

Emergent relations in pigeons following training with temporal samples

Edson M. Huziwara · Saulo M. Velasco ·
Gerson Y. Tomanari · Deisy G. de Souza ·
Armando D. Machado

© Psychonomic Society, Inc. 2012

Abstract In two experiments, we investigated emergent conditional relations in pigeons using a symbolic matching-to-sample task with temporal stimuli as the samples and hues as the comparisons. Both experiments comprised three phases. In Phase I, pigeons learned to choose a red keylight (R) but not a green keylight (G) after a 1-s signal. They also learned to choose G but not R after a 4-s signal. In Phase II, correct responding consisted of choosing a blue keylight (B) after a 4-s signal and a yellow keylight (Y) after a 16-s signal. Comparisons G and B were both related to the same 4-s sample, whereas comparisons R and Y had no common sample. In Phase III, R and G were presented as samples, and B and Y were presented as the comparisons. The choice of B was correct following G, and the choice of Y was correct following R. If a relation between comparisons that shared a common sample were to emerge, then responding to B given G would be more likely than responding to Y given R. The results were generally consistent with this prediction, suggesting, for the first time in pigeons, the emergence of novel relations that involve temporal stimuli as nodal samples.

Keywords Emergent relations · Temporal stimuli · Pigeons

E. M. Huziwara · S. M. Velasco · G. Y. Tomanari
Universidade de São Paulo, São Paulo, Brazil

D. G. de Souza
Universidade Federal de São Carlos, São Carlos, Brazil

E. M. Huziwara · S. M. Velasco · G. Y. Tomanari · D. G. de Souza
Instituto Nacional de Ciência e Tecnologia sobre Comportamento,
Cognição e Ensino, São Paulo, Brazil

A. D. Machado
Universidade do Minho, Minho, Portugal

E. M. Huziwara (✉)
Rua Corrego da Mata, n. 590, apto. 901, Bloco 1,
Belo Horizonte, 31030-030 MG, Brazil
e-mail: huziwara.edson@gmail.com

In a matching-to-sample (MTS) procedure, two stimuli from Set A (e.g., A1 and A2) and two stimuli from Set B (e.g., B1 and B2) may be conditionally related as follows. In each trial, one stimulus from Set A is presented as the sample, and both stimuli from Set B are presented as comparisons. Given A1, responses to B1 but not B2 are arbitrarily defined as correct. Given A2, responses to B2 but not B1 are arbitrarily defined as correct (i.e., AB training). Next, a second conditional discrimination can be trained with stimuli from Sets B and C. In this case, the correct responses are (1) C1 but not C2 given sample B1 and (2) C2 but not C1 given sample B2 (BC training). Correct responses are reinforced, and incorrect responses are extinguished.

After the AB and BC training, a subject may exhibit conditional discriminations that were not directly trained. For example, a subject may match the same pair of previously related stimuli when their sample–comparison roles are reversed. Thus, having learned the AB and BC relations, the subject correctly matches the stimuli in the symmetry tests (i.e., BA and CB). A subject may also correctly match the samples from the AB training to the comparisons from the BC training (i.e., AC transitivity tests). This pattern of responding demonstrates the emergence of an untrained conditional relation between two stimulus sets, A and C, that had never been paired directly but had been related to a common stimulus set, B.

The experimental procedure under which the comparison stimuli from AB training become the sample stimuli during BC training is called a *linear-series* training structure. In addition to the linear series, two other experimental arrangements could be used to organize a training procedure. In a *many-to-one* training structure, the samples come from two or more stimulus sets, but the comparisons come from a single stimulus set (e.g., AB and CB training). Conversely, in a *one-to-many* training structure, the samples come from a single stimulus set, but the comparisons come from two or more stimulus sets (e.g., BA and BC training). In both cases,

emergent relations could be evaluated through transitivity and symmetry tests (e.g., Green & Saunders, 1998).

Sidman and Tailby (1982) proposed combined tests to evaluate emergent relations. According to their proposal, an emergent relation between two comparisons following a one-to-many training structure could be considered evidence of the simultaneous emergence of more than one relational property. Following AB and AC training, for example, one could infer the emergence of transitivity and symmetry by testing BC and CB (i.e., combined tests). The argument has two steps. First, from the trained AB and AC relations, the BA and CA relations emerge, demonstrating symmetry. Second, from the (emergent) BA and (trained) AC relations, the BC relation emerges. Similarly, from the (emergent) CA and (trained) AB relations, the CB relation emerges, demonstrating transitivity. Thus, symmetry and transitivity can both be inferred from the emergence of an untrained relation between comparisons mediated by a common sample.

Sidman (1990) went one step farther by proposing that these combined tests can also be used to evaluate reflexivity. This relational property consists of matching each stimulus to itself (i.e., AA, BB, and CC identity tests). For example, after AB and AC training, the emergence of the BC relation also implies reflexivity, because Stimulus Set B must be recognized as the same when it is presented as samples (i.e., test trials) and comparisons (i.e., training trials).

When exposed to combined tests, verbally able humans have yielded systematically positive results (e.g., Devany, Hayes, & Nelson, 1986; Lazar, Davis-Lang, & Sanchez, 1984; Sidman, 1971; Sidman & Cresson, 1973; Sidman & Tailby, 1982; Spradlin, Cotter, & Baxley, 1973; Spradlin & Saunders, 1986). Nonhuman subjects, however, have yielded mostly negative results (e.g., Hogan & Zentall, 1977; Lionello-DeNolf & Urcuioli, 2002; Lipkens, Kop, & Matthijs, 1988; Sidman et al., 1982; Tomonaga, Matsuzawa, Fujita, & Yamamoto, 1991; Yamamoto & Asano, 1995). Reflexivity (Sweeney & Urcuioli, 2010), symmetry (Frank & Wasserman, 2005; Garcia & Benjumea, 2006; Urcuioli, 2008; Vasconcelos & Urcuioli, 2011; Velasco, Huziwara, Machado, & Tomanari, 2010; Yamamoto & Asano, 1995), and transitivity (D'Amato, Salmon, Loukas, & Tomie, 1985; Kuno, Kitadate, & Iwamoto, 1994) have been reported in separate experiments with nonhuman animals. However, only one study with a California sea lion has obtained positive results with combined tests (Schusterman & Kastak, 1993). Specifically, after a history of reinforced symmetrical responding (i.e., multiple-exemplar training with an AB + BC linear series arrangement), the sea lion passed the CA equivalence test.

The difficulties in showing emergent relations with animals could stem from biological constraints shared by all animals—namely, limitations in linguistic ability (cf.

Devany et al., 1986; Hayes, 1989; Horne & Lowe, 1996). Those difficulties may also conceivably stem from the procedures used in the original studies, in which one or more of their features may have prevented the animals from learning or expressing emergent relations (Dube, McIlvane, Callahan, & Stoddard, 1993). If the latter argument is true, researchers should continue to explore ways of training and testing emergent relations in animals that circumvent the problems with the current procedures.

In the present study, we followed a new approach to test emergent conditional relations. Proposed by Velasco et al. (2010), the approach compares the acquisition of two conditional relations, one between two stimuli presumably associated by prior training, and another between two stimuli presumably not associated by prior training. The four stimuli are all equally familiar to the subjects. Because the two conditional relations are taught during reinforced trials, the approach of Velasco et al. solves some of the problems commonly found when test trials are performed in extinction. According to some authors (cf. Antonitis, 1951; Dube & McIlvane, 1996; Epstein, 1983, 1985; Galvão, Calcagno & Sidman, 1992; Kuno et al., 1994; Lerman & Iwata, 1996; Schusterman & Kastak, 1993; Sidman et al., 1982; Wilson & Hayes, 1996), extinction during test trials is similar to the consequences of incorrect choices during training trials, in that in both cases the reinforcer is omitted. If this similarity changes the sources of stimulus control established during prior training, the subjects are likely to fail the tests for emergent relations (see Dube & McIlvane, 1996; Galvão et al., 1992; Kuno et al., 1994; Sidman et al., 1982). Because the approach of Velasco et al. compares the speed of learning of two conditional relations within subjects, it increases the study's sensitivity with respect to the more common between-subjects design (see Velasco et al., 2010, for further discussion of design issues in tests of emergent relations).

The present study introduced another novelty to the study of emergent relations in animals—specifically, the use of a cross-modal MTS task with temporal stimuli as samples and hues as comparisons. Time is a fundamental property of the environment in which behavior occurs (Catania, 1991), and many animals show some form of behavioral sensitivity to time (e.g., Richelle & Lejeune, 1980). Additionally, temporal stimuli may be particularly salient (e.g., Staddon, 2001); therefore, they may reduce extraneous sources of stimulus control in the MTS task. For these reasons, we asked whether temporal stimuli would be particularly effective as nodal elements, mediating the association between two nontemporal stimuli and enabling the emergence of new relations.

To summarize, following the combined-tests approach of Sidman and Tailby (1982), and using the procedure of Velasco et al. (2010) of reinforced test trials in a within-subjects design, in the present study we examined whether pigeons would show an emergent relation

between two visual stimuli previously related to a common temporal stimulus.

The present report contains two experiments, each comprising three phases (Table 1). In Phase I, pigeons learned to choose a red keylight (R) but not a green keylight (G) after a 1-s signal, and G but not R after a 4-s signal. In Phase II, the pigeons learned to choose a blue keylight (B) but not a yellow keylight (Y) after a 4-s signal, and Y but not B after a 16-s signal. Notably, the G and B comparisons were correct following the same samples (4 s), whereas the R and Y comparisons were correct following different samples (1 and 16 s, respectively). In Phase III, to test for the emergence of a direct relation between the G and B hues, the R and G hues were presented as samples, and the B and Y hues were presented as comparisons. Choices of B but not Y were reinforced following G, and choices of Y but not B were reinforced following R. The emergence of an untrained relation between the G and B comparisons previously related to the same sample would be indicated by (1) response accuracy above chance on GB trials and near chance on RY trials during the first session of Phase III or (2) faster acquisition of the GB relation than of the RY relation across the sessions of Phase III.

However, a potential problem can be foreseen with the experiment summarized in Table 1. During Phase III, we wanted to assess how the sample hues would control choice; however, in addition to hue, the samples would inevitably have some duration. Given that the pigeons learned to choose the comparisons during Phases I and II on the basis of sample durations, the sample duration might mask any differential effects of sample hue (e.g., Honig & Urcuioli, 1981; Oliveira & Machado, 2008). Thus, in Experiment 1 we examined how sample duration would affect choice. The results of Experiment 1 suggested that the duration associated with the sample hues indeed could bias responding during Phase III. In Experiment 2, control by sample duration was minimized in order to unmask any potential control by the sample's hue, thereby assessing the hypothesized emergence of an untrained conditional relation.

Experiment 1

To reveal any potential control by sample hue during Phase III, ideally one should eliminate all of the other competing sources of stimulus control, including sample duration. Unfortunately, sample duration cannot be eliminated. Therefore, in Experiment 1 we varied sample duration to determine how, in this MTS procedure, it would control choice during Phase III. On the basis of the results, we determined the sample duration that would minimize temporal control and then used that sample duration in Experiment 2.

To understand how control by sample duration may be revealed during Phase III, consider the case in which *only* sample duration affects choice. Regardless of sample hue, when the sample lasts 4 s, the pigeons will be more likely to choose B than Y, because during Phase II the choice of B was reinforced and the choice of Y was extinguished after the 4-s samples. When the sample duration lasts 16 s, however, the pigeons will be more likely to choose Y than B, because during Phase II the choice of Y was reinforced and the choice of B was extinguished after the 16-s samples (e.g., Machado & Arantes, 2006; Machado & Keen, 1999; Machado & Pata, 2005). To predict the results for other sample durations, we would need to consider the effects of temporal generalization (e.g., samples shorter than 4 s should also yield a preference for B over Y).

Other cases are obviously possible. For example, because the sample in Phase III consists of a keylight hue presented for the first time on the center key, temporal discrimination might be disrupted, and the preference for the B and Y comparisons might fall to indifference. In this case, the pattern of choices would not vary with sample duration. Finally, and perhaps more realistically, both sample hue and duration might influence choice, albeit to different degrees (e.g., Oliveira & Machado, 2008). In this case, the pattern of choices as a function of sample duration would also help determine the sample duration that would minimize differential temporal control, and that value could be used in Experiment 2.

Table 1 Samples and comparisons during each experimental phase. S+ and S− identify the correct and incorrect comparisons, respectively, following each sample stimulus

Phase I			Phase II			Phase III		
Comparisons			Comparisons			Comparisons		
Samples	Red	Green	Samples	Blue	Yellow	Samples	Blue	Yellow
1 s	S+	S−	4 s	S+	S−	Red	S−	S+
4 s	S−	S+	16 s	S−	S+	Green	S+	S−

Method

Subjects

Eight adult pigeons (*Columba livia*) without experience in temporal discrimination tasks were used in Experiment 1. All of the pigeons were individually housed in stainless-steel cages with water and grit continuously available. A 13-h:11-h light:dark cycle was in effect in the colony room, with lights on at 8:00 am. The pigeons were maintained at 80 % of their free-feeding body weights throughout the experiment.

Apparatus

Two Lehigh Valley chambers, measuring $32 \times 36 \times 34$ cm, were used. Three circular response keys, each with a 2.5-cm diameter, were located 22 cm above the chamber floor and spaced 4.5 cm apart, center to center. The keys could be illuminated with red, green, blue, and yellow lights. The center key could also be illuminated with a white light. A food hopper was accessible through a 6×5 cm opening 8.5 cm below the central key. A 7.5-W houselight located 25 cm above the wall opposite the intelligence panel provided general illumination. A computer programmed in the C++ language controlled the experimental events and recorded the data.

Procedure

As Table 2 shows, two mappings between the sample durations and the comparison hues were used during

Table 2 Sample–comparison mappings used in Experiments (Exp.) 1 and 2

Exp.	Group	Pigeon	Phase I		Phase II		Phase III	
			1 s	4 s	4 s	16 s	R	G
1	1 s	P201	G	R	B	Y	B	Y
		P300	G	R	B	Y	B	Y
		P440	R	G	B	Y	Y	B
		P726	R	G	B	Y	Y	B
	16 s	P271	G	R	B	Y	B	Y
		P435	G	R	B	Y	B	Y
		P454	R	G	B	Y	Y	B
		P619	R	G	B	Y	Y	B
2	8 s	P366	G	R	B	Y	B	Y
		P604	G	R	B	Y	B	Y
		P746	R	G	B	Y	Y	B
		P876	R	G	B	Y	Y	B

During Phases I and II, the sample was a white light 1, 4, or 16 s in duration. During Phase III, the sample was a red (R) or green (G) light. The Group column shows the minimum sample durations during Phase III. B, blue; Y, yellow

Phase I. For half of the pigeons, the mapping was “1 s→R, 4 s→G” and for the other half, the mapping was “1 s→G, 4 s→R.” During Phase II, the mapping was “4 s→B, 16 s→Y” for all pigeons. Therefore, Phase III had two sets of trained relations: R→B and G→Y or G→B and R→Y. However, for clarity, we will describe the remainder of the procedure and the experimental results as if all of the pigeons had the mappings used in Table 1 (“1 s→R, 4 s→G” in Phase I, “4 s→B, 16 s→Y” in Phase II, and G→B and R→Y in Phase III).

Phase I (see Table 1) Given a choice between R and G, the pigeons learned to choose R after a 1-s signal and G after a 4-s signal. Each trial began with the offset of the houselight and onset of the central key with a white light. Pecks at the center key during the sample had no programmed consequences. At the end of the sample (1 or 4 s), the center key was turned off, and the side keys were illuminated with R and G keylights. The left–right location of the keylights varied randomly across trials, but with the constraint that at the end of the session both keylights had been presented equally often on both keys. One peck at any of the side keys turned all of the lights off. If the choice was correct, the food hopper was raised and illuminated, and the pigeon had access to grain for 2 s. After the food delivery, a 30-s intertrial interval (ITI) with the houselight on followed. If the choice was incorrect, the ITI followed immediately, and the trial was repeated (i.e., correction procedure). If three consecutive errors occurred, only the correct comparison was presented during the choice period in the fourth trial (i.e., forced choice). Sessions ended after 40 reinforcers, and training continued until the pigeon achieved at least 85 % correct responses in each of the two conditional relations over five consecutive sessions.

Phase II Given a choice between B and Y, the pigeons learned to choose B after a 4-s signal and Y after a 16-s signal. All of the other experimental details were the same as in Phase I, including the correction procedure. Next, the trials used during Phases I and II were mixed. Each session comprised 40 trials, of which ten trials of each conditional relation already trained (i.e., 1 s→R, 4 s→G, 4 s→B, and 16 s→Y) were presented in a pseudorandom order. The learning criterion was set at 80 % correct responses for each of the four conditional relations over five consecutive sessions.

Phase III The R and G keylights were presented on the center key as samples, and the B and Y keylights were presented on the side keys as comparisons. The choice of B was reinforced following the G sample (i.e., GB relation), and the choice of Y was reinforced following the R sample (i.e., RY relation). Thus, during Phase III, the pigeon’s choices were reinforced depending on the sample hue. The

reinforcement parameters and the correction procedure were the same as during the previous phases.

The pigeons were divided into two groups, Group 1S and Group 16S, each defined by the contingency with respect to the sample duration. For Group 1S, when the center keylight was illuminated with the sample hue, at least 1 s had to elapse and at least one peck on the center key had to be emitted (i.e., a conjunctive fixed-time [FT] 1-s fixed-ratio 1 [FR1] schedule) before the sample keylight was turned off and the side keys were illuminated with the comparison keylights. For Group 16S, the center keylight was turned on for at least 16 s, and the pigeon had to peck the key at least once (i.e., a conjunctive FT 16-s FR1 schedule) to produce the comparison stimuli. The conjunctive schedule ensured that the pigeon observed the sample (FR1) and that its duration would be relatively short for Group 1S and relatively long for Group 16S. Phase III lasted six sessions.

If sample duration strongly controls choice, at least during the first session of Phase III, we predicted that Group 1S would choose B more frequently than Y and that Group 16S would choose Y more frequently than B, *regardless of the sample hue*. Therefore, for Group 1S, the GB relation would yield a higher percentage of correct responses than the RY relation, but the opposite would be the case for Group 16S.

Results and discussion

All of the pigeons learned the four baseline conditional discriminations. As Table 3 shows, the average numbers of sessions to meet the learning criteria were 35.9 (range, 12–59) for the first two discriminations in Phase I, 37.4 (range, 19–67) for the two new discriminations in Phase II, and 20.8 (range, 5–54)

Table 3 Average numbers of sessions (Sess.), trials per session, and percentages correct for each discrimination during Experiments (Exp.) 1 and 2

Exp.	Phase	Sess.	Trials ^a	Percentage Correct			
				1 s→R	4 s→G	4 s→B	16 s→Y
1	I	35.9	53.8	91.8	91.8		
	II	37.4	60.0			92.9	92.6
	I/II	20.8	46.1	88.1	90.9	92.9	95.1
2	I	26.5	55.4	91.3	91.5		
	II	25.3	56.5			93.5	92.8
	I/II	15.8	46.4	95.5	90.5	95.0	95.5

Phase “I/II” represents the sessions in which the four discriminations were intermixed within the same session

^a Correction trials included

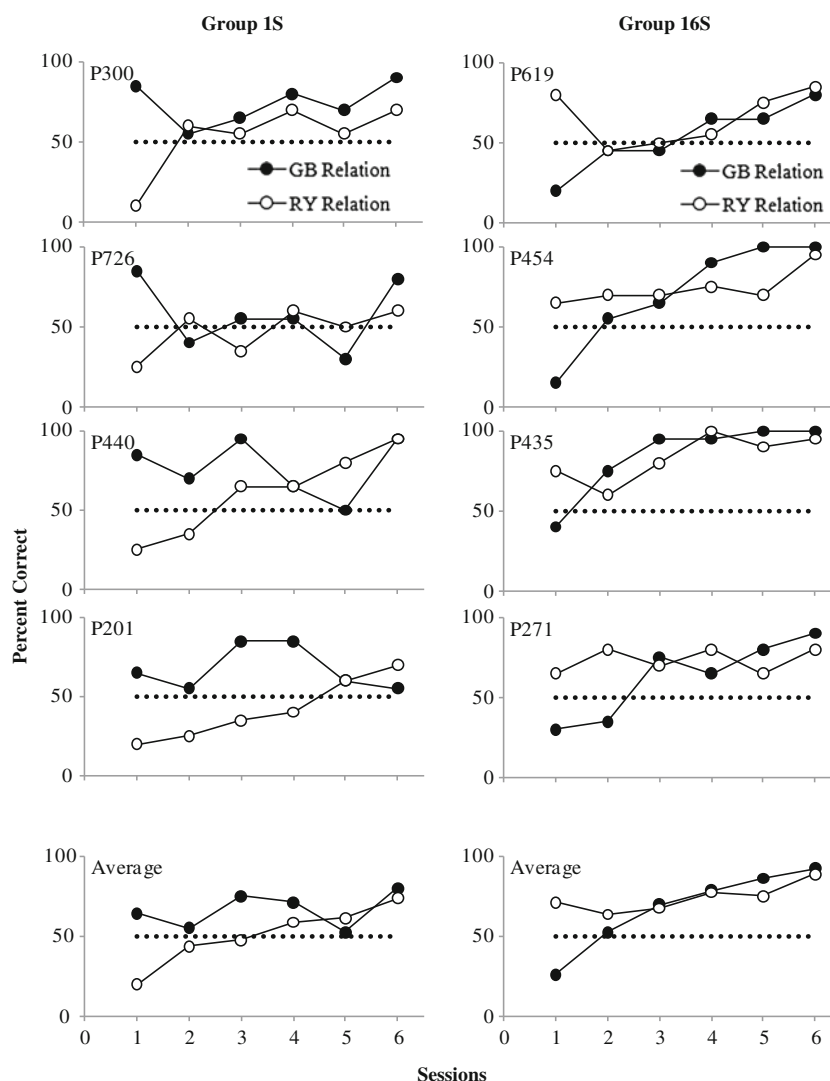
when all four discriminations were combined in the same session. During the last five sessions before Phase III, the overall proportions of correct responses were consistently high across pigeons and across conditional relations (average, .92; range, .87–.95).

Figure 1 shows individual and average data from Phase III. For Group 1S (left panels), the percentage of correct responses during the first session was substantially higher for the GB relation than for the RY relation (average: 64 % vs. 20 %, respectively). In subsequent sessions, the percentages for the RY relation increased, but the percentages for the GB relation did not show a consistent trend across pigeons. By the last session, the percentages of correct responses did not differ appreciably between the two relations. A two-way repeated measures analysis of variance (ANOVA) revealed a significant effect of relation [$F(1, 3) = 28.6$, $p = .013$] in which scores tended to be higher for the GB relation than for the RY relation, a significant effect of session [$F(5, 15) = 6.48$, $p = .002$] in which performance tended to improve across sessions, and a significant Relation \times Session interaction [$F(5, 15) = 5.59$, $p = .004$] in which performance improved mostly for the RY relation.

In contrast, for Group 16S (right panels), the percentage of correct responses during the first session was substantially lower for the GB than for the RY relation (average: 21.2 % vs. 71.3 %, respectively). In subsequent sessions, these percentages increased, particularly for the GB relation. By the last session, the percentages of correct responses did not differ appreciably between the two relations. A two-way repeated measures ANOVA revealed a nonsignificant effect of relation [$F(1, 3) = 4.3$, $p = .13$], in that the overall difference between the two relations was small (68 % vs. 74 %), a significant effect of session [$F(5, 15) = 19.9$, $p < .001$] in which performance tended to improve across sessions, and a significant Relation \times Session interaction [$F(5, 15) = 7.0$, $p = .001$] in which performance improved mostly for the GB relation.

The results of the present experiment are consistent with the results of other studies on temporal discrimination (e.g., Church, 2002; Church & Deluty, 1977; Machado & Keen, 1999; Stubbs, 1968). Consider, for example, the study by Machado and Keen. In a standard temporal bisection task, pigeons learned to choose B after 4-s samples and Y after 16-s samples. When presented with the intermediate sample durations of 5.9, 8.0, and 11.3 s, the pigeons preferred B after the 5.9-s samples and Y after the 11.3-s samples. These results are consistent with temporal generalization; 5.9 s is closer to 4 s than to 16 s (hence, the preference for B), and 11.3 s is closer to 16 s than to 4 s (hence, the preference for Y). In Phase III of the present

Fig. 1 Individual and average matching accuracies in Phase III of Experiment 1. The filled circles represent accuracy for the GB relation, and the open circles represent accuracy for the RY relation



experiment, the pigeons from Group 1S experienced sample durations closer to 4 s than to 16 s; therefore, during the first session they selected the comparison (B) that had been associated with the 4-s sample. Given the time-based preference for B over Y, the GB relation yielded a higher percentage of correct responses than did the RY relation. In contrast, the pigeons from Group 16S experienced sample durations closer to 16 s than to 4 s; therefore, during the first session, they preferred the comparison (Y) that was associated with the 16-s sample. The time-based preference for Y over B yielded a percentage of correct responses that was higher for the RY relation than for the GB relation. Consistent with timing studies, these results suggest that the pigeons' choices in Phase III were controlled mainly by sample duration and not by sample hue. However, temporal control appeared to have been quickly reduced, considering that the percentage of correct responses for the GB relation in Group 1S decreased for all pigeons

during the second session of Phase III and that the percentage of correct responses for the RY relation in Group 16S decreased for two of the four pigeons.

In the study by Machado and Keen (1999), the results for the 8-s test samples showed that although some pigeons were more likely to choose B, others were more likely to choose Y. The mean preference for one of the two comparisons was closer to indifference than when the sample durations were 5.9 and 11.3 s. This result is consistent with the usual finding in temporal discrimination studies that the point of subjective equality (PSE) tends to be at or close to the geometric mean of the two training durations (i.e., $8 = \sqrt{4 \times 16}$; see, e.g., Church, 2002; Church & Deluty, 1977; Gibbon, 1981; Machado, 1997).

From the three facts that (1) the results from Phase III did not reveal severe disruption of choice performance when the sample hue was introduced, (2) the data from the first session of Phase III revealed orderly control by sample duration, and (3) in temporal

discriminations, the PSE tends to be at or close to the geometric mean of the trained durations, we hypothesized that a sample duration at the geometric mean would maximize control by sample hue, thereby exposing the differences between the GB and RY relations from the previous conditional training (Phases I and II). In Experiment 2, we tested this hypothesis.

Experiment 2

The results of Experiment 1 showed that sample durations strongly controlled choice during Phase III, perhaps masking any control by sample hue. To determine whether sample hue differentially controls choice (i.e., a necessary condition to evaluate emergent relations), we needed a sample duration that would not bias choice. During Phase III of Experiment 2, a conjunctive FT 8-s FR1 schedule was used, such that each sample stimulus (G or R) was on for at least 8 s and until the pigeon pecked the center key at least once. When we minimized the bias caused by sample duration, any difference between the GB and RY relations in choice percentages during the first test session, or in the speed of learning during the test sessions, would suggest that the sample hues influenced choice of the comparison stimuli. The relations hypothetically learned during Phases I and II could then be expressed behaviorally.

Method

Subjects and apparatus

Four experimentally naive adult pigeons (*Columba livia*) were used in Experiment 2. They were housed and maintained similarly to those from Experiment 1, and the apparatus was the same as in Experiment 1.

Procedure

All of the procedural details remained the same as in Experiment 1, with the exception that a conjunctive FT 8-s FR1 schedule was in effect during the sample in Phase III. Although the sample–comparison mappings varied across pigeons (Table 2), we will continue to present the results as if all of the pigeons had the mappings “1 s→R, 4 s→G” and “4 s→B, 16 s→Y.”

Results and discussion

Table 3 shows, on average, that the pigeons required 26.5 sessions (range, 12–47) to learn the two Phase I discriminations, 25.3 sessions (range, 21–30) to learn the two Phase

II discriminations, and 15.8 sessions (range, 6–25) to reach the accuracy criterion when all four discriminations were included in the same session. During the last five sessions that preceded Phase III, the average proportions of correct choices for the four discriminations ranged from .88 to .98. The overall proportion of correct choices was above .94 for all of the pigeons. These results are consistent with those from Experiment 1.

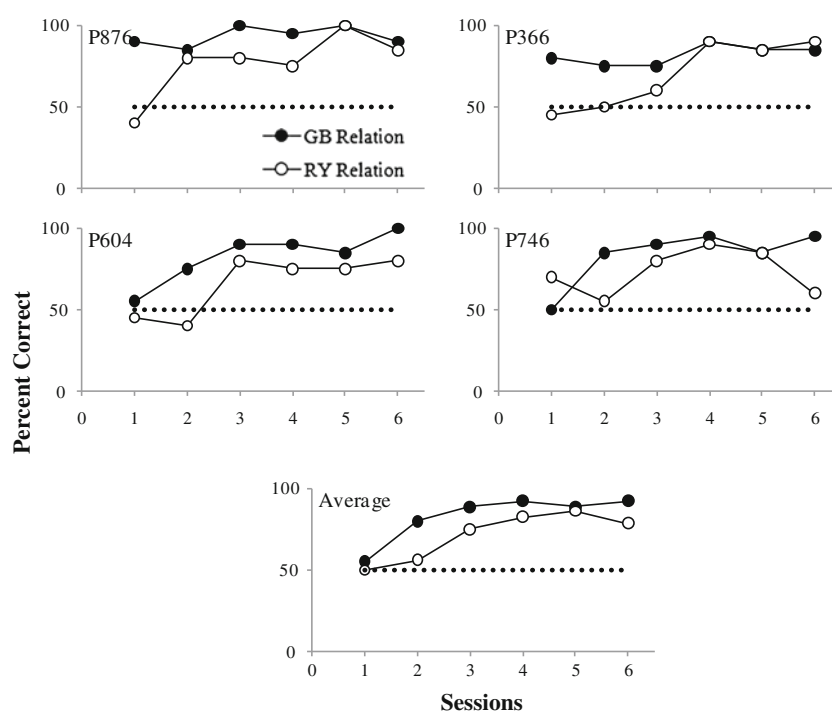
During Phase III, the average sample duration was always close to the minimum scheduled value of 8 s. To illustrate, during the first test session, the median was 8.07 s for the four pigeons, and the percentages of trials on which the sample duration was between 8 and 8.5 s were 95 %, 92.5 %, 97.5 %, and 90 % for subjects P876, P366, P604, and P746, respectively. Because the pigeons pecked during the sample, the obtained sample duration was close to the minimum scheduled value.

Figure 2 shows the individual and average choice data during Phase III. In the first session, P366 and P876 performed well above chance on the GB relation (80 % and 90 %, respectively), but near chance on the RY relation (45 % and 40 %, respectively). P604 showed chance performance on both the GB and RY relations, and P746 showed slightly better performance on the RY relation than on the GB relation.

Considering the results from all pigeons and sessions (i.e., four pigeons × six sessions per pigeon = 24 data points), the percentages of correct responses on the GB relation were higher than the percentages of correct responses on the RY relation in 18 cases, were equal to them in four cases, and were lower in two cases ($p < .01$, binomial test). The visual impression of overall faster learning of the GB than of the RY relation was corroborated by a two-way repeated measures ANOVA with Relation and Session as factors. The ANOVA yielded a strong effect of relation [$F(1, 3) = 64.0, p = .004$] in which overall performance was better on the GB relation than on the RY relation, a strong effect of session [$F(5, 15) = 11.2, p < .001$] in which performance tended to improve across sessions, but a nonsignificant Relation × Session interaction [$F(5, 15) = 0.7$].

To compare the results from Experiments 1 and 2, Fig. 3 shows the mean proportions of choices of B, given a choice between B and Y, following the two samples G (GB relation) and R (RY relation) during the first session of Phase III. For convenience when interpreting the results, we refer to the pigeons from Experiment 2 as *Group 8S*. For these pigeons, the probability of choosing B was higher following the G sample than following the R sample (.69 vs. .50). A one-tailed t test for dependent samples yielded a significant result [$t(3) = 2.85, p < .05$]. In contrast, no such difference existed for Groups 1S (.80 vs. .80) and 16S

Fig. 2 Individual and average matching accuracies in Phase III of Experiment 2. The filled circles represent accuracy for the GB relation, and the open circles represent accuracy for the RY relation



(.26 vs. .29). Moreover, the probability of correctly choosing B for Group 8S (.69) was between the probabilities for Groups 1S (.80) and 16S (.26).

This pattern of results suggests that the pigeons' choices in Experiment 2 were not attributable to a time-based bias for one of the comparisons, as were the choices for the two groups from Experiment 1. Rather, the pigeons' choices were differentially controlled by the sample hue, and consequently, the GB relation yielded more correct responses than did the RY relation.

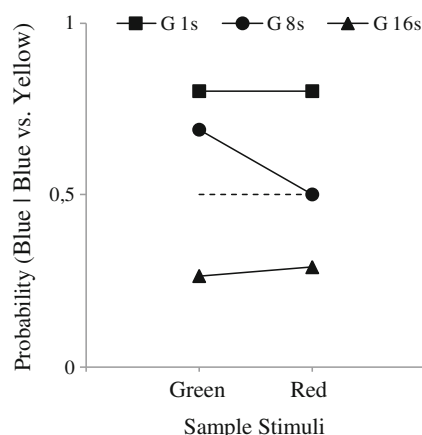


Fig. 3 Mean proportions of choices of blue in Phase III, given a choice between blue and yellow, as a function of the sample stimulus. The filled squares and triangles correspond to the pigeons from Groups 1S and 16S of Experiment 1, respectively, and the circles correspond to the pigeons from Experiment 2 (Group 8S)

General discussion

Experiment 1 showed that choice is not simply disrupted when the sample stimulus changes in the critical test phase, and in addition, that very short or long sample durations bias choice in that phase. Relatively short and long sample durations may mask control by sample hue. For these reasons, in Experiment 2 we attempted to reduce the biasing effect of sample duration, thereby maximizing the control by sample hue. According to timing studies with animals (e.g., Church & Deluty, 1977; Machado & Keen, 1999), the PSE between two durations is typically at their geometric mean. Therefore, in Experiment 2, the sample duration was set at approximately 8 s (i.e., the geometric mean of 4 and 16 s). The results showed consistent *acquisition functions* across all four pigeons, such that the GB relation was learned faster than the RY relation.

However, a few studies have found a PSE that was slightly greater than the geometric mean (e.g., Zentall, Weaver, & Clement, 2004), predicting a bias for B in the present case (i.e., the comparison associated with the 4-s sample). In fact, if the PSE were greater than 8 s (e.g., PSE = 10 s), at least for some pigeons, we would predict that these pigeons would prefer B over Y following test samples shorter than 10 s, and Y over B following test samples longer than 10 s. Because the samples were approximately 8 s long in Phase III of Experiment 2, the pigeons should have shown a bias for B similar to the one shown by subjects P876 and P366

during the first test session (see Fig. 2). Could variations in the PSE explain our findings?

To answer this question, we need to compare the results obtained in the first session of Experiment 2 with those obtained in Experiment 1, particularly from Group 1S. Both P876 and P366 from Experiment 2 and the four pigeons from Group 1S showed a bias for B. Critically, however, regarding P876 and P366, the bias for B occurred only following the G samples; following the R samples, the choice for B was close to chance (see Fig. 2), whereas in the four pigeons from Group 1S, the bias for B occurred following both G and R samples (see Fig. 1). More generally, whereas the results from Experiment 1 are consistent with a time-based account, in which the horizontal segments in Fig. 3 reveal no effect of sample hue, the results from Experiment 2 are not consistent with a time-based account, in which the sloped segment in Fig. 3 shows a moderate effect of sample hue. If we consider the evidence provided by the acquisition functions of Experiment 2 (i.e., GB learned faster than RY), we can conclude that the data from Experiment 2 are not attributable to variations in the PSE.

The use of temporal samples to produce emergent comparison-comparison relations seems to require two seemingly incompatible conditions. First, for the sample to function as a node, and thereby support the emergence of new relations, the sample duration must exert strong control over choice behavior. However, for the comparison-comparison emergent relations to be expressed behaviorally, the sample duration must not exert strong temporal control over choice behavior. However, the seeming incompatibility vanishes when we realize that the two conditions apply at different moments—the sample must exert strong control during training, but little or no control during testing.

Our findings suggest that in MTS tasks temporal stimuli may be an interesting alternative to the usual sample stimuli that is worth exploring. Other reasons for using temporal stimuli stem from the work of Miller and Barnet (1993), who have claimed that when two or more stimuli are paired, the temporal properties of the pairing (e.g., the various intervals between the stimuli) are always encoded. Time, in other words, seems to play a prominent role in all forms of associative learning. If this is the case, temporal stimuli may generate a powerful source of control in conditional discrimination training and should be used more often in experiments studying symbolic behavior.

The present results appear to be consistent with the common-coding effect reported in experiments that have shown emergent relations in pigeons (e.g., Edwards, Jagielo, Zentall, & Hogan, 1982; Urcuioli, Zentall, Jackson-Smith, & Steirn, 1989; Zentall, 1996; Zentall, Clement, & Weaver, 2003; Zentall, Steirn, Sherburne, & Urcuioli, 1991; for a detailed discussion, see Zentall, 1998). Urcuioli, Zentall, and DeMarse (1995) evaluated derived sample-comparison relations in pigeons following

many-to-one versus one-to-many matching training. In their Experiment 1, illustrated in the top panel of Fig. 4, pigeons received AB and CB training during Phase I and AD training during Phase II. When they were tested for the emergent CD relation during Phase III, they showed positive results. The authors proposed the following explanation: First, consistent with the secondary-generalization hypothesis of Hull (1939), when a subject learns to select a comparison stimulus (e.g., B1) conditionally upon a sample stimulus (e.g., A1), the subject then produces this comparison stimulus implicitly or covertly when exposed to the same sample stimulus. Second, during Phase I, the A and C stimuli acquire the function to produce covertly the B stimuli, the common code. Third, in Phase II, a mediating link between the B stimuli, covertly produced upon presentation of A, and the D stimuli is formed. Fourth, during Phase III, the C stimuli covertly produce the B stimuli (due to Phase I training). Because the B stimuli were connected to the D stimuli (due to Phase II training), the pigeons express the emergent CD relation. The authors concluded that the CD relation supports the common-coding hypothesis.

A similar behavioral process may explain the results from Experiment 2. As the bottom panel of Fig. 4 suggests, the pigeons may have learned to produce G covertly in Phase I in the presence of the 4-s sample. Then, in Phase II, a mediating link between G and B could have been formed during the 4-s samples, in which G would be covertly produced by the 4-s samples, and choices of B would have been reinforced in the presence of G. The GB link could then explain the better results on the GB than on the RY relation.

Formulations similar to the common-coding effect have been advanced by Hall's (1994, 1996) model of associatively activated stimulus representations. According to the model, when two representations are activated concurrently, a

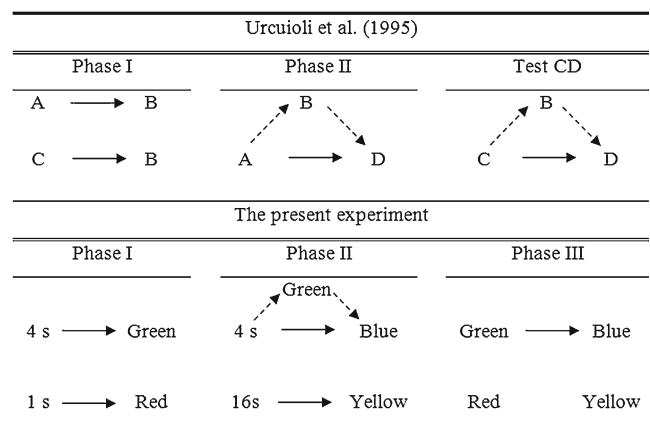


Fig. 4 Summary of the procedures used in Urcuioli et al. (1995) and the present experiment. Explicitly trained or tested relations are indicated by solid arrows, and covertly generated comparison stimuli are indicated by dashed arrows

link between them is formed. Thereafter, a representation can be activated directly or indirectly—directly by presenting the corresponding stimulus, and indirectly by activating the other representation. To apply the model to our experiment, consider that during Phase I a link was established between the representations of the 4-s sample stimulus and the G comparison stimulus; we call this *Link 1*. Next, during Phase II, a second link (*Link 2*) was established between the representations of the 4-s sample stimulus and the B comparison stimulus. At that moment, the 4-s stimulus activated two representations, the representation of B via Link 2 and the representation of G via Link 1. According to Hall's model, the concurrent activation of the G and B representations would link them (*Link 3*). In Phase III, when G was presented as the sample, the representation of B was activated via Link 3, and the animal preferred B (see also Hall, Mitchell, Graham, & Lavis, 2003).

After the one-to-many AB and AC training, positive results on the combined BC and CB tests demonstrated the emergence of transitivity, symmetry, and reflexivity. Altogether, these three emergent performances provide behavioral evidence of equivalence-class formation (Sidman, 1994; Sidman et al., 1982; Sidman, & Tailby, 1982). For this reason, the combined tests are also called *equivalence tests*.

In the present experiment, the tests focused on only one of the two possible emergent relations. That is, only the relations that involved the G and R stimuli as samples and the B and Y stimuli as comparisons were tested. Without testing the converse relations, with B and Y as samples and G and R as comparisons, we do not have evidence of all of the necessary emergent performances to conclude that the pigeons formed equivalence classes. Using the same pigeons to test all of the relations would have required a significantly greater number of sessions, but then the effects of reinforcement learning during Phase III would probably mask the transfer effects of Phases I and II. The solution might be a mixed between-within design, in which one group of pigeons would be tested for one relation (as in Exp. 2) and another group would be tested for the other relation (B and Y as samples and R and G as comparisons).

Most experiments using matching-to-sample tasks with nonhuman animals, including pigeons, have shown that stimulus location also controls performance (cf. Iversen, 1997; Iversen, Sidman, & Carrigan, 1986; Lionello & Urcuioli, 1998; Lionello-DeNolf & Urcuioli, 2000, 2002; Sidman, 1992). To illustrate, when subjects are trained using fixed positions (e.g., samples on the center key and comparisons on the two side keys) and then tested using variable positions (e.g., samples presented on one of the side keys and the comparisons presented on the remaining two keys), accuracy is significantly lower than when samples and comparisons maintain fixed positions (e.g., Iversen, 1997; Iversen et al., 1986; Lionello & Urcuioli, 1998). According to Lionello and

Urcuioli, control by stimulus location may explain the negative results obtained in tests for emergent relations with the three-key paradigm because, in some cases at least, the locations of samples and comparisons changed during the tests.

In the present study, stimuli R and G were presented on the side keys during Phase I and on the center key during Phase III. Although this change in location is the opposite of that made by Lionello and Urcuioli (1998)—in their study, the stimuli previously presented on the center key were later presented on the side keys—one would expect a similar disruption of the emergent performance evaluated in Phase III. However, at least for pigeons P876 and P366, no such disruption took place. Furthermore, considering the general acquisition curves of Phase III, any potential control by stimulus location was not sufficiently strong to reduce the higher accuracy levels achieved in the GB relation when compared to the RY relation.

According to Saunders and collaborators (cf. Saunders & Green, 1999; Saunders, Drake, & Spradlin, 1999; Saunders, Saunders, Williams, & Spradlin, 1993), conditional discriminations embed two types of simple discriminations: successive discriminations between samples presented across trials, and simultaneous discriminations between comparisons presented on each trial. According to this viewpoint, negative results in transitivity and symmetry tests following one-to-many training could occur if the training did not include some of the necessary simple discriminations. The argument is as follows: During AB training, the subject learns a successive discrimination between samples A1 and A2 and a simultaneous discrimination between comparisons B1 and B2. During AC training, the subject learns in addition to the successive discrimination between the A samples a simultaneous discrimination between comparisons C1 and C2. Missing from the training, however, is the successive discrimination between the B1 and B2 samples that is necessary for the BC relation to emerge. In other words, for a novel relation between the comparisons to emerge, all relevant conditional discriminations, with their embedded simple discriminations, must be trained. Because the one-to-many structure does not train all of the embedded discriminations, negative results could be obtained in the tests for the emergent relations between the comparisons.

In the present study, stimulus set B (the R and G hues) and stimulus set C (the B and Y hues) were presented simultaneously as comparisons during Phases I and II, but the successive discrimination between stimulus set B was not in place before Phase III. However, in contrast with Saunders and Green's (1999) hypothesis, P876 and P366 achieved high accuracy levels in the GB relation. It seems that, for these subjects, the simple simultaneous discriminations trained in Phase I produced the repertoire of simple discrimination necessary for appropriate performance during Phase III.

Another question addressed by our experiment concerns the testing of emergent relations using reinforced trials and a within-subjects design (Velasco et al., 2010). In the standard conditional discrimination training procedure, correct responses are reinforced but incorrect responses have no scheduled consequences, except in some cases the repetition of the trial (i.e., correction method). However, these differential response outcomes do not occur during testing (e.g., Kuno et al., 1994; Lipkens et al., 1988) because an extinction procedure is usually in effect during the test trials (e.g., Sidman et al., 1982). The use of extinction during the test trials may produce various effects, including changes in the source of stimulus control, resurgence of behavior, and an increase in behavioral variability (cf. Antonitis, 1951; Dube & McIlvane, 1996; Epstein, 1983, 1985; Galvão et al., 1992; Kuno et al., 1994; Lerman & Iwata, 1996; Schusterman & Kastak, 1993; Sidman et al., 1982; Wilson & Hayes, 1996). The testing protocol used in the present experiment overcomes this problem because it assesses emergent conditional relations using reinforced test trials. Specifically, it compares the acquisition of two conditional relations, one comprising stimuli associated by prior training, and another comprising stimuli not associated by prior training. In our experiments, an emergent conditional relation was evaluated in Phase III, considering that G and B had a common association with 4-s samples, whereas R and Y had no such association.

The use of reinforced tests to assess the emergence of conditional relations has been proposed in previous studies (e.g., Edwards et al., 1982; Schusterman & Kastak, 1993; Urcuioli et al., 1989; Zentall et al., 1991). However, in those studies, the acquisition of different conditional discriminations had to be compared across groups, which not only requires a large number of subjects but also may obscure the effects of training in individual subjects (see Velasco et al., 2010). The within-subjects design used in the present study also overcomes these problems.

In summary, our results clearly indicated the emergence of a new relation between two visual comparison stimuli, a relation that could be established only through their common relation with a temporal stimulus. As far as we know, this is the first demonstration of cross-modal emergent relations involving temporal and visual stimuli in pigeons. Although it was not possible to produce strong evidence of equivalence-class formation, the present findings suggest that more than one emergent performance was simultaneously established. This is a noteworthy fact, considering that only one experiment with mammals (Schusterman & Kastak, 1993), and no experiment with birds, has reported the simultaneous establishment of emergent performances. Experiments using similar stimuli or similar training and testing protocols may help to clarify the controversial question about equivalence-class formation in nonverbal subjects.

Author note This research was supported by doctoral grants to E.M.H. (FAPESP 60678-4/05 and CAPES 0103/08 0) and S.M.V. (CNPq 142544/2005-1 and CAPES 4457 07 2). G.Y.T. and D.G.d.S. were supported by a research productivity fellowship from The National Council for Scientific and Technological Development (CNPq, Brazil). A.D.M. was funded by the Portuguese Foundation for Science and Technology (FCT). All of the Brazilian authors are currently affiliated with the National Institute of Science and Technology on Behavior, Cognition and Teaching, supported by FAPESP (Grant No. 08/57705-8) and CNPq (Grant No. 573972/2008-7). The data were collected in the Animal Learning and Behavior Laboratory, University of Minho, Portugal, and were presented at the 36th Annual Meeting of the Brazilian Society of Psychology (SBP), Goiânia, Brazil, in October 2009. E.M.H. is currently a college professor at the Universidade Federal de Minas Gerais.

References

- Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology*, 42, 273–281.
- Catania, A. C. (1991). Time as a variable in behavior analysis. In I. H. Iversen & K. A. Lattal (Eds.), *Experimental analysis of behavior—Part 2* (pp. 1–18). Amsterdam, The Netherlands: Elsevier.
- Church, R. (2002). Temporal learning. In R. Gallistel & H. Pashler (Eds.), *Stevens' handbook of experimental psychology* (Learning, motivation, and emotion, Vol. 3, pp. 365–393). New York, NY: Wiley.
- Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology. Animal Behavior Processes*, 3, 216–228. doi:10.1037/0097-7403.3.3.216
- D'Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1985). Symmetry and transitivity of conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). *Journal of the Experimental Analysis of Behavior*, 44, 35–47.
- Devany, M. J., Hayes, S. C., & Nelson, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior*, 46, 243–257.
- Dube, W. V., & McIlvane, W. J. (1996). Some implications of a stimulus control topography analysis for emergent stimulus class. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 197–218). Amsterdam, The Netherlands: Elsevier.
- Dube, W. V., McIlvane, W. J., Callahan, T. D., & Stoddard, L. T. (1993). The search for stimulus equivalence in nonverbal organisms. *Psychological Record*, 43, 761–778.
- Edwards, C. A., Jagielo, J. A., Zentall, T. R., & Hogan, D. E. (1982). Acquired equivalence and distinctiveness in matching to sample by pigeons: Mediation by reinforcer-specific expectancies. *Journal of Experimental Psychology. Animal Behavior Processes*, 8, 244–259.
- Epstein, R. (1983). Resurgence of previously reinforced behavior during extinction. *Behaviour Analysis Letters*, 3, 391–397.
- Epstein, R. (1985). Extinction-induced resurgence: Preliminary investigations and possible applications. *Psychological Record*, 35, 143–153.
- Frank, A. J., & Wasserman, E. A. (2005). Associative symmetry in the pigeon after successive matching-to-sample training. *Journal of the Experimental Analysis of Behavior*, 84, 147–165.
- Galvão, O. F., Calcagno, S., & Sidman, M. (1992). Testing for emergent performances in extinction. *Experimental Analysis of Human Behavior Bulletin*, 10, 18–20.
- Garcia, A., & Benjumea, S. (2006). The emergence of symmetry in a conditional discrimination task using different responses as proprioceptive samples in pigeons. *Journal of the Experimental Analysis of Behavior*, 86, 65–80.

- Gibbon, J. (1981). On the form and location of the Psychometric Bisection Function for time. *Journal of Mathematical Psychology*, 24, 58–87. doi:10.1016/0022-2496(81)90035-3
- Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229–262). New York, NY: Plenum Press.
- Hall, G. (1994). Pavlovian conditioning: Laws of association. In N. J. Mackintosh (Ed.), *Handbook of perception and cognition: Animal learning and cognition* (pp. 15–43). Orlando, FL: Academic Press.
- Hall, G. (1996). Learning about associatively activated stimulus representations: Implications for acquired equivalence and perceptual learning. *Animal Learning & Behavior*, 24, 233–255. doi:10.3758/BF03198973
- Hall, G., Mitchell, C., Graham, S., & Lavis, Y. (2003). Acquired equivalence and distinctiveness in human discrimination learning: Evidence for associative mediation. *Journal of Experimental Psychology: General*, 132, 266–276. doi:10.1037/0096-3445.132.2.266
- Hayes, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, 51, 385–392.
- Hogan, D. E., & Zentall, T. R. (1977). Backward associations in the pigeon. *The American Journal of Psychology*, 90, 3–15.
- Honig, W. K., & Urcuioli, P. J. (1981). The legacy of Guttman and Kalish (1956): 25 years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior*, 36, 405–445.
- 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.
- Hull, C. L. (1939). The problem of stimulus equivalence in behavior theory. *Psychological Review*, 46, 9–30. doi:10.1037/h0054032
- Iversen, I. H. (1997). Matching-to-sample performance in rats: A case of mistaken identity? *Journal of the Experimental Analysis of Behavior*, 68, 27–45.
- Iversen, I. H., Sidman, M., & Carrigan, P. (1986). Stimulus definition in conditional discriminations. *Journal of the Experimental Analysis of Behavior*, 45, 297–304.
- Kuno, H., Kitadate, T., & Iwamoto, T. (1994). Formation of transitivity in conditional matching to sample by pigeons. *Journal of the Experimental Analysis of Behavior*, 62, 399–408.
- Lazar, R., Davis-Lang, D., & Sanchez, L. (1984). The formation of visual stimulus equivalence in children. *Journal of the Experimental Analysis of Behavior*, 41, 251–266.
- Lerman, D. C., & Iwata, B. A. (1996). Developing a technology for the use of operant extinction in clinical settings: An examination of basic and applied research. *Journal of Applied Behavior Analysis*, 29, 345–382.
- Lionello, K. M., & Urcuioli, P. J. (1998). Control by sample location in pigeons' matching to sample. *Journal of the Experimental Analysis of Behavior*, 70, 235–251.
- Lionello-DeNolf, K. M., & Urcuioli, P. J. (2000). Transfer of pigeons' matching to sample to novel sample locations. *Journal of the Experimental Analysis of Behavior*, 73, 141–161.
- Lionello-DeNolf, K. M., & Urcuioli, P. J. (2002). Stimulus control topographies and test of symmetry in pigeons. *Journal of the Experimental Analysis of Behavior*, 78, 467–495.
- Lipkens, R., Kop, P. F., & Matthijs, W. (1988). A test of symmetry and transitivity in the conditional discrimination performances of pigeons. *Journal of the Experimental Analysis of Behavior*, 49, 395–409.
- Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological Review*, 104, 241–265. doi:10.1037/0033-295X.104.2.241
- Machado, A., & Arantes, J. (2006). Further tests of the Scalar Expectancy Theory (SET) and the Learning-to-Time (LeT) model in a temporal bisection task. *Behavioural Processes*, 72, 195–206. doi:10.1016/j.beproc.2006.03.001
- Machado, A., & Keen, R. (1999). Learning to time (LeT) or scalar expectancy theory (SET)? A critical test of two models of timing. *Psychological Science*, 10, 285–290.
- Machado, A., & Pata, P. (2005). Testing the scalar expectancy theory (SET) and the learning to time model (LeT) in a double bisection task. *Animal Learning & Behavior*, 33, 111–122. doi:10.3758/BF03196055
- Miller, R. R., & Barnet, R. C. (1993). The role of time in elementary associations. *Current Directions in Psychological Science*, 2, 106–111.
- Oliveira, L., & Machado, A. (2008). The effect of sample duration and cue on a double temporal discrimination. *Learning and Motivation*, 39, 71–94.
- Richelle, M., & Lejeune, H. (1980). *Time in animal behaviour*. Oxford, U.K.: Pergamon Press.
- Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion and modification in preschool children. *Journal of the Experimental Analysis of Behavior*, 71, 195–214.
- Saunders, R. R., & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior*, 72, 117–137.
- Saunders, K. J., Saunders, R. R., Williams, D. C., & Spradlin, J. E. (1993). An interaction of instruction and training design on stimulus class formation: Extending the analysis of equivalence. *Psychological Record*, 43, 725–744.
- Schusterman, R. J., & Kastak, D. (1993). A California sea lion (*Zalophus californianus*) is capable of forming equivalence relations. *Psychological Record*, 43, 823–839.
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, 14, 5–13.
- Sidman, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 93–114). Hove, U.K.: Erlbaum.
- Sidman, M. (1992). Adventitious control by the location of comparison stimuli in conditional discriminations. *Journal of the Experimental Analysis of Behavior*, 58, 173–182.
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston, MA: Authors' Cooperative.
- Sidman, M., & Cresson, O., Jr. (1973). Reading and crossmodal transfer of stimulus equivalence in severe retardation. *American Journal of Mental Deficiency*, 77, 515–523.
- Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior*, 37, 23–44.
- 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.
- Spradlin, J. E., Cotter, V. W., & 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.
- Spradlin, J. E., & Saunders, R. R. (1986). The development of stimulus classes using match-to-sample procedure: Sample classification versus comparison classification. *Analysis and Intervention in Developmental Disabilities*, 6, 41–58.
- Staddon, J. E. R. (2001). *Adaptive dynamics: The theoretical analysis of behavior*. Cambridge, MA: MIT Press.
- Stubbs, A. (1968). The discrimination of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior*, 11, 223–238.
- Sweeney, M. M., & Urcuioli, P. J. (2010). Reflexivity in pigeons. *Journal of the Experimental Analysis of Behavior*, 94, 267–282.
- Tomonaga, M., Matsuzawa, T., Fujita, K., & Yamamoto, J. (1991). Emergence of symmetry in a visual conditional discrimination by chimpanzees (*Pan troglodytes*). *Psychological Reports*, 68, 51–60.

- Urcuioli, P. J. (2008). Associative symmetry, antisymmetry, and a theory of pigeons' equivalence-class formation. *Journal of the Experimental Analysis of Behavior*, *90*, 257–282.
- Urcuioli, P. J., Zentall, T. R., & DeMarse, T. (1995). Transfer to derived sample–comparison relations by pigeons following many-to-one versus one-to-many matching with identical training relations. *Quarterly Journal of Experimental Psychology*, *48B*, 158–178. doi:[10.1080/14640749508401445](https://doi.org/10.1080/14640749508401445)
- Urcuioli, P. J., Zentall, T. R., Jackson-Smith, P., & Steirn, J. N. (1989). Evidence for common coding in many-to-one matching: Retention, intertrial interference, and transfer. *Journal of Experimental Psychology. Animal Behavior Processes*, *15*, 264–273. doi:[10.1037/0097-7403.15.3.264](https://doi.org/10.1037/0097-7403.15.3.264)
- Vasconcelos, M., & Urcuioli, P. J. (2011). Associative symmetry in a spatial sample-response paradigm. *Behavioural Processes*, *86*, 305–315.
- Velasco, S. M., Huziwara, E. M., Machado, A., & Tomanari, G. Y. (2010). Associative symmetry by pigeons after few-exemplar training. *Journal of the Experimental Analysis of Behavior*, *94*, 283–295. doi:[10.1901/jeab.2010.94-283](https://doi.org/10.1901/jeab.2010.94-283)
- Wilson, K. G., & Hayes, S. C. (1996). Resurgence of derived stimulus relations. *Journal of the Experimental Analysis of Behavior*, *66*, 267–281.
- Yamamoto, J., & Asano, T. (1995). Stimulus equivalence in a chimpanzee (*Pan troglodytes*). *Psychological Record*, *45*, 3–21.
- Zentall, T. R. (1996). An analysis of stimulus class formation in animals. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 15–34). Amsterdam, The Netherlands: Elsevier/North-Holland.
- Zentall, T. R. (1998). Symbolic representation in animals: Emergent stimulus relations in conditional discrimination learning. *Animal Learning & Behavior*, *26*, 363–377. doi:[10.3758/BF03199229](https://doi.org/10.3758/BF03199229)
- Zentall, T. R., Clement, T. S., & Weaver, J. E. (2003). Symmetry training in pigeons can produce functional equivalences. *Psychonomic Bulletin & Review*, *10*, 387–391. doi:[10.3758/BF03196496](https://doi.org/10.3758/BF03196496)
- Zentall, T. R., Steirn, J. N., Sherburne, L. M., & Urcuioli, P. J. (1991). Common coding in pigeons assessed through partial versus total reversals of many-to-one conditional and simple discriminations. *Journal of Experimental Psychology. Animal Behavior Processes*, *17*, 194–201. doi:[10.1037/0097-7403.17.2.194](https://doi.org/10.1037/0097-7403.17.2.194)
- Zentall, T. R., Weaver, J. E., & Clement, T. S. (2004). Pigeons group time intervals according to their relative duration. *Psychonomic Bulletin & Review*, *11*, 113–117. doi:[10.3758/BF03206469](https://doi.org/10.3758/BF03206469)