Behavioural Processes 95 (2013) 18–30

Contents lists available at SciVerse ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc

The context effect as interaction of temporal generalization gradients: Testing the fundamental assumptions of the Learning-to-Time model

Ana Catarina Vieira de Castro[a,*], Armando Machado[a], Gerson Yukio Tomanari[b]

[a] School of Psychology, University of Minho, 4710 Braga, Portugal
[b] Institute of Psychology, University of São Paulo, Av. Prof. Mello Moraes, 1721 São Paulo, SP 05508-030, Brazil

ARTICLE INFO

Article history:
Received 1 October 2012
Received in revised form 14 January 2013
Accepted 30 January 2013

Keywords:
Context effect
Double bisection task
Learning-to-Time model
Temporal discrimination
Temporal generalization
Pigeons

ABSTRACT

To test the Learning-to-Time model, six pigeons learned two temporal bisection tasks. In one task they learned to choose a Red key over a Green key following 2-s samples and the Green key over the Red key following 6-s samples; in another task, they learned to choose a Blue key over a Yellow key following 6-s samples and the Yellow key over the Blue key following 18-s samples. After each task was learned, temporal generalization gradients were obtained with samples ranging from 0.7 s to 51.4 s. Finally, preference for Green over Blue – the keys associated with the common 6-s duration, was determined as a function of sample duration. Two issues were examined, whether the preference for Green over Blue increased with sample duration, a transposition-like effect reported before, and whether the preference for Green over Blue could be predicted from the generalization gradients for Green and Blue previously obtained. Results showed that preference for Green over Blue increased with sample duration and that the general shape of the function could be predicted from the generalization gradients. The Learning-to-Time model accounted well for the major trends in the data.

This article is part of a Special Issue entitled: SQAB 2012.

1. Introduction

To contrast timing models, Machado and Keen (1999) developed the double temporal bisection task. This task is a modified version of the temporal bisection task used extensively to study the time sense of animals (e.g., Church and Deluty, 1977; Fetterman and Killeen, 1991) and humans (e.g., Allan and Gibbon, 1991; Wearden, 1991). In a temporal bisection task, a pigeon, say, sees a sample light illuminated for either 1 s or 4 s and then chooses between two comparison keys, one Red and one Green. The pigeon receives food if it chooses Red following 1-s samples and if it chooses Green following 4-s samples. We represent these contingencies as "1 s → Red; 4 s → Green". After the animal learns the discrimination it is presented with intermediate sample durations during test trials. The prototypical finding is that the proportion of Green choices increases monotonically with the sample duration; indifference between the two comparisons occurs close to the geometric mean of the training durations (Catania, 1970; Church and Deluty, 1977; Crystal, 2002; Fetterman and Killeen, 1991; Platt and Davis, 1983; Stubbs, 1968; for summaries, see Gallistel, 1990; Richelle and Lejeune, 1980).

In the double bisection task (see Fig. 1) the subjects learn not one but two bisection tasks. The pigeon first learns the task described above, "1 s → Red; 4 s → Green", which we call *Type 1*, and then it learns the task "4 s → Blue; 16 s → Yellow", which we call *Type 2*. After the second task is learned, the two tasks are integrated within the same session. Finally, on test trials, the pigeon is exposed to samples of different durations and to a new key combination, Green and Blue, the keys previously reinforced following the common 4-s sample. The critical finding is that preference for Green over Blue increases as the sample duration ranges from 1 s to 16 s. Machado and Keen (1999); see also Arantes, 2008; Arantes and Machado, 2008; Machado and Arantes, 2006; Machado and Pata, 2005; Oliveira and Machado, 2008, 2009; Vieira de Castro and Machado, 2012) referred to this finding as a context effect because it suggests that the sample contexts in which the Green and Blue keys were trained, Green with a shorter sample, and Blue with a longer sample, affected how the preference for each key varied with the sample duration.

The Learning-to-Time model (LeT; Machado, 1997; Machado et al., 2009), a behavioral model developed on the basis of earlier work by Killeen and Fetterman (1988), can account for the context effect. The model consists of a serial organization of behavioral states, a vector of associative links connecting the behavioral states to the operant responses, and the operant responses themselves. The model assumes that at the onset of the sample only the first state is active but, as time elapses, the activation spreads

* Corresponding author. Tel.: +351 253601396.
E-mail addresses: castro.anacatarina@gmail.com (A.C. Vieira de Castro), armandom@psi.uminho.pt (A. Machado), tomanari@usp.br (G.Y. Tomanari).

0376-6357/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.beproc.2013.01.007

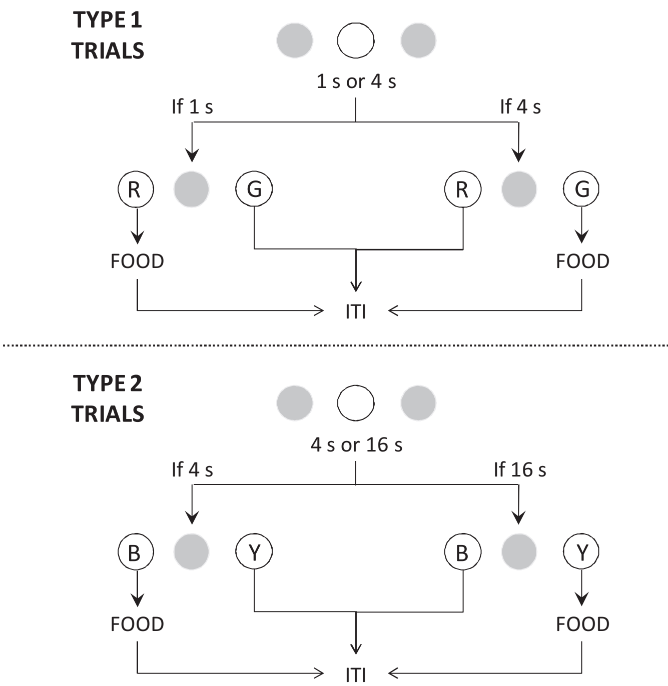

Fig. 1. Structure of the two bisection tasks in a double bisection task. On the *Type 1* task, pecking Red or Green is reinforced after 1-s or 4-s samples, respectively. On the *Type 2* task, pecking Blue or Yellow is reinforced following 4-s or 16-s samples, respectively.

from each state to the next. In addition, each state is linked with the operant responses, and the degree of the linking changes with training, increasing towards 1 with reinforcement and decreasing towards 0 with extinction. The strength of a response at a given moment depends on the most active state at that moment and on the strength of the link between that state and the response.

The model's predictions for the critical test between Green and Blue derive from the strengths of the associative links acquired during training. Table 1 summarizes the details. Let S1, S4, and S16 represent three sets of behavioral states, the sets most likely to be active at the end of 1-s, 4-s and 16-s samples, respectively. Initially, these states are equally linked to the two critical comparisons, Green and Blue; the link strength is represented by the "+" symbol in the table. During the *Type 1* task, choices of Green are reinforced after 4-s samples and extinguished after 1-s samples. Hence, the link of S4 with Green increases to "++" whereas the link of S1 with Green decreases to 0. The link of S16 with Green does not change appreciably from its initial value because S16 is rarely active during the *Type 1* task. During the *Type 2* task, choices of Blue are reinforced after 4-s samples and extinguished after 16-s samples. Consequently the link of S4 with Blue increases to "++" whereas the link of S16 with Blue decreases to 0. The link of S1 with Blue retains its initial value. Given this profile of associative links, consider what happens on test trials: (a) If the sample is 1-s long, S1 will be the most active set of states at the moment of choice, and because S1 is coupled less with Green than with Blue, the animal is unlikely to choose Green; (b) if the sample is 4-s long, S4 will be the most active set of states at the moment of choice, and because it is equally coupled with both keys the animal is indifferent between them; and (c) if the sample is 16-s long, S16 will be the most active set of states at the moment of choice, and because it is coupled more with Green than with Blue, the animal is likely to choose Green. In summary, as Table 1 shows, on the critical test between Green and Blue, LeT predicts that the preference for Green should increase with sample duration.

Table 1
Strength of the links between the behavioral states and the responses during training in a double bisection task, as predicted by LeT. S1, S4 and S16 represent the most likely active sets of behavioral states at the end of the 1-s, 4-s, and 16-s samples, respectively. The arrows show the direction of change. "0", "+", and "++" stand for weak, moderate, and strong links.

Response	Behavioral states		
	S1	S4	S16
Green	+ → 0	+ → ++	+
Blue	+	+ → ++	+ → 0

The context effect is robust. A series of studies using different versions of the double bisection task have produced the context effect. For example, Arantes (2008) used successive rather than simultaneous discrimination tasks; Arantes and Machado (2008) did not combine the *Type 1* and *Type 2* tasks within the same session; Oliveira and Machado (2008) used visually distinct sample stimuli during the *Type 1* and *Type 2* tasks; Oliveira and Machado (2009) presented the choice keys since sample onset; and Machado and Pata (2005) conducted the test trials under nondifferential reinforcement instead of extinction. Despite these variations, the context effect was obtained.

In a recent study, Vieira de Castro and Machado (2012) tested LeT's account of the context effect, namely, that it stems from the profiles of the associative links acquired during training (see Table 1). They reasoned that the link profile for Green could be revealed by the temporal generalization gradient for Green obtained after the animal learned the *Type 1* bisection task and, similarly, the link profile for Blue could be revealed by the temporal generalization gradient for Blue obtained after the animal learned the *Type 2* bisection task. In addition, the interaction of the two generalization gradients should predict how preference for Green over Blue varies with the sample duration (i.e., the context effect).

The authors used a simplified go/no-go version of the double bisection task wherein they trained only the two operants critical to the account. Pigeons learned to choose Green after 4-s samples (S^+) and not to choose it after 1-s samples (S^-). We represent these contingencies as "1 s → Not Green; 4 s → Green". They also learned to choose Blue after 4-s samples (S^+) but not after 16-s samples (S^-), "4 s → Blue; 16 s → Not Blue". After learning each task, a temporal generalization gradient was obtained for each operant, with durations ranging from 1 s to 16 s. Finally, the experimenters gave the pigeons a choice test between the two operants, Green and Blue, while varying the sample duration. Results showed that preference for Green over Blue increased with sample duration, the context effect, and that the interaction of the two separate generalization gradients predicted the preference data.

In the present study we asked whether the results obtained by Vieira de Castro and Machado (2012) using a simplified go/no-go double bisection task would extend to the standard, simultaneous double bisection task. In other words, would the results obtained with the successive discriminations "1 s → Not Green; 4 s → Green" and "4 s → Blue; 16 s → Not Blue" also hold with the simultaneous discriminations "1 s → Red; 4 s → Green" and "4 s → Blue; 16 s → Yellow"? An affirmative answer would mean that the two generalization gradients – one for Green and obtained from choices between Green and Red, and one for Blue and obtained from choices between Blue and Yellow – would predict how the preference for Green over Blue changes with sample duration.

Moreover, in all previous studies with the double bisection task, the sample durations used during the test trials were always within the range of sample durations used during training. For example, when the animals were trained to discriminate between 1- and 4-s samples, and between 4- and 16-s samples, the sample durations used on test trials ranged from 1 s to 16 s. Therefore, the function relating preference for Green over Blue to sample duration is known

only for sample durations within the training range. In the present study we examined the preference function following durations outside the training range.

Pigeons were exposed to a double bisection task with 2-s, 6-s, and 18-s samples, and Red, Green, Blue and Yellow comparison keys. First they learned the task "2 s → Red; 6 s→ Green" and then we varied the sample duration from 0.7 s to 51.4 s to obtain a generalization gradient for Green. Next, they learned the task "6 s → Blue; 18 s → Yellow" and then we varied the sample duration from 0.7 s to 51.4 s to obtain a generalization gradient for Blue. Finally the two tasks were presented within the same session and a preference test between Green and Blue was conducted following sample durations ranging from 0.7 s to 51.4 s. We asked whether the generalization gradients for Green and Blue, obtained separately, predict how the preference for Green over Blue varies with sample duration.

2. Methods

2.1. Subjects

Six adult pigeons (*Columba livia*) participated in the experiment. Four of them (P03, P06, P12 and P18) had experimental histories, but not in temporal discriminations. The other two were experimentally naïve. They were maintained at 85% of their free-feeding body weights throughout the experiment and were housed in individual home cages with water and grit continuously available. A 13:11 h light/dark cycle, beginning at 7:00 am, was in effect in the pigeon colony.

2.2. Apparatus

Two standard experimental chambers for pigeons from Med Associates® were used. The front panel of each chamber contained three keys, 2.6 cm in diameter, centered on the wall 24 cm above the floor and 6 cm apart, center to center. The keys could be illuminated from behind with red, green, blue and yellow lights. Directly below the center key and 6 cm above the floor was a hopper opening measuring 6 cm × 5 cm. When raised the hopper was illuminated with a white light and allowed access to mixed grain. On the opposite wall of the chamber, a 7.5-W houselight provided general illumination. An outer box equipped with a ventilating fan enclosed the experimental chamber. A personal computer controlled all experimental events and recorded the data.

2.3. Procedure

After the two naïve pigeons learned to peck the keys through autoshaping, all pigeons were exposed to the two bisection tasks, one using 2-s and 6-s samples and referred to as *Type 1*, and one using 6-s and 18-s samples and referred to as *Type 2*. Half of the birds (P03, P04, and P06) learned the *Type 1* task first and the *Type 2* task next, whereas the other half learned the tasks in the opposite order. For three pigeons (P03, P04 and P12) the 2-s and 6-s samples of the *Type 1* task were associated with Red and Green, respectively, and the 6-s and 18-s samples of the *Type 2* task were associated with Blue and Yellow, respectively; for the other three pigeons, the reverse assignment was in effect, that is, the 2-s and 6-s samples of the *Type 1* task were associated with Yellow and Blue, respectively, and the 6-s and 18-s samples of the *Type 2* task were associated with Green and Red, respectively. However, for clarity we describe the procedure and the experimental results as if all pigeons had learned first the *Type 1* task with the assignment "2 s → Red; 6 s → Green" and then the *Type 2* task with the assignment "6 s → Blue; 18 s → Yellow".

The general structure of a training trial was as follows. After a 20-s ITI spent in the dark, the houselight was turned on. After the sample duration elapsed, the houselight was turned off and the two side keys were illuminated with different colors. A peck at a side key turned all lights off and if the choice was correct the hopper was activated. Hopper duration varied across birds from 3 s to 12 s in order to maintain body weight with minimal extra session feeding. After food delivery, the ITI followed. If the choice was incorrect the ITI started immediately and the trial was repeated. If the bird made three consecutive errors, only the correct key was presented (correction procedure). As explained below, the trial structure was sometimes modified (e.g., on some trials, choices were not reinforced, even if correct).

The experiment was divided into three phases, 1–3, each subdivided into three conditions, Training, Pretesting, and Testing. Throughout Phase 1, the *Type 1* task "2 s → Red; 6 s → Green" was in effect. During the Training condition, all trials were reinforced provided the choice was correct. Each session comprised 60 trials (excluding repeated trials), 30 for each sample. The left-right position of the red and green keylight colors varied randomly across trials with the constraint that, at the end of the session, each color appeared equally often on the two side keys. When the birds met the criterion of at least 80% correct choices following each sample for three consecutive sessions, they advanced to the following condition. Training continued for 30 sessions on the average (range: 15–45).

During the Pretesting condition extinction trials were introduced to adapt the birds to the lower rate of food that would be in effect during the subsequent condition. In addition to ending without food, even after a correct choice, extinction trials were not repeated after an incorrect choice. Sessions comprised 60 trials, 30 for each sample of which 4 were extinction trials and 26 were regular trials. Pretesting lasted until the learning criterion was met ($M = 4$ sessions; range: 3–7).

During the Testing condition, each session comprised 60 trials, 52 training trials (26 for each sample) and 8 test trials. The sample duration on test trials equalled 0.7, 1.2, 3.5, 10.4, 30.4 or 51.4 s. The test durations were presented in two blocks of sessions. Block 1 included the durations of 0.7, 1.2, 30.4 and 51.4 s, each presented twice per session (once for each left/right assignment of the comparison keylights). Block 2 included the geometric means of the two pairs of training durations, 3.5 s and 10.4 s, each presented four times per session (twice for each left/right assignment of the comparison keylights). If during any test session the percentage of correct choices following one of the training durations fell below 75%, the bird was returned to the Pretesting condition for at least 2 sessions and until its performance again met the criterion. Series of 4 sessions with Block 1 and 2 sessions with Block 2 alternated until a total of 18 test sessions were completed. Because no bird required intercalated Pretesting sessions, Testing lasted 18 sessions for all of them.

Throughout Phase 2, the Type 2 task "6 s → Blue; 18 s → Yellow" was in effect. All procedural details remained as in Phase 1. The average (range) number of sessions during the Training, Pretesting and Testing conditions were, respectively, 30 sessions (8–48), 5 sessions (3–8), and 21 sessions (18–33).

During the Training condition of Phase 3, *Type 1* and *Type 2* tasks alternated across sessions for a minimum of 6 sessions and until the pigeons met the learning criterion on both tasks. Next, both tasks were presented within the same session. At this point each session comprised 80 trials, 40 *Type 1* and 40 *Type 2* (20 for each sample duration). Training lasted until the birds averaged at least 80% correct choices (excluding repeated trials) for three consecutive sessions on each of the four basic samples ($M = 30$ sessions; range: 10–58).

During the Pretesting condition, the session structure remained the same except that some trials were conducted in extinction. Sessions comprised 80 trials, 40 *Type 1* and 40 *Type 2*; of the 20

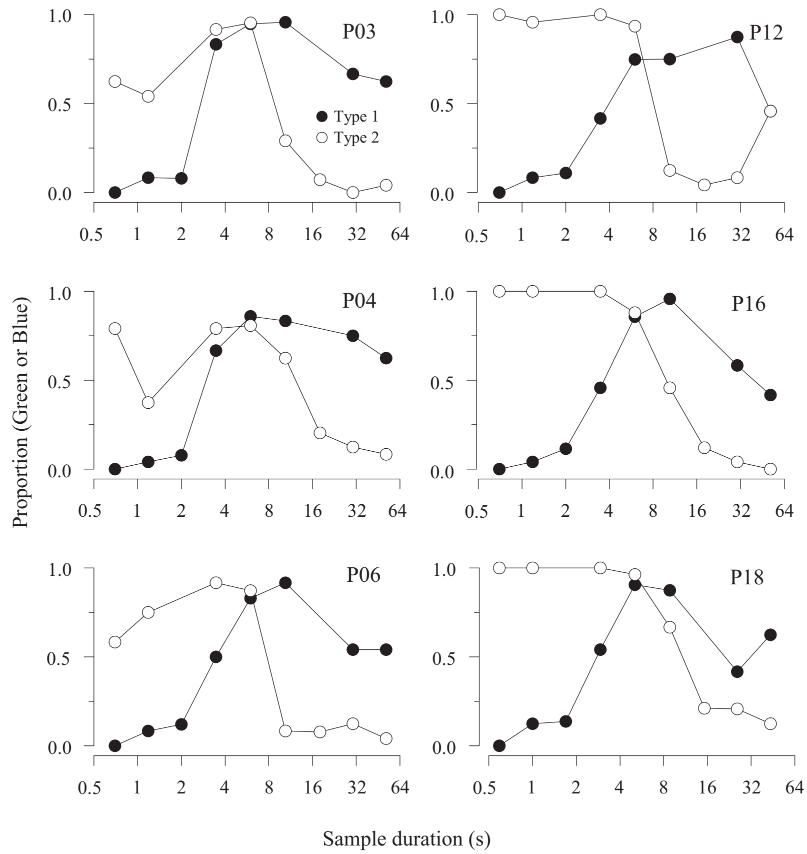

Fig. 2. Generalization gradients obtained following training on the *Type 1* task (2 s vs. 6 s, filled circles) and the *Type 2* task (6 s vs. 18 s, empty circles). Note the logarithmic scale in the *x*-axis.

trials with each sample, 4 were extinction trials. Pretesting lasted 5 sessions on average (range: 3–7).

During the Testing condition, each session comprised 80 trials, 64 training trials (32 *Type 1* and 32 *Type 2*) and 16 test trials. During the test trials, the sample durations remained as in Phases 1 and 2, but a new pair of key colors, namely Green and Blue, was presented for the first time. Each test duration occurred twice per session, one for each left/right assignment. If during any test session the percentage of correct choices following one of the four training samples fell below 75%, the bird returned to Pretesting for at least 2 sessions and until its performance was greater than 80%. The Testing condition continued until 12 test sessions were completed (with intercalated Pretesting sessions, Testing lasted 22 sessions on the average; range 12–45).

3. Results

3.1. Training

All pigeons learned the two basic discriminations. The number of sessions required to learn the *Type 1* task was, on the average, greater than the number of sessions required to learn the *Type 2* task ($M = 35$ and 26, respectively), but the difference was not statistically significant [$t(5) = 1.50$, $p = .19$]. Concerning the order in which the tasks were learned, the first and second tasks required the same ($M = 30$) average number of sessions to learn (range: 15–45, for the first task; and 8–48, for the second task).

During the last three sessions of separate training (i.e., *Type 1* task alone, and *Type 2* task alone), the proportion of correct responses averaged .84 on the *Type 1* task (range: .82–.85) and .87 on the *Type 2* task (range: .81–.92). In Phase 3, the last three training sessions of combined training (*Type 1* task + *Type 2* task) yielded an overall average proportion of correct responses of .90 (range: .80–.96).

The total number of training sessions required to learn the two discriminations (*Type 1* task in Phase 1 + *Type 2* task in Phase 2 + combined tasks in Phase 3) ranged from 35 to 129 ($M = 90$). This number is significantly higher compared to previous studies with simultaneous double bisection tasks (e.g., 29–34 in Machado and Keen, 1999; 27–65 in Machado and Pata, 2005; 26–45 in Machado and Arantes, 2006; 28–68 in Oliveira and Machado, 2008). In fact, if we exclude pigeon P03, which required a total of 35 sessions, a number more consistent with the aforementioned studies, the range was between 72 and 129. The reason for the slower learning may be the fact that the ratio of sample durations used in the present study (1:3) was smaller than the ratio used in most of the previous studies (1:4); according to Weber's law for timing, smaller ratios imply harder discriminations and therefore slower learning. Oliveira and Machado (2009) also used a 1:3 ratio and their pigeons learned the discriminations in fewer sessions (26–38) than in the present study. However, in their study, the choice keys were available during the sample and the authors reported that the pigeons developed patterns of pecking the side keys, which may have facilitated the discriminations.

3.2. Stimulus generalization tests

Fig. 2 shows the temporal generalization gradients obtained following training on the *Type 1* task ($S^- = 2$ s, $S^+ = 6$ s, for pecking Green) and *Type 2* task ($S^+ = 6$ s, $S^- = 18$ s for pecking Blue). The proportion of responses to Green or Blue is plotted against the sample duration. The gradients for the *Type 1* task (filled circles) were

similar for the six pigeons. The proportion of responses to Green was lowest following 0.7-s samples and increased to a maximum value that occurred either following 6-s samples (P04 and P18) or following a slightly longer sample (P03, P06, P12 and P16). For still longer samples the proportion of responses to Green decreased. For the *Type 2* task there was more variability across pigeons, especially for samples shorter than 6 s. For pigeons P12, P16 and P18, the proportion of responses to Blue was high following 6-s and shorter samples, whereas for pigeons P03, P04 and P06 it was high following 3.5-s and 6-s samples but decreased for shorter samples. For samples longer than 6 s the proportion of responses to Blue decreased towards zero for all pigeons except P12. For this subject the proportion of responses to Blue decreased to 0 but then, for the longest sample, it increased again to about .50.

The top panel of Fig. 3 shows the average generalization gradients for the two discriminations. The average gradient for the *Type 1* task reproduced well the individual data. The proportion of responses to Green was lowest for the shortest sample, increased monotonically with sample duration from 0.7 s to 10.4 s and then decreased towards .50 for the longest samples. Concerning the *Type 2* gradient, the proportion of responses to Blue was lowest following the longest samples, increased monotonically as the sample duration decreased from 18 s to 3.5 s, and then decreased slightly following the shortest samples. The larger error bars at the extreme sample durations reflect the individual differences described above.

To compare the two gradients directly, the middle panel of Fig. 3 re-plots the data with the gradient for the *Type 2* task reflected around the vertical line $t = 6$ s. The average gradients have similar shapes. The proportion of responses to Green (or Blue) was low on the S^- side of the gradient, increased from the S^- value to the S^+ value and its neighbour, and then decreased again on the S^+ side of the gradient. Despite the overall similarity, some differences between the two average gradients are noticeable. First, on the S^- side, the gradient for the *Type 1* task reached zero, but the gradient for the *Type 2* task never fell below .10. Second, at the geometric mean of the training durations (i.e., 3.5 s for the *Type 1* task and 10.4 s for the *Type 2* task), the gradient for the *Type 1* task was above the gradient for the *Type 2* task. Third, in the S^+ side of the gradient, the proportion of responses to Blue (*Type 2* task) did not decrease as much as the proportion of responses to Green (*Type 1* task). A two-way, repeated measures ANOVA with sample duration and type of discrimination as factors yielded significant effects of sample duration [$F(7,35) = 96.11$, $p < .001$] and of its interaction with type of discrimination [$F(7,35) = 3.60$, $p < .05$], but no effect of type of discrimination [$F(1,5) = 1.69$, $p = .25$].

One of the most robust properties of temporal discriminations is the scalar property, which in temporal bisection tasks is expressed by the fact that the individual psychometric functions obtained with different training samples overlap when the *x*-axis equals test duration divided by the geometric mean of the two training durations (Church, 2003; Gibbon, 1977, 1991; Lejeune and Wearden, 2006). To determine whether the present results also show the scalar property, the bottom panel of Fig. 3 compares the average proportion of 'Long' responses on the two bisections tasks, that is, the average proportion of responses to Green on the *Type 1* test trials and the average proportion of responses to Yellow on the *Type 2* test trials. All test durations from one task, t, were divided by the geometric mean (GM) of the training durations used in that task. The results show that the two curves superimposed except, perhaps, at the extremes. At the individual level (data not shown) there was some variability. For pigeons P03, P16, and P18 the two gradients superimposed, but for the other three they did not (for pigeons P06 and P12 the gradient for the *Type 2* task was steeper than the gradient for the *Type 1* task, but the opposite was the case for pigeon P04). Although superimposition was not met for all pigeons, there was no tendency for one gradient to be steeper than the other.

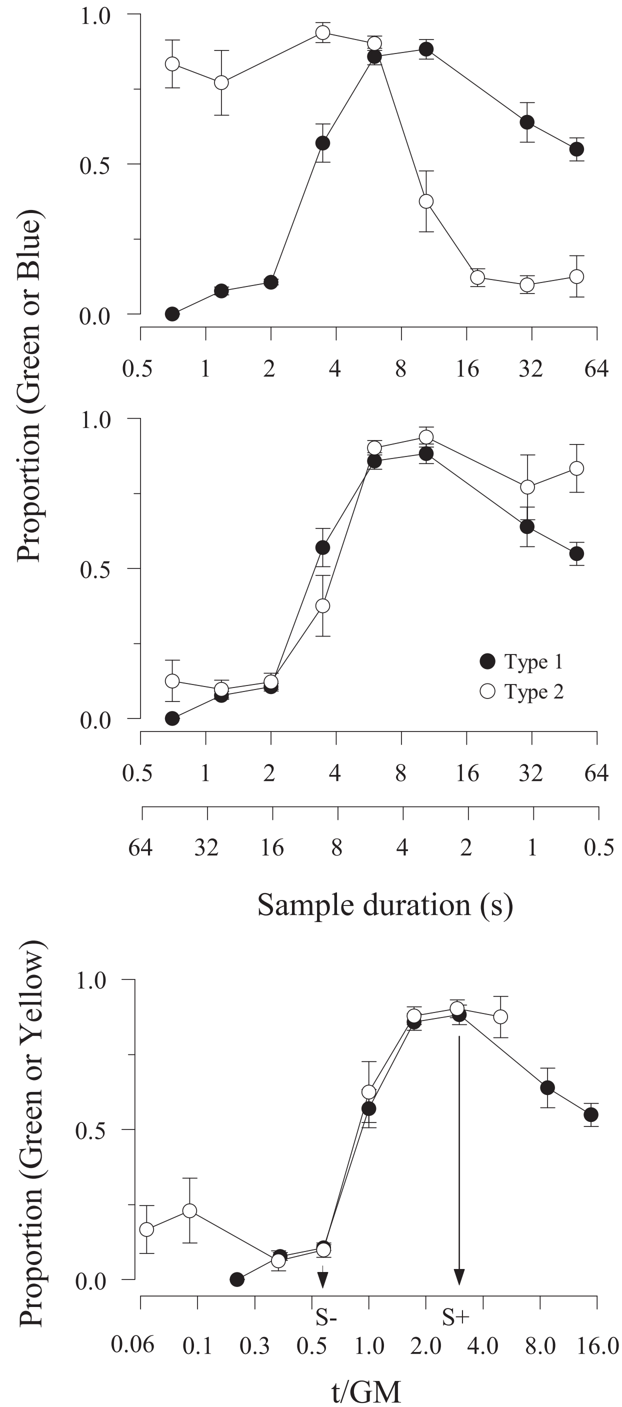

Fig. 3. Top panel. Average generalization gradients obtained following training on the *Type 1* task (2 s vs. 6 s, filled circles) and the *Type 2* task (6 s vs. 18 s, empty circles). Middle panel. The gradient for the *Type 2* task is reflected along the vertical line t = 6. The gradient for the *Type 1* task is plotted according to the upper *x*-axis and the gradient for the *Type 2* task is plotted according to the lower *x*-axis. Note the logarithmic scale on the *x*-axis. Bottom panel. Proportion of Green choices on the *Type 1* task and Yellow choices on the *Type 2* task as a function of relative sample duration. For each discrimination, the sample durations were divided by the geometric mean of the training durations. t/GM is expressed in base 2 log units. The vertical bars show the SEM.

Hence we conclude that the two gradients did not differ systematically. An ANOVA conducted with the data from the five samples common to the two tasks (i.e., in the bottom panel of Fig. 3, the t/GM range from 0.33 to 3), and with sample duration and type of discrimination as repeated factors, confirmed that the gradients superimposed significantly. There was a significant effect of sample duration [$F(4,20) = 227.44$, $p < .001$], but no significant effect of type of discrimination [$F(1,5) = .15$, $p = .72$] and, more important for the present analysis, no significant effect of the interaction between sample duration and type of discrimination [$F(4,20) = .20$, $p = .93$].

To further analyze the superimposition of the generalization gradients we compared the indifference points and the Weber fractions for the two discriminations. The indifference point corresponds to the sample duration that yields a proportion of 'Long' responses equal to .5. The indifference points were estimated by linear interpolation from the individual relative generalization gradients and averaged 0.98 for both the *Type 1* task (3.3 s in the absolute scale) and the *Type 2* task (9.7 s in the absolute scale), $t(5) = .02$, $p = .98$. The Weber fraction corresponds to the ratio between the difference limen and the indifference point. The difference limen equals half the difference between the sample durations that yield the proportions of 'Long' responses equal to .75 and .25, estimated also by linear interpolation from the individual relative generalization gradients. The Weber fractions averaged 0.35 for the *Type 1* task and 0.28 for the *Type 2* task, but the difference was not statistically significant [$t(5) = 1.2$, $p = .28$]. The fact that the indifference points and the Weber fractions did not differ across tasks is consistent with the overlap of the generalization gradients when plotted in a common scale (bottom panel of Fig. 3). Additionally, although the average indifference points were slightly below the geometric means of the training durations, they did not differ significantly from them [*Type 1* task: $t(5) = -.29$, $p = .78$; *Type 2* task: $t(5) = -.21$, $p = .84$]. We conclude that the temporal generalization gradients showed two of the most robust properties of timing data, superimposition and indifference close to the geometric mean.

3.3. Stimulus-response generalization testing

One of the major goals of the present study was to investigate the shape of the preference function in the critical test between Green and Blue when the test range included sample durations outside the training range. Two questions were of critical interest. First, when testing included sample durations significantly outside the training range, was the context effect preserved, that is, did preference for Green over Blue still increase as the sample ranged from 2 s to 18 s? And second, what was the pigeons' preference for sample durations outside the training range, that is, for the shorter samples of 0.7 and 1.2 s and for the longer samples of 30.4 and 51.4 s? Fig. 4 shows the individual and average preference functions.

Consider first the results within the training range (i.e., between the arrows in Fig. 4). For five pigeons (except P03), the preference for Green tended to increase with sample duration. Pigeon P03 displayed the most discrepant performance, for it showed an approximately equal preference for Green following all samples. On the average (see bottom panel), preference for Green increased monotonically from about .35 following 2-s samples to about .60 following 18-s samples. A repeated-measures ANOVA revealed a significant effect of sample duration, $F(3,15) = 7.93$, $p < .005$. We conclude that in the present experiment the context effect, although somewhat weaker than in previous studies, was reproduced.

The results for the full range of test durations revealed that the individual preference functions tended to increase with the sample duration. The preference functions for all pigeons except P03 had their minimum value following one of the shorter samples and their maximum value following one of the longer samples. The average function (see bottom panel) also showed that the preference for Green increased with sample duration. A repeated-measures ANOVA showed that the effect of sample duration was statistically significant, $F(7,35) = 7.05$, $p < .001$.

The fact that pigeon P03 did not show the context effect and that the magnitude of the effect was weaker than in previous studies (see Arantes, 2008; Arantes and Machado, 2008; Machado and Arantes, 2006; Machado and Keen, 1999; Machado and Pata, 2005; Oliveira and Machado, 2008, 2009; Vieira de Castro and Machado, 2012) suggested a finer analysis of the pigeons' performances across test sessions. Because the test trials ended without food it could be the case that pigeons began to learn, in the course of testing, that the Green and Blue keys presented together signaled extinction, which learning in turn could drive preference towards indifference. The more general issue is whether performance changed systematically across test sessions.

Fig. 5 displays the preference functions for the first three (empty circles) and last three (filled circles) test sessions. The top six panels show that, in the first sessions, and within the training range (signaled by the arrows), the preference for Green tended to increase with sample duration – the context effect, albeit to different degrees. In the last sessions, the trend was reversed for pigeons P03 and P18. In general, then, the context effect was stronger during the first test sessions. This trend is clearly visible in the average curves (bottom panel): the steepness of the preference function decreased from the first to the last sessions. Non-parametric Friedman tests showed that the effect of sample duration on the preference for Green was statistically significant in the first sessions [$\chi^2(3) = 13.35$, $p < 0.005$] but not in the last sessions [$\chi^2(3) = 1.39$, $p = 0.71$]. Additionally, the average data shows that the flattening of the preference function across testing was mainly due to an increase in the preference for Green for the shorter durations.

3.4. Predicting preference functions from generalization gradients

The second major goal of the present study was to assess whether the separate generalization gradients for Green and Blue predicted the choice between Green and Blue. To that end we used the individual data from the stimulus generalization tests (Fig. 2) to compute the predicted choice functions. Specifically, for each sample duration, we divided the proportion of responses to Green in the gradient for the *Type 1* task by the sum of the proportions to Green in the gradient for the *Type 1* task and to Blue in the gradient for the *Type 2* task. For example, to compute the predicted preference for 6-s samples for pigeon P03, we divided the proportion for Green represented by the filled circle at 6 s in the top left panel of Fig. 2 by the sum of this proportion and the proportion for Blue represented by the empty circle at 6 s in the same panel. Because the preference functions tended to become flatter across testing (see Fig. 5) we compared the predicted preference functions with the preference functions obtained during the first three test sessions. Fig. 6 shows the individual and average functions plotted together.

The predicted preference functions showed an increasing preference for Green over Blue with sample duration (see top six panels of Fig. 6, empty circles). Specifically, the preference for Green equalled zero following 0.7-s samples, was approximately .50 following 6-s samples and continued to increase for longer samples to a maximum value between .80 and 1. For pigeon P12 the preference for Green reached its maximum following 30.4-s samples, but decreased to .50 following the longest sample. Additionally, all six predicted functions expressed the context effect – within the training range, the preference for Green increased with sample duration.

For all pigeons, the predicted and obtained preference functions were positively correlated. Pigeon P16 showed the strongest

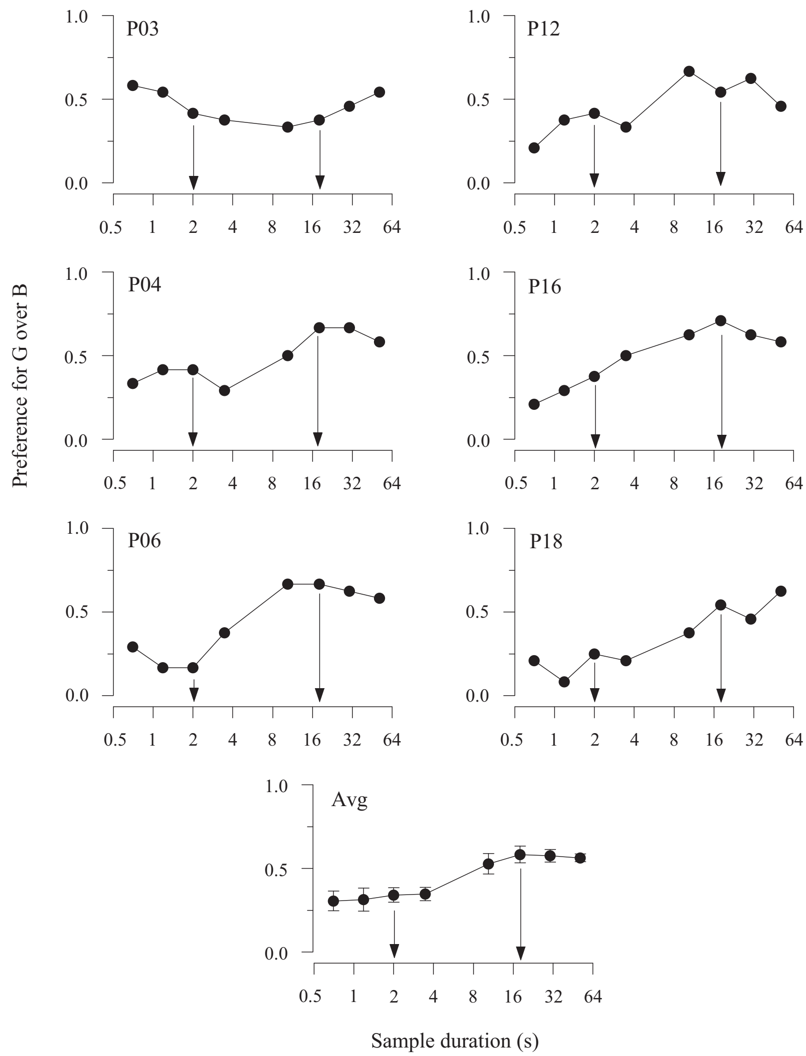

Fig. 4. Preference for Green over Blue as a function of sample duration. The top six panels show the individual data and the bottom panel shows the average data. The arrows delimit the training range. The vertical bars show the SEM. Note the logarithmic scale on the *x*-axis.

correlation (r = .94), followed by pigeons P12 and P04 (r = .82 and .81, respectively). For pigeons P18 and P06 the two functions correlated to a lesser degree (r = .72 and r = .61, respectively). The functions for pigeon P03 were the most dissimilar (r = 0.35). Although the correlation between the predicted and obtained preference functions was strong for three of the six pigeons, the predicted functions tended to be steeper than the obtained functions, a result clearly visible in the average data (see bottom panel). Despite the differences in steepness, the two average functions were strongly correlated (r = .93).

4. Discussion

Six pigeons learned two temporal discrimination tasks. In the first (*Type 1*), they learned to choose Red over Green following 2-s samples and to choose Green over Red following 6-s samples. In the second (*Type 2*), they learned to choose Blue over Yellow following 6-s samples and to choose Yellow over Blue following 18-s samples. After each task was learned, a stimulus generalization test was conducted by varying the sample duration from 0.7 s, a value significantly below the shortest training sample of 2 s to 51.4 s, a value significantly above the longest training sample of 18 s. Thereafter the two discriminations were included within the same session and a stimulus-response generalization test was conducted in which the pigeons chose between Green and Blue following samples ranging from 0.7 s to 51.4 s. The goal of the study was to extend Vieira de Castro and Machado's (2012) approach to a simultaneous double bisection task in which test durations included values outside the training range. More specifically, the present study asked whether the context effect would hold for sample durations within the range of trained durations and what would be the pigeons' preference for sample durations outside that range. In addition, it asked whether the temporal generalization gradients for Green and Blue could predict the preference data.

The results showed that in the test between Green and Blue five of the six pigeons displayed an increasing preference for Green as a function of sample duration (Fig. 4). The context effect, evaluated within the training range, was shown for the same five pigeons. The pigeon that did not display the context effect when the data included all test sessions (P03, see Fig. 4, top left panel) showed a weak effect during the first test sessions (see Fig. 5, top left panel, empty circles). Because the tests were conducted in extinction it is conceivable that the pigeons learned during testing that Green and Blue presented together signaled extinction and started to choose the comparison keys independent of sample duration. In fact, the slope of the preference functions tended to decrease across testing (compare empty and filled circles in Fig. 5).

On the average, the preference for Green increased from about .35 following 2-s samples to about .60 following 18-s samples

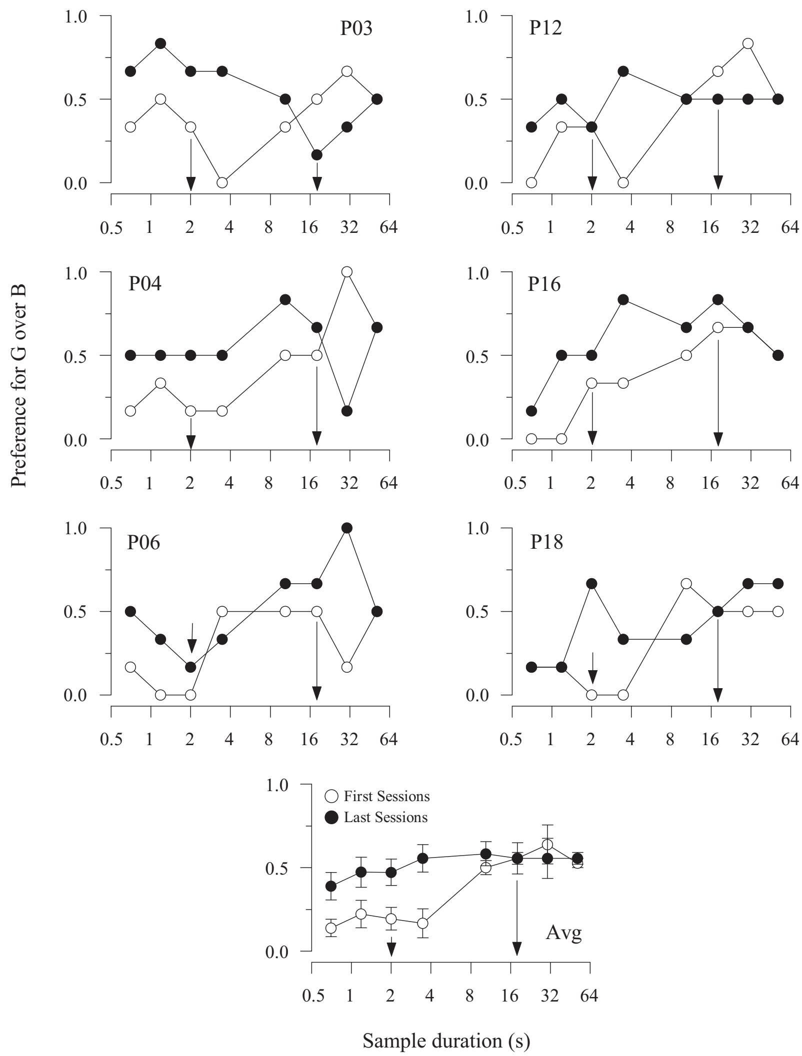

Fig. 5. Preference for Green over Blue as a function of sample duration across the stimulus-response generalization test. The open circles show the data for the first 3 sessions and the filled circles show the data for the last 3 sessions. The top six panels show the individual data and the bottom panel shows the average data. The arrows delimit the training range. The vertical bars show the SEM. Note the logarithmic scale on the *x*-axis.

(bottom panel of Fig. 4). Usually, in studies with the double bisection task, the preference for Green increases from about .10 following the shorter duration to about .80 following the longer duration (see Arantes, 2008; Arantes and Machado, 2008; Machado and Arantes, 2006; Machado and Keen, 1999; Machado and Pata, 2005; Oliveira and Machado, 2008; Vieira de Castro and Machado, 2012). Even if we consider only the data from the first test sessions to exclude the possible effects of extinction, the preference for Green increased from about .20 to about .55 within the training range (empty circles in the bottom panel of Fig. 5). Therefore, the context effect was reproduced in a double bisection task when the test range included durations outside the training range, but it was weaker than in previous studies. One reason for the weaker effect may be related to the difficulty of the discriminations. Here we used a 1:3 training ratio in contrast to the 1:4 training ratio used in the majority of the previous studies. Oliveira and Machado (2009) was the only other study to use a 1:3 training ratio and they also obtained a weaker context effect – the preference for Green increased with sample duration from about .30 to about .60.

The results also showed that the preference functions computed from the temporal generalization gradients predicted well the positive trend of the preference functions for the first test sessions. However, the predicted functions were always steeper than the obtained functions (see Fig. 6). For comparison purposes we analyzed the correlation between the average predicted preference function and the average obtained preference function from the present study and from Vieira de Castro and Machado (2012). The correlation between the predicted and obtained functions was strong in both studies, .93 in the present study, and .96 in Vieira de Castro and Machado's (2012) study. However, in Vieira de Castro and Machado's (2012) study the predicted and obtained preference functions had similar slopes and, therefore, the exact values of the predicted functions were closer to the exact values of the obtained functions than in the present study. In summary, in the present study the temporal generalization gradients were also good predictors of the overall trend of the preference data, but overestimated the effect of sample duration.

The reasons for the differences between the predicted and the obtained preference functions remain uncertain. One possibility is generalization decrement because the stimulus conditions during the stimulus-response generalization test (Green and Blue comparison keys) were different from the stimulus conditions during the tests that yielded the generalization gradients (Red and Green on the *Type 1* task, and Blue and Yellow on the *Type 2* task). Another possibility is that the introduction of test values clearly outside the training range may alter the animals' performance during the test. In fact, in Vieira de Castro and Machado's (2012) study, the test range included only values intermediate to the training

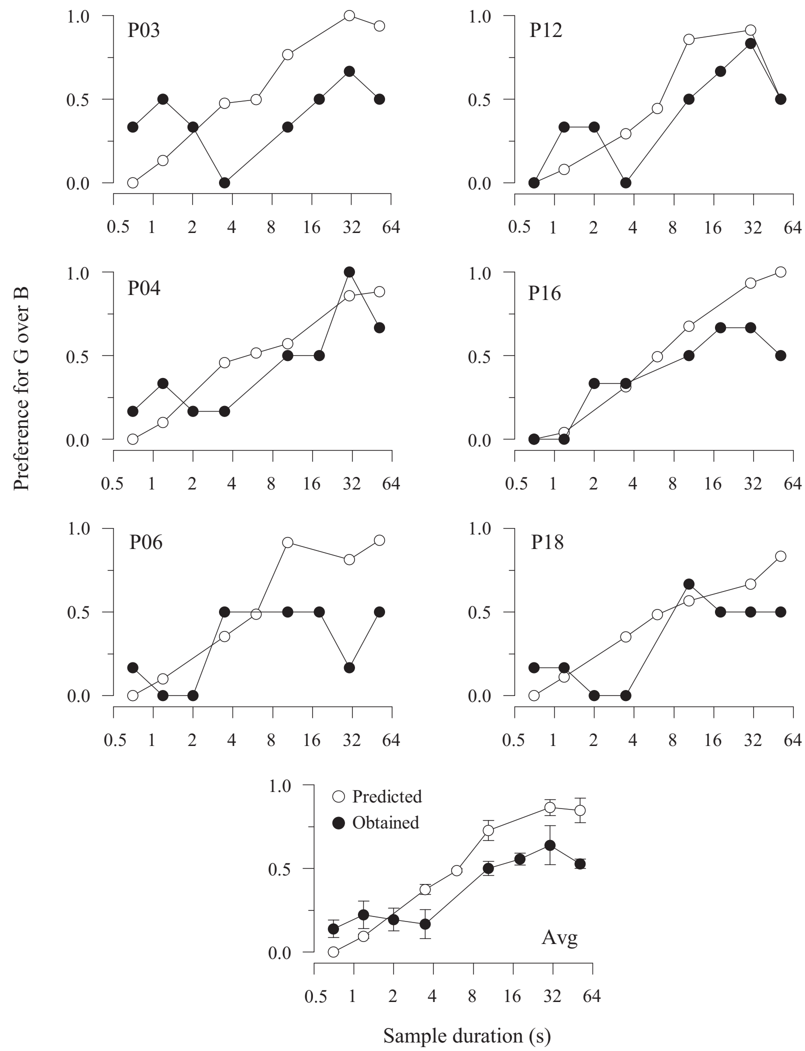

Fig. 6. The filled circles show the preference for Green over Blue obtained in the first 3 sessions of the stimulus-response generalization test. The open circles show the preference for Green over Blue predicted from the stimulus generalization tests. The top six panels show the individual data and the bottom panel shows the average data. The vertical bars show the SEM. Note the logarithmic scale on the *x*-axis.

durations and the predicted preference functions matched substantially better the obtained functions. Also in support of this view, Siegel (1986, Experiment 2) found that the location of the indifference point in a simple bisection task changed with the range of test durations. Specifically, Siegel found that the indifference point was well described by the geometric mean of the training durations when the test range involved only intermediate durations, but it was significantly below the geometric mean of the training durations when values outside the training range were introduced. The reasons for these changes remain unclear.

A parallel purpose of the present study concerned the investigation of the shape of the temporal generalization gradients. Only a few studies have reported temporal generalization gradients outside the training interval following intradimensional discrimination training, and the shape of the obtained gradients was not consistent across them (see Elsmore, 1971; Russell and Kirkpatrick, 2007; Siegel, 1986; Spetch and Cheng, 1998; Mellgren et al., 1983; Vieira de Castro and Machado, 2012). The generalization gradients obtained in the present study fit into two major categories of temporal generalization gradients found in the literature. In the first category, the proportion of responses (to Green or Blue) is low on the S^- side of the gradient, increases to the highest value at the S^+ or at a duration further removed from the S^-, and decreases again for durations far removed from the S^+. All six gradients for the *Type 1* task and half of the gradients for the *Type 2* task (pigeons P03, P04 and P06) belong to this category. In the second category, the proportion of responses is low on the S^- side of the gradient and then increases and remains high for all durations on the S^+ side. The gradients for the *Type 2* task for pigeons P12, P16 and P18 belong to this category.

Curiously, the pigeons that learned the *Type 1* task in the first place (see left panels of Fig. 2) had gradients for the *Type 2* task that decreased for the shortest durations, whereas the pigeons that learned the *Type 2* task in the first place (see right panels of Fig. 2) had gradients for the *Type 2* task that remained high for the shortest durations. The following hypothesis may explain these findings. Consider a pigeon that before learning the *Type 2* task "6 s → Blue; 18 s → Yellow" learned that 2-s samples are associated with Red. In the stimulus generalization test with Blue and Yellow, when the shortest durations of 0.7 s and 1.2 s are presented, Red would be the most probable response because these durations are close to 2 s. But because Red is not available for choice, the pigeon may be indifferent between Blue and Yellow. For this pigeon, then, the gradient for Blue would be around .50 following the shortest samples, similar to the results displayed in the left panels of Fig. 2 (empty circles). On the other hand, consider a pigeon that learns only the *Type 2* task "6 s→ Blue; 18 s→ Yellow". When presented with samples 0.7-s and 1.2-s long and the Blue and Yellow keys, this pigeon

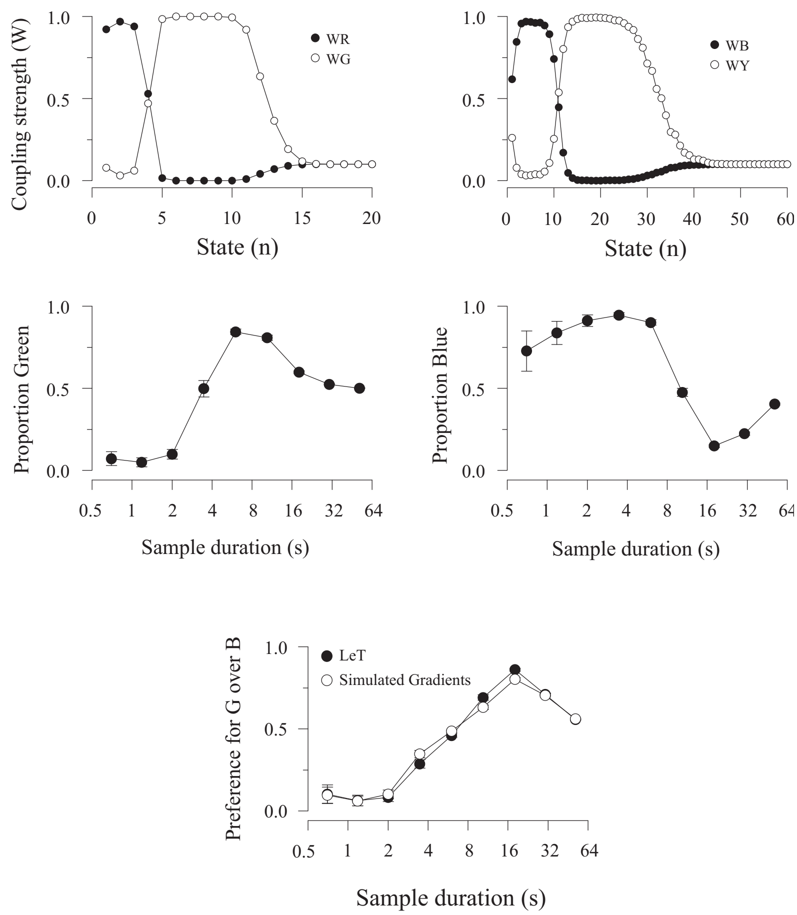

Fig. 7. Results from the simulations of the Learning-to-Time model (Machado et al., 2009). Top panel. Strength of the couplings between the behavioral states and the responses Red and Green (left panel) and Blue and Yellow (right panel) at the end of training. Middle panels. Temporal generalization gradients for Green (left panel) and Blue (right panel) obtained during the stimulus generalization tests. Bottom panel. Preference for Green over Blue in the stimulus-response generalization test. The filled circles show the preference function generated by the model and the empty circles show the preference function predicted from the generalization gradients. The vertical bars on the middle and bottom panels show the standard deviation. Note the logarithmic scale on the *x*-axis of the middle and bottom panels.

prefers Blue because these durations (a) have not been associated with any other key, and (b) are closer to the 6-s duration associated with Blue than to the 18-s duration associated with Yellow. For this pigeon, then, the gradient for Blue would remain high following the shortest samples, similar to the results displayed in the right panels of Fig. 2 (empty circles).

Extended to the gradients for the *Type 1* task (filled circles in Fig. 2), the foregoing hypothesis would predict that a pigeon that had learned the *Type 2* task previously, when given a choice between Red and Green following a long sample would be indifferent because the long sample, close to 18 s, would have been associated with Yellow, a color not available for choice. Hence, according to the hypothesis the gradient for this pigeon would be close to .50 following the longest samples, similar to the results displayed in the right panels of Fig. 2. However, the hypothesis predicts incorrectly that a pigeon that learned only the *Type 1* task would <u>not</u> show a decreasing proportion of Green choices following the longest samples, unlike the results displayed in the left panels of Fig. 2 (filled circles). It is possible that choice proportions depend not only on previous training (e.g., whether a sample duration has been associated with a choice key that is now unavailable), but also on the absolute distance between the trained and tested sample durations (e.g., the difference between 0.7 s and 6 s is much smaller than the difference between 51.4 s and 18 s).

4.1. Data versus model

In what follows we compare our findings with the results of a simulation of the Learning-to-Time (LeT) model. Consider first the training on one of the temporal discriminations, say the *Type 1* task. According to LeT, on each trial, the onset of the sample initiates the serial activation of the behavioral states. The activation spreads across successive states at rate λ, a value sampled at trial onset from a normal distribution with mean $\mu = 1$ per second and standard deviation $\sigma = 0.4$ per second. At the end of the sample, one state, say, n^*, is active. This state is linked with the available responses, Red and Green, and the strengths of the links are represented by $WR(n^*)$ and $WG(n^*)$, respectively. According to the model, the probability of choosing Green, say, equals the ratio $WG(n^*)/[WG(n^*) + WR(n^*)]$. At the end of the trial, the strengths of the links change according to the trial outcome, increasing with reinforcement and decreasing with extinction. If the link between the active state and one response increases, the link between the active state and the other response decreases and, conversely, if the link between the active state and one response decreases, the link between that state and the other response increases (response competition). To illustrate, if at the end of the sample the animal chooses Red and the response is reinforced, $WR(n^*)$ increases by the amount $\Delta WR(n^*) = \beta(1 - WR(n^*))$ and $WG(n^*)$ decreases by the amount $\Delta WG(n^*) = -\beta WG(n^*)$), where $\beta = 0.15$ is the reinforcement

parameter; if the animal chooses Red and the response is extinguished, WR(n^*) decreases by the amount ΔWR(n^*) = $-\alpha$WR(n^*) and WG(n^*) increases by the amount ΔWG(n^*) = α(1 − WG(n^*)), where $\alpha = 0.06$ is the extinction parameter[1].

In the present simulation the coupling of the behavioral states with Red and Green was initialized at 0.1 (i.e., WR(n) = WG(n) = 0.1, for n = 1,2,3,. . .). After 30 sessions of training, with Red reinforced following 2-s samples and Green reinforced following 6-s samples, the strength of the couplings changed to the values depicted in the top left panel of Fig. 7. The states around n = 2, more likely to be active after 2-s samples, are strongly coupled with Red (first three filled circles close to (1) and weakly coupled with Green (first three empty circles close to 0). The states around n = 6, more likely to be active after 6-s samples, are strongly coupled with Green but weakly coupled with Red. The remaining states, almost never active during the *Type 1* task, maintained their initial coupling of 0.1 with the two responses.

For the *Type 2* task the simulation process was the same except that the sample durations were 6 s and 18 s and two different vectors linked the states to the Blue and Yellow responses, WB(n) and WY(n), respectively. The top right panel of Fig. 7 shows that, after 30 sessions of training, the states around n = 6, more likely to be active after 6-s samples, are strongly coupled with Blue and weakly coupled with Yellow. The states around n = 18, more likely to be active after 18-s samples, are strongly coupled with Yellow but weakly coupled with Blue. The states further down the series retained their initial coupling with the two responses.

The middle panels of Fig. 7 show the stimulus generalization gradients for Green and Blue generated by the model. The simulated gradients reflect the profiles of the associative links displayed in the upper panels. The gradient for Green (*Type 1* task – middle left panel) is low following 2 s and shorter samples, then increases until 6-s samples and decreases again, reaching .50 for longer samples. The gradient for Blue (*Type 2* task – middle right panel) starts around .75 following 0.7-s samples, then increases slightly until 3.5-s and 6-s samples, decreases to a minimum value following 18-s samples and finally increases again, reaching about .40 following the 51.4-s samples.

Finally, in the stimulus-response generalization test, when Green and Blue occur together, the animal chooses Green with probability WG(n^*)/(WG(n^*)+WB(n^*)). The filled circles in the bottom panel of Fig. 7 show the preference function predicted by LeT. The preference for Green is low following 2-s samples and shorter samples, increases monotonically from 2-s to 18-s samples – the context effect, and decreases for even longer samples, reaching .55 following the 51.4-s samples.

One of the major goals of the present study was to determine if the temporal generalization gradients for Green and Blue obtained following training on the two tasks "2 s → Red; 6 s → Green" and "6 s → Blue; 18 s → Yellow" could predict the data from the final preference test between Green and Blue, as was previously found by Vieira de Castro and Machado (2012). There is, however, one important difference between the procedure of the present study and the procedure used by Vieira de Castro and Machado (2012), which makes this type of reasoning more difficult in the present case. Vieira de Castro and Machado (2012) used a go/no-go version of the double bisection task in which they reinforced Green following 4-s samples but not following 1-s samples and reinforced Blue following 4-s samples but not following 16-s samples. The generalization gradients that they obtained after the pigeons learned each task expressed WG(n) and WB(n) directly. That is, the gradients expressed directly the strength of the links used to compute the preference for Green over Blue during the final test. In the present study the generalization gradients express directly, not WG(n) and WB(n), but the ratios WG(n)/(WG(n) + WR(n)) and WB(n)/(WB(n) + WY(n)). Computing a preference function based on the relative strengths of Green over Red and of Blue over Yellow is not equivalent to predicting preference based on the absolute strengths of Green and Blue alone[2]. Be that as it may, to determine whether in LeT the separate generalization gradients predict well the context effect, we performed the same analysis that we carried out for the pigeons. That is, for each sample duration, we divided the proportion of responses to Green over Red by the proportion of responses to Green over Red plus the proportion of responses to Blue over Yellow. The result was the function depicted by the empty circles in the bottom panel of Fig. 7. The filled circles in the same panel show the results generated by the model in the tests between Green and Blue. The two functions were very similar. We conclude that, in LeT, the generalization gradients for Green over Red and Blue over Yellow are reliable predictors of the preference for Green over Blue.

Fig. 8 allows a direct comparison between LeT's predictions and the data from the present study. The two top panels show the stimulus generalization gradients for Green (left panel) and Blue (right panel) generated by the model (filled circles) and produced by the pigeons (empty circles). The simulated gradient for Green reproduced well the average of the gradients produced by the pigeons. The simulated gradient for Blue approximated the average of the gradients obtained in the present study except for the two longest samples.

The gradients produced by the pigeons for the *Type 2* task showed appreciable variability across subjects. For example, the gradient for pigeon P12 (see Fig. 2) differed from the remaining five gradients because it increased to .50 for the 51.4-s samples, a result in agreement with the present simulation. The greatest variability across subjects, however, was found for the shortest samples: Whereas some gradients remained high, others decreased towards .50 (see Fig. 2). Interestingly, the model also generated the highest variability for the 0.7-s and 1.2-s samples. Because we ran a high number of simulations (n = 100) this variability is not expressed in the error bars (SEM) in the top right panel of Fig. 8. However it can be seen in the middle right panel of Fig. 7 that the standard deviation in the simulated gradient for the *Type 2* task was higher for the two shortest samples. The variability in the simulated gradient for the *Type 2* task is due to the fact that, during training with 6-s and 18-s samples, the probability that the first behavioral states are active at the end of the sample is a rare but possible event. When the event happens during training, the links between these first states and the Blue and Yellow responses change substantially due to the values of the learning parameters β and α. In particular, the links with Blue will tend to be strengthened and the links with Yellow will tend to be weakened, because the probability of these first behavioral states being active following 18-s samples is practically zero. Subsequently, during generalization tests, these subjects will be more likely to choose Blue after very short samples because these samples activate the first states. However, if the rare event does not happen during training, that is, if the first states are not activated during the 6-s or 18-s samples, the links between these states and the two responses retain their initial (and equal) value. During the generalization tests, these subjects will be indifferent between Blue and Yellow after very short samples. Hence, according to Let, the variability across subjects following the shortest samples stems from the fact that some subjects prefer Blue whereas other subjects are indifferent between Blue and Yellow.

[1] The values of the parameters remained constant during the entire simulation and were similar to those used in previous studies (see Machado et al., 2009).

[2] Relative strengths can predict the test involving Green and Blue if the total link strength in the two discriminations is equal, i.e., if WG(n) + WR(n) = WB(n) + WY(n).

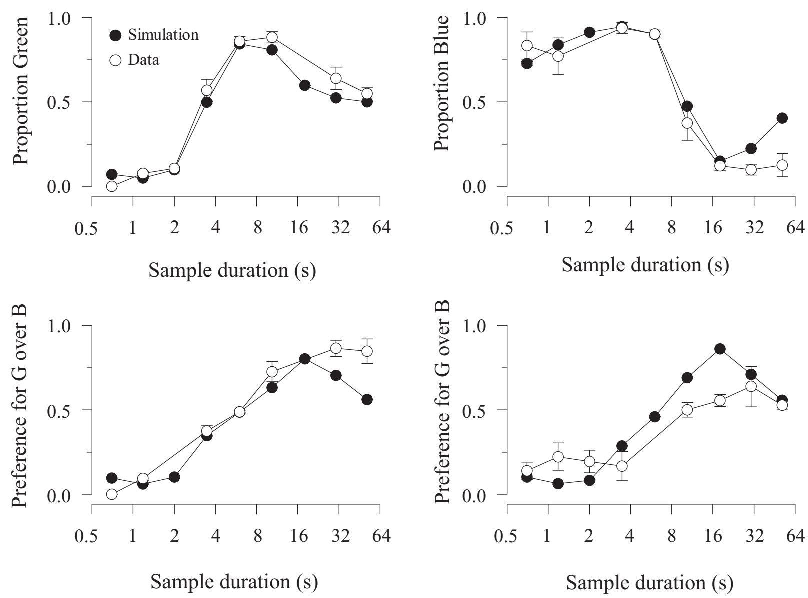

Fig. 8. Comparison between LeT's predictions and the data from the present study. Top panels. Temporal generalization gradients for Green (left panel) and Blue (right panel) predicted by LeT (filled circles) and produced by the pigeons (empty circles). Bottom left panel. Preference for Green over Blue computed from LeT's simulated gradients (filled circles) and from the obtained gradients (empty circles). Bottom right panel. Preference for Green over Blue generated by LeT (filled circles) and obtained in the present study (empty circles). The vertical bars show the SEM. Note the logarithmic scale on the *x*-axis.

The bottom left panel of Fig. 8 shows the preference functions predicted from the generalization gradients for LeT (filled circles) and for the birds (empty circles). The two functions are somewhat different, especially at the two longest durations. These differences reflect the differences in the gradients for the *Type 2* task generated by LeT and by the pigeons, for the model predicted that the proportion of responses to Blue should increase for the longest samples but only one pigeon (P12, see Fig. 2) showed this pattern.

Finally, the bottom right panel of Fig. 8 displays the preference for Green over Blue generated by LeT (filled circles) and the average preference for Green over Blue produced by our pigeons (empty circles). The model accounted well for the overall trend of the average data. However, two differences are important. First, LeT predicted a preference function steeper than the obtained function. And second, whereas LeT predicted that the preference for Green should be highest following 18-s samples and then decrease significantly for longer samples, the obtained function showed a maximum following 30.4-s samples and then decreased only slightly for the longest sample duration.

In summary, LeT accounts reasonably well for the present findings. The simulated generalization gradients are consistent with the gradients produced by our pigeons, except for the durations longer than 18 s in the *Type 2* task. In the critical test between Blue and Green, model and data also yielded similar preference functions, although the pigeons' average function was shifted to the right relative to the model's preference function.

To conclude, the present study showed that, in a temporal double bisection task in which the test range was extended beyond the training range (1) the preference for Green over Blue increased with sample duration; (2) the context effect (i.e., the preference for Green increasing with sample duration within the training range) was reproduced, although it was weaker than in previous studies; and (3) the temporal generalization gradients induced by the two temporal discriminations predicted well the overall pattern of preference, but overestimated the strength of the effect of sample duration in the preference for Green over Blue. LeT reproduced the major features of the data.

Acknowledgements

The authors thank Saulo Missiaggia Velasco for all the valuable help given during the data collection at the University of São Paulo. The authors also thank the students from the Animal Learning and Behavior laboratory of the University of Minho for their helpful comments on the paper. Ana Catarina Vieira de Castro was supported by a PhD fellowship and Armando Machado by a grant from the Portuguese Foundation for Science and Technology (FCT). Gerson Yukio Tomanari was supported by a grant from the Brazilian National Council for Science and Technology (CNPq).

References

Allan, L.G., Gibbon, J., 1991. Human bisection at the geometric mean. Learn. Motiv. 22, 39–58.
Arantes, J., 2008. Comparison of Scalar Expectancy Theory (SET) and the Learning to-Time (LeT) model in a successive temporal bisection task. Behav. Process. 78, 269–278.
Arantes, J., Machado, A., 2008. Context effects in a temporal discrimination task: Further tests of the Scalar Expectancy Theory and Learning-to-Time models. J. Exp. Anal. Behav. 90, 33–51.
Catania, A.C., 1970. Reinforcement schedules and the psychophysical judgments: a study of some temporal properties of behavior. In: Schoenfeld, W.N. (Ed.), The Theory Of Reinforcement Schedules. Appleton-Century-Crofts, New York, pp. 1–42.
Church, R.M., A concise introduction to scalar timing theory 2003. In: Warren, H., Meck (Eds.), Functional and Neural Mechanisms Of Interval Timing. CRC Press, Boca Raton, FL, pp. 3–22.
Church, R.M., Deluty, M.Z., 1977. Bisection of temporal intervals. J. Exp. Psychol.: Anim. Behav. Process. 3, 216–228.
Crystal, J.D., 2002. Timing inter-reward intervals. Learn. Motiv. 33, 311–326.
Elsmore, T.F., 1971. Control of responding by stimulus duration. J. Exp. Anal. Behav. 16, 81–87.
Fetterman, J.G., Killeen, P.R., 1991. Adjusting the pacemaker. Learn. Motiv. 22, 226–252.
Gallistel, C.R., 1990. The Organization of Learning. Bradford Books/MIT Press, Cambridge, MA.
Gibbon, J., 1977. Scalar expectancy theory and Weber's law in animal timing. Psychol. Rev. 84, 279–325.
Gibbon, J., 1991. Origins of scalar timing theory. Learn. Motiv. 22, 3–38.
Killeen, P.R., Fetterman, J.G., 1988. A behavioral theory of timing. Psychol. Rev. 95, 274–295.
Lejeune, H., Wearden, J.H., 2006. Scalar properties in animal timing: conformity and violations. Q. J. Exp. Psychol. 59, 1875–1908.

Machado, A., 1997. Learning the temporal dynamics of behavior. Psychol. Rev. 104, 241–265.
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. Behav. Process. 72, 195–206.
Machado, A., Keen, R., 1999. Learning to Time (LeT) or Scalar Expectancy Theory (SET)? A critical test of two models of timing. Psychol. Sci. 10, 285–290.
Machado, A., Malheiro, M.T., Erlhagen, W., 2009. Learning to time: A perspective. J. Exp. Anal. Behav. 92, 423–458.
Machado, A., Pata, P., 2005. Testing the Scalar Expectancy Theory (SET) and the Learning-to-Time model (LeT) in a double bisection task. Learn. Behav. 33, 111–122.
Mellgren, R.L., Mays, M.Z., Haddad, N.F., 1983. Discrimination and generalization by rats of temporal stimuli lasting for minutes. Behav. Process. 14, 75–91.
Oliveira, L., Machado, A., 2008. The effect of sample duration and cue on a double temporal discrimination. Learn. Motiv. 39, 71–94.
Oliveira, L., Machado, A., 2009. Context effect in a temporal bisection task with the choice keys available during the sample. Behav. Process. 81, 286–292.
Platt, J.R., Davis, E.R., 1983. Bisection of temporal intervals by pigeons. J. Exp. Psychol.: Anim. Behav. Process. 9, 160–170.
Richelle, M., Lejeune, H., 1980. Time in Animal Behavior. Pergamon Press, Oxford, UK.
Russell, R., Kirkpatrick, K., 2007. The role of temporal generalization in a temporal discrimination task. Behav. Process. 74, 115–125.
Siegel, S.F., 1986. A test of the similarity rule model of temporal bisection. Learn. Motiv. 17, 59–75.
Spetch, M.L., Cheng, K., 1998. A step function in pigeons' temporal generalization in the peak shift task. Anim. Learn. Behav. 26, 103–118.
Stubbs, D.A., 1968. The discrimination of stimulus duration by pigeons. J. Exp. Anal. Behav. 11, 223–238.
Vieira de Castro, A.C., Machado, A., 2012. The interaction of temporal generalization gradients predicts the context effect. J. Exp. Anal. Behav. 97, 263–279.
Wearden, J.H., 1991. Human performance on an analogue of the interval bisection task. Q. J. Exp. Psychol.: Compar. Physiol. Psychol. 43, 59–81.