carried out. The average number of trials needed to reach a stable performance criterion was 156, with a range between 120 and 219 trials. Four subjects had to repeat the equivalence test (Subjects 48, 49, 50, and 51), performing it successfully when they were again administered the mixed training block A-A, A-B, A-C.

Table 4
Experiment 4: Number of Correct Responses per Block of Trials, Percentage of Correct Responses in the Equivalence Test, Criterion Chosen in the Selection Phase, and Percentage of Correct Responses in the Nonchosen Criterion Test

	Subjects														
Tasks	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
AA	21/21	21/21	21/21	21/21	21/21	21/21	21/21	21/21	21/21	21/21	21/21	21/21	21/21	21/21	21/21
AB	20/21	16/21 21/21	17/21 21/21	15/21 20/21	7/21 12/21 17/21 21/21	17/21 15/21 21/21	15/21 21/21	13/21 21/21	16/21 21/21	14/21 17/21 20/21	15/21 21/21	14/21 19/21	12/21 21/21	8/21 15/21 20/21	0/21 12/21 19/21
AC	16/21 20/21	17/21 21/21	18/21 20/21	19/21	17/21 21/21	16/21 20/21	20/21	17/21 21/21	21/21	21/21	17/21 20/21	18/21 20/21	19/21	12/21 21/21	18/21 17/21 20/21
ABC	35/36	36/36	35/36 36/36	27/36 34/36 35/36	33/36 36/36	35/36 36/36	36/36	35/36	36/36	36/36	36/36	36/36	25/36 34/36	35/36	36/36
CB			11/15 73%	8/15 53%	12/15 80%	11/15 73%									
Equivalence Test	14/15 93%	15/15 100%	14/15 93%	14/15 93%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	14/15 93%	14/15 93%
EQ-EQ / SIM	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	SIM	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ
Correct Responses in Nonchosen Criterion	14/18 78%	14/18 78%	18/18 100%	12/18 67%	8/18 44%	14/18 78%	18/18 100%	9/18 50%	14/18 78%	15/18 83%	14/18 78%	16/18 89%	8/18 44%	0/18 0%	17/18 94%

As to the selection-test trials,the experimenters observed a considerable change in these subjects when their results are compared with those obtained in the three previous experiments. A large majority of them responded in a consistent way on the basis of the arbitrary equivalence-equivalence criterion (14 subjects out of 15, or 93%). Results indicate that an intensification of the training trials provided the subjects a stable baseline on which to lead their subsequent responses based on the equivalence relations they had learned. With those subjects, the assessment of the nonarbitrary physical similarity criterion was particularly interesting, since a high number of them (10 of those who chose the equivalence-equivalence criterion) gave an efficient performance. Despite selecting an arbitrary criterion, subjects showed their capability to respond when faced with the nonarbitrary characteristics of the stimuli. Finally, as a question to be dealt with in the final experiment of the series, the possibility was considered of the influence of the reflexivity training administered to the subjects at the beginning of the blocks of trials. In other words, to what extent do the participants respond on the basis of a nonarbitrary criterion of physical similarity because they have been trained to do so during the training procedure.

Experiment 5

In this final experiment, still maintaining the intensity of the training on which the previous experiment was based, the conditional discrimination training was explicitly manipulated by withdrawing the A-A block. In this way, the reflexive property was not trained.

Method

Subjects

Fifteen subjects with characteristics identical to those described in previous experiments were assigned to this task, according to the counterbalance criterion also set out previously.

Apparatus and Procedure

The stimuli used during the performance of this experiment were the series of meaningless syllables shown previously (Figure 4). The task was similar to that described in Experiment 4 except for one difference: The training in equivalence classes consisted of two blocks of 21 trials, A-B and A-C, plus the block of combined trials A-B, A-C. After an equivalence test was conducted, the following step was the choice of compound-stimuli trials; and, finally, the assessment of the criterion was not chosen.

Results and Discussion

All the subjects passed the training and the equivalence test without difficulty, with only three of them needing to repeat it (Subjects 63, 64, and 65),and therefore receiving a further block of training A-B, A-C. The average number of trials necessary to reach the criterion was 126, with a range between 99 and 219 trials.

Table 5
Experiment 5: Number of Correct Responses per Block of Trials, Percentage of Correct Responses in the Equivalence Test, Criterion Chosen in the Selection Phase, and Percentage of Correct Responses in the Nonchosen Criterion Test

	Subjects														
Tasks	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
AB	19/21	12/21 20/21	20/21	3/21 3/21 18/21 21/21	15/21 19/21	13/21 19/21	15/21 21/21	11/21 20/21	12/21 16/21 20/21	9/21 11/21 16/21 19/21	11/21 15/21 19/21	10/21 15/21 20/21	19/21	17/21 17/21 18/21 19/21	11/21 19/21
AC	14/21 17/21 21/21	19/21	19/21	16/21 18/21 21/21	14/21 21/21	19/21	20/21	15/21 20/21	20/21	19/21	19/21	18/21 21/21	16/21 21/21	14/21 21/21	16/21 21/21
ABC	35/36	36/36	35/36 33/36	36/36 36/36	33/36 34/36	34/36	36/36	35/36	35/36	36/36	36/36	36/21	36/36	35/36	36/36
CB Equivalence Test	15/15 100%	14/15 93%	12/15 80% 15/15 100%	12/15 80% 15/15 100%	10/15 67% 15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%	15/15 100%
EQ-EQ / SIM	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ	EQ-EQ
Correct Responses in Nonchosen Criterion	7/18 39%	14/18 78%	6/18 33%	7/18 39%	10/18 56%	8/18 44%	12/18 67%	11/18 61%	16/18 89%	13/18 72%	9/18 50%	11/18 61%	9/18 50%	11/18 61%	5/18 28%

The most striking results relating to this experiment appear in the selection phase, where all the subjects chose the arbitrary equivalence-equivalence criterion. For assessing the criterion not chosen during the previous phase, in this case the similarity criterion, only two of the subjects (Subjects 62 and 69) performed the task in a way that can be considered efficient (more than 75% correct answers).

The difference between the subjects of the previous experiment cannot be considered as centered on their choice of criterion, since both cases were very similar. However, the exclusion of training in the A-A relation might have had an effect on the later assessment of the similarity criterion, since the participants appeared to have ignored that reply option.

To summarize, the chart below shows the analysis of the subjects' choices (equivalence-equivalence or similarity criterion) through the five experiments performed (Figure 5).

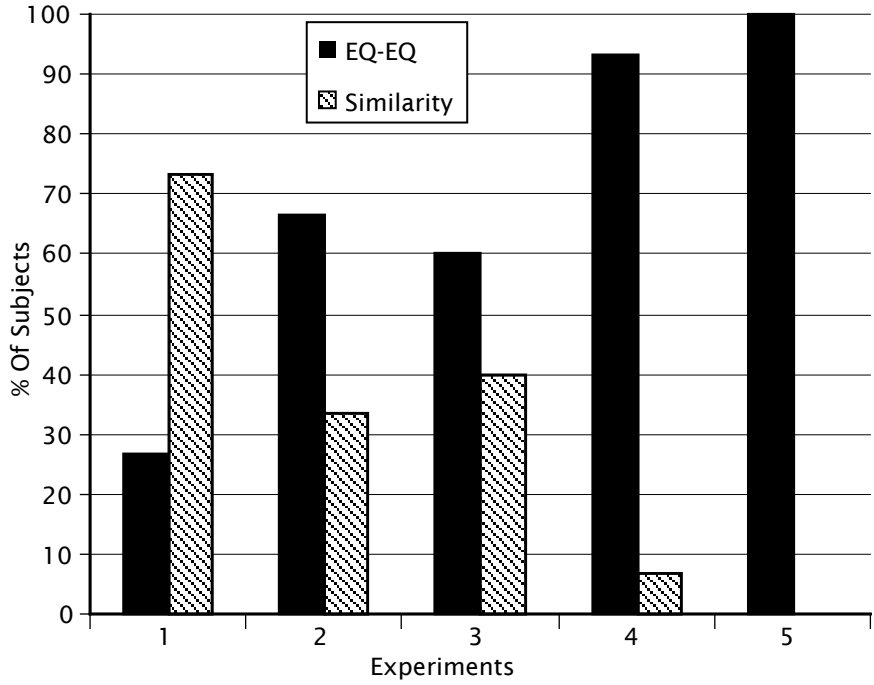

Figure 5. Percentage of subjects' choices through Experiments 1 to 5 (equivalence-equivalence: black bars; similarity: striped bars).

From the analysis of the group results relating to the choice of criterion during the competition test, three interesting conclusions arise. On the one hand, more than 70% of the subjects participating in Experiment 1 chose the physical similarity criterion, whereas 26% opted for the equivalence-equivalence criterion. As to the subjects participating in Experiments 4 and 5, 91% and 100%, respectively, chose the equivalence-equivalence criterion when both were available. A one-way analysis of variance, with the choices the subjects made (equivalence-equivalence or similarity) as the main factor, showed a significant difference between groups: $F(4, 70) = 8.34$, $p = .000$. Post hoc analyses

(the Tukey 1953 Honestly Significant Differences test, or HSD) revealed significant differences (a) between the choice of criterion made by participants in Experiments 1 and 4 (p = .000), and (b) between the criterion the subjects chose in Experiments 1 and 5 (p = .000). These are established as the groups with the highest differences in choosing the criterion (equivalence-equivalence or similarity). Further, a significantly higher number of participants joining Experiments 4 and 5 chose the equivalence-equivalence criterion during the selection phase.

General Discussion

The results reported show that the subjects will choose, with a higher probability, an arbitrary criterion based on the learned relations,if certain training conditions support it. The effect of the manipulations conducted was best observed in Experiments 4 and 5, where the large majority of the subjects clearly based their performance on an arbitrary criterion of equivalence-equivalence. First, the number of training trials had been increased in both groups. This fact points to the hypothesis of the intensity of the training in conditional discriminations as a fundamental factor for adequate performance in equivalence-equivalence-based tasks (Bohórquez et al., 2002). Appropriate training of the relations to be tested might avoid the competition effect with a nonarbitrary criterion of physical similarity between the stimuli (Stewart, Barnes-Holmes, Roche, & Smeets, 2002). A change of tendency is evident between the participants in Experiments 3 and 4 of the study, with the subjects in the latter experiment performing a larger number of training trials. This tendency towards responses based on the nonarbitrary criterion of equivalence-equivalence was maintained in all the subjects who participated in Experiment 5.

On the other hand, and following the evolution of each of the experimental groups assessed, the exclusion of reflexivity training can not be considered as having an effect on the subjects' choices, since the differences between the subjects participating in Experiments 4 and 5 are not significant in that sense. However, from the moment the equivalence test was included, more of the subjects began to show equivalence-equivalence responding, based on the trained relations (if the choices of the subjects participating in Experiment 1 are compared with choices in Experiment 2 and subsequent experiments). In any case, evidence shows no need to carry out the equivalence test before the test of the equivalence-equivalence relations to obtain good performance in this last criterion. Barnes et al. (1997) successfully tested relations based on equivalence-equivalence responding before performing an equivalence test and concluded that this last test is not an indispensable prerequisite. Still et al. (2001) reported that only half the subjects passed a BC-BC test when it was presented before the equivalence test. Moreover, Carpentier et al. (2003a) also found deteriorated BC-BC performances in the groups where the equivalence test was presented last. The equivalence test may be considered a required element that facilitates more complex subsequent relations, although it cannot be considered a prerequisite. Even when it doesn't seem necessary to test equivalence relations, the number of subjects demonstrating equivalence-equivalence responding in the present study increased when Experiments 1 and 2 were compared. Some experimental data has been reported about novel conditional responding from repeated testing without any kind of explicit reinforcement (Spradlin, Cotter, & Baxley, 1973). That is the case of the delayed emergence

of derived conditional discriminations (see Sidman, 1994); perhaps subjects learn something about equivalence-equivalence relations during the equivalence test.

The criterion of similarity, presented from a first moment together, with an arbitrary criterion based on trained equivalence relations, appears to acquire more strength when these relations have not been learned correctly. The nonarbitrary criterion of similarity has developed its importance as a general criterion through the history of reinforcement of the subject. It is acquired considerably before the rest of the properties (symmetry and transitivity) necessary for a performance based on equivalence. (For a review of the acquisition of the symmetric response, see Boelens & Van der Broek, 2000; Boelens, Van der Broek, & Calmeyn, 2003; García & Benjumea, 2001, 2006; Luciano, Herruzo, & Barnes-Holmes, 2001.) This is probably why subjects who do not find a stable criterion with which to respond will resort to the nonarbitrary criterion by which they have responded successfully in the majority of their previous interactions.

The types of stimuli used in the results of the subjects'choices don't appear to be an influence, since there were no appreciable differences between Experiments 2 and 3. There are antecedents for quite good performances in equivalence-equivalence tasks using both syllables (Barnes-Holmes, Hegarty, & Smeets, 1997; Stewart, Barnes-Holmes, Roche, & Smeets, 2001, 2002); color images (Stewart et al., 2001, 2002; Stewart et al., 2002); and black and white images (Carpentier, Smeets, & Barnes-Holmes, 2002; García, Bohórquez, Gómez, Gutiérrez, & Pérez, 2001; García, Gutiérrez, Bohórquez, Gómez, & Pérez, 2002). These groups shared the same characteristics except for the type of stimuli used. In Experiment 2, the training was carried out with all the designed figures (Figure 1), and in Experiment 3 meaningless syllables were used that were also created for the purpose (Figure 4). As to the number of trials necessary to pass the training in each of these groups, no substantial differences were found (111 and 131 trials on average, respectively). Similarly, no effects were found for the counterbalance performed at the moment of assigning the stimuli to the different classes of equivalence.

The present study creates a relatively stable baseline for performances based on equivalence-equivalence responding. This is true exclusively under certain conditions, all of which were manipulated during the training of the conditional discriminations A-A, A-B, and A-C. When equivalence relations training is intensive (e. g., using 21 trials per block in the present study) and an equivalence test is carried out before an equivalence-equivalence test is performed, the subjects will respond with more probability on the basis of an arbitrary criterion in which they have to relate equivalence relations with other equivalence relations. This is true even when a nonarbitrary criterion based on similarity is competing, probably chosen in conditions in which (a) the reflexive property is explicitly trained, and (b) the relations acquired are not manifested through an equivalence test. An interesting conclusion related to reflexivity training can be found in the assessment of the criterion (equivalence-equivalence or physical similarity) not chosen during the competition test. The withdrawal of the A-A training block may cause a smaller number of subjects to perform the similarity test correctly after having chosen the equivalence-equivalence criterion. In Experiment 4 (which includes A-A training), 10 of the subjects performed in a way that can be considered more than by chance, compared with three subjects in Experiment 5 (in which the A-A training had been withdrawn),

with this criterion considered as responding to 14 of the 18 trials of the test. An argument might be made that despite the existence of this criterion and the subjects' ease in identifying it, the withdrawal of the reflexivity training caused subjects to try to respond using an arbitrary criterion based more on the trained equivalence relations than on the nonarbitrary criterion of similarity. This, on the other hand, is the only one that existed in the second part of the test trials for the groups that performed these choices.

However, another explanation would be worth considering for the effect found when the reflexivity training was withdrawn. The behavior observed in the subjects in Experiment 5 during the choice test (equivalence-equivalence and similarity in a situation of competition) may be interpreted as a selection (their behavior is guided by an arbitrary criterion), but also as a rejection (Carrigan & Sidman, 1992; Sidman, 1987). That the subjects chose the choice opposed to similarity, does not mean that they followed, or even identified, the equivalence-equivalence criterion. This interpretation earns force if the poor performance of the subjects carrying out the similarity test alone is considered. This result suggests that they continued responding in opposition to the similarity criterion and therefore chose the comparison sharing no element with the sample.

As suggested earlier, this pattern of response may originate in the withdrawal of the reflexivity training. During this variant of the training procedure (Experiment 5), the subjects never have to match stimuli depending on physical properties, which may lead to rejection of this response criterion when it is available. The data obtained in Experiment 4 (in which A-A training was included) does not appear to support this hypothesis, since the number of subjects choosing the arbitrary criterion in the competition situation was very similar to that obtained in Experiment 5. Nevertheless, the results in the similarity test showed a number of differences. Whereas in Experiment 4, 11 of the 14 subjects who chose equivalence-equivalence passed the similarity test, in Experiment 5 only 2 out of 15 did so. The subjects joining Experiment 4 clearly did not respond in opposition to the similarity criterion, whereas those in Experiment 5 may have done so.

One way of checking whether the subjects were really following the equivalence-equivalence criterion (and not the nonsimilarity criterion) might have been to expose them to a test in which only this response criterion was available, and not just to one in which the only possible criterion was the opposite of the one chosen in the competition situation. Future studies will check on the effectiveness of the procedure developed here in situations in which the effect of the response by opposition is controlled.

References

BARNES, D., HEGARTY, N., & SMEETS, P. M. (1997). Relating equivalence relations to equivalence relations: A relational framing model of complex human functioning. *The Analysis of Verbal Behavior, 14,* 57–83.

BOELANS, H., & VAN DER BROEK, M. (2000). Influencing children's symmetric responding in matching-to-sample tasks. *The Psychological Record*, *50*, 655–670.

BOELANS, H., VAN DER BROEK, M., & CALMEYN, S. (2003). Is children's symmetric matching to sample the product of symmetric experiences with spoken names? *The Psychological Record*, *53*, 593–616.

BOHORQUEZ., C., GARCIA, A., GUTIERREZ, M. T., GOMEZ, J., & PEREZ, V. (2002). Efecto del entrenamiento en reflexividad y la evaluación de equivalencia en la competencia entre relaciones arbitrarias y no arbitrarias en el paradigma de equivalencia-equivalencia. [Effect of training on reflexivity and the assessment of equivalence in competition between arbitrary and non-arbitrary relations in the equivalence-equivalence paradigm]. *International Journal of Psychology and Psychological Therapy, 2,* 41–56.

BUSH, K. M., SIDMAN, M., & DE ROSE, T. (1989). Contextual control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior, 51*, 29–45.

CARPENTIER, F., SMEETS, P. M., & BARNES-HOLMES, D. (2000). Matching compound samples with unitary comparisons: Derived stimulus relations in adults and children. *The Psychological Record, 50*, 671–685.

CARPENTIER, F., SMEETS, P. M., & BARNES-HOLMES, D. (2002a). Establishing transfer of compound control in children: A stimulus control analysis. *The Psychological Record, 52,* 139–158.

CARPENTIER, F., SMEETS, P. M., & BARNES-HOLMES, D. (2002b). Matching functionally same relations: Implications for equivalence-equivalence as a model for analogical reasoning. *The Psychological Record, 52*, 351–370 .

CARPENTIER, F., SMEETS, P. M., & BARNES-HOLMES, D. (2003a). Equivalence-equivalence: Matching stimuli with same discriminative functions. *The Psychological Record, 53,* 145–162.

CARPENTIER, F., SMEETS, P. M., & BARNES-HOLMES, D. (2003b). Equivalence-equivalence as a model of analogy: Further analyses. *The Psychological Record, 53,* 349–372.

CARPENTIER, F., SMEETS, P. M., BARNES-HOLMES, D., & STEWART, I. (2004). Matching derived functionally-same relations: Equivalence-equivalence and classical analogies. *The Psychological Record, 54,* 255–273.

CARR, D., & BLACKMAN, D. E. (1996). Equivalence relations, naming, and generalized symmetry. *Journal of the Experimental Analysis of Behavior, 65*, 245–247.

CARRIGAN, P. F., JR., & SIDMAN, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58,* 183–204.

COWLEY, B. J., GREEN, G., & BRAUNLING-MCMORROW, D. (1992). Using stimulus equivalence procedures to teach name-face matching to adults with brain injuries. *Journal of the Experimental Analysis of Behavior, 25*, 461–475.

DENAVY, J. M., HAYES, S. C., & NELSON, R. O. (1986). Equivalence class formation in language-able and language-disable children. *Journal of the Experimental Analysis of Behavior, 46,* 243–257.

DE ROSE, J. C., SOUZA, D. G., ROSSITO, A. L., & DE ROSE, T. (1992). Stimulus equivalence and generalization in reading after matching to sample by exclusion. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 69–82). Reno, NV: Context Press.

FIELDS, L., VERHAVE, T., & FATH, S. J. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of the Experimental Analysis of Behavior, 42,* 143–157.

GARCIA, A., & BENJUMEA, S. (2001). Pre-requisitos ontogenéticos para la emergencia de relaciones simétricas [ontogenetic prerequirements for the emergence of symmetrical relationships]. *International Journal of Psychology and Psychological Therapy, 1,* 115–136.

GARCIA, A., & BENJUMEA, S. (2002). Orígenes, ampliación y aplicaciones de la equivalencia de estímulos. [origins, widening and applications of the equivalence of stimuli]. *Apuntes de Psicología, 20,* 171–186.

GARCIA, A., & BENJUMEA, S. (2006). The emergence of symmetry in a conditional discrimination task using different responses as propioceptive samples in pigeons. *Journal of the Experimental Analysis of Behavior, 86,* 65–80.

GARCIA, A., BOHORQUEZ, C., GOMEZ, J., GUTIERREZ, M. T., & PEREZ, V. (2001). Ensombrecimiento entre relaciones arbitrarias y no arbitrarias en el paradigma de equivalencia-equivalencia. [Overshadowing between arbitrary and non-arbitrary relations in the equivalence-equivalence paradigm]. *Suma Psicológica, 8*, 251–270.

GARCIA, A., GUTIERREZ, M. T., BOHORQUEZ, C., GOMEZ, J., & PEREZ, V. (2002). Competencia entre relaciones arbitrarias y relaciones no arbitrarias en el paradigma de equivalencia-equivalencia. [Competition between arbitrary and non-arbitrary relations in the equivalence-equivalence paradigm]. *Apuntes de Psicología, 20*, 205–224.

GARCIA, A., GOMEZ, J., GUTIERREZ, M. T., & PUCHE, A. (2001). Formación y ampliación de clases de equivalencia aplicadas al tratamiento de un niño autista. [Training and widening of equivalence classes applied to the treatment of an autistic child.] *Análisis y Modificación de Conducta, 27, 114*, 649–669.

GARCIA, A., GOMEZ, J., PEREZ, V., BOHORQUEZ, C., & GUTIERREZ, M. T. (2003). Efectos del orden de presentación entre criterios de respuesta basados en relaciones de semejanza y de equivalencia-equivalencia. [Effects of the order of presentation between response criteria based on similarity and equivalence-equivalence relations]. *Acción Psicológica, 2*, 239–249.

GOMEZ, J., GARCIA, A., PEREZ, V., GUTIERREZ, M. T., & BOHOROQUEZ, C. (2004). Aportaciones del análisis conductual al estudio de la conducta emergente: algunos fenómenos experimentales. [Contributions of behaviour analysis to the study of emerging behaviour: Some experimental phenomena]. *International Journal of Psychology and Psychological Therapy, 4*, 161–191.

HALL, S. S., DEBERNARDIS, G. M., & REISS, A. L. (2006). The acquisition of stimulus equivalence in individuals with fragile X syndrome. *Journal of Intellectual Disability Research, 50,* 643–651.

HAYES, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes y P. N. Chase (Eds.), *Dialogues on Verbal Behavior* (pp. 19–40). Reno, NV: Context Press.

HORNE, P. J., & LOWE, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, 65, 185–241.

KAMIN, L. J. (1968). "Attention-like" processes in classical conditioning. In M.R. Jones (Ed.), *Miami Symposium on the Prediction of Behavior: Aversive stimulation*. Miami: University of Miami Press.

KAMIN, L. J. (1969). Predictability, surprise, attention, and conditioning. In B.A. Campbell & R.M. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts.

LAZAR, R. (1977). Extending sequence-class membership with matching to sample. *Journal of the Experimental Analysis of Behavior, 27*, 381–392.
LIPKENS, R., HAYES, S. C., & HAYES, L. J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology, 56,* 201–239.
LUCIANO, M. C., HERRUZO, J., & BARNES-HOLMES, D. (2001). Generalization of say-do correspondence. *The Psychological Record, 51*, 111–131.
LYNCH, D. C., & CUVO, A. J. (1995). Stimulus equivalence instruction of fraction-decimal relations. *Journal of the Experimental Analysis of Behavior, 28*, 115–126.
MAGUIRE, R. W., STROMER, R., MACKAY, H. A., & DEMIS, C. A. (1994). Matching to complex samples and stimulus class formation in adults with autism and young children. *Journal of Autism and Developmental Disorders, 24*, 753–772.
MARKAM, M. R., & DOUGHER, M. J. (1993). Compound stimuli in emergent stimulus relations: Extending the scope of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 60,* 529–542.
MARKAM, M. R., DOUGHER, M. J., & AUGUSTSON, E. M. (2002). Transfer of operant discrimination and respondent elicitation via emergent relations of compound stimuli. *The Psychological Record, 52,* 325–350.
MAYDAK, M., STROMER, R., MACKAY, H. A., & STODDARD, L. T. (1995). Stimulus classes in matching to sample and sequence production: The emergence of numeric relations. *Research in Developmental Disabilities, 16,* 179–204.
O'DONNELL, J., & SAUNDERS, K. J. (2003). Equivalence relations in individuals with language limitations and mental retardation. *Journal of the Experimental Analysis of Behavior, 80,* 131–157.
PEREZ-GONZALEZ, L. A. (1994). Transfer of relational stimulus control in conditional discriminations. *Journal of the Experimental Analysis of Behavior, 61*, 487–503.
PEREZ-GONZALEZ, L. A., & MORENO-SERRANO, V. (1999). Formación de clases de equivalence en ancianos. [Formation of equivalence classes in old people.] *Psicothema, 11*, 325–336.
PILGRIM, C., CHAMBERS, L., & GALIZIO, M. (1995). Reversal of baseline relations and stimulus equivalence: II. Children. *Journal of the Experimental Analysis of Behavior, 63*, 239–254.
SAUNDERS, K. J., & SPRADLIN, J. E. (1993). Conditional discrimination in mentally retarded subjects: Programming acquisition and learning set. *Journal of the Experimental Analysis of Behavior, 60,* 571–585.
SCHENK, J. J. (1993). Emergent conditional discriminations in children: Matching to compound stimuli. *The Quarterly Journal of Experimental Psychology, 46B*, 345–365.
SIDMAN, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14*, 5–13.
SIDMAN, M. (1986a). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213–245). Hillsdale, NJ: Lawrence Erlbaum Associates.
SIDMAN, M. (1987). Two choices are not enough. *Behavior Analysis, 22*, 11–18.
SIDMAN, M. (1990b). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behavior analysis in theory and practice: Contributions and controversies* (pp. 93–114). Hillsdale, NJ: Lawrence Erlbaum Associates.

SIDMAN, M. (1994). *Equivalence relations and behavior: A research history*. Boston: Authors Cooperative.

SIDMAN, M., CRESSON, O., & WILLSON-MORRIS, M. (1974). Acquisition of matching to sample via mediated transfer. *Journal of the Experimental Analysis of Behavior, 22,* 261–273.

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.

SMEETS, P. M., SCHENK, J. J., & B, D. (1995). Establishing arbitrary stimulus classes via identity-matching training and non-reinforced matching with complex stimuli. *The Quarterly Journal of Experimental Psychology, 48B*, 311–328.

SPRADLIN, J. E., COTTER, V. W. Y BAXLEY, N. (1973). Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. *American Journal of Mental Deficiency, 77,* 556–566.

STEWART, I., BARNES-HOLMES, D., ROCHE, B., & SMEETS, P. M. (2001). Generating derived relational networks via the abstraction of common physical properties: A possible model of analogical reasoning. *The Psychological Record, 51,* 381–408.

STEWART, I., BARNES-HOLMES, D., HAYES, S. C., & LIPKENS, R. (2001). Relations among relations: Analogies, metaphors, and stories. In S. C. Hayes, D. Barnes-Holmes & B. Roche (Eds.), *Relational Frame Theory: A post-Skinnerian account of human language and cognition* (pp. 73-86). New York: Plenum.

STEWART, I., BARNES-HOLMES, D., ROCHE, B., & SMEETS, P. M. (2002). Stimulus equivalence and non-arbitrary relations. *The Psychological Record,* 5, 77–88.

STEWART, I., BARNES-HOLMES, D., ROCHE, B., & SMEETS, P. M. (2002). A functional-analytic model of analogy: a relational frame analysis. *Journal of the Experimental Analysis of Behavior*, *78*, 375–396.

STROMER, R., MACKAY, H. A., & STODDARD, L. T. (1992) Classroom applications of stimulus equivalence technology. *Journal of Behavioral Education, 2*, 225–256.

STROMER, R., & OSBOURNE, J. G. (1982). Control of adolescents' arbitrary matching-to-sample by positive and negative stimulus relations. *Journal of the Experimental Analysis of Behavior, 37,* 329– 348.

STROMER, R., & STROMER, J. B. (1990a). The formation of arbitrary stimulus classes in matching to complex samples. *The Psychological Record, 40*, 51–66.

STROMER, R., & STROMER, J. B. (1990b). Matching to complex: Further study of arbitrary stimulus classes. *The Psychological Record, 40*, 505–516.

TUKEY, J. (1953). *The Problem of Multiple Comparisons*. Princeton University. Princeton NJ.

WULFERT, E., & HAYES, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50,* 125–144.

ZENTALL, T. R., GALIZIO, M., & CRITCHFIELD, T. S. (2002). Categorization, Concept learning, and behavior analysis: An introduction. *Journal of the Experimental Analysis of Behavior*, *78*, 237–248.