Available online at www.sciencedirect.com
ScienceDirect
Learning and Motivation 39 (2008) 71–94
Learning and Motivation
www.elsevier.com/locate/l&m

The effect of sample duration and cue on a double temporal discrimination ☆

Luís Oliveira, Armando Machado *

Instituto de Educação e Psicologia, Universidade do Minho, 4710 Braga, Portugal

Received 22 March 2007; received in revised form 8 June 2007
Available online 6 September 2007

Abstract

To test the assumptions of two models of timing, Scalar Expectancy Theory (SET) and Learning to Time (LeT), nine pigeons were exposed to two temporal discriminations, each signaled by a different cue. On half of the trials, pigeons learned to choose a red key after a 1.5-s horizontal bar and a green key after a 6-s horizontal bar; on the other half of the trials, they learned to choose a blue key after a 6-s vertical bar and a yellow key after a 24-s vertical bar. During subsequent test trials, they were exposed to the horizontal or vertical bar, for durations ranging from 1.5 to 24 s, and given a choice between novel key combinations: red vs. yellow, or green vs. blue. Results showed a strong effect of sample duration—as the test signal duration increased, preference for green over blue increased and preference for red over yellow decreased. The effect of sample cue was obtained only on the green-blue test trials. These effects are discussed in light of SET and LeT.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Double temporal bisection; Learning-to-Time model; Psychometric function; Scalar Expectancy Theory; Timing; Key peck; Pigeons

☆ The authors thank Francisco Silva for helpful comments. Research supported by a grant from the Portuguese Foundation for Science and Technology (FCT) to the second author.
* Corresponding author.
E-mail address: armandom@iep.uminho.pt (A. Machado).

0023-9690/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.lmot.2007.06.001

Introduction

When the occurrence of an important biological event such as food is predictable on the basis of a temporal cue, birds and mammals typically change their behavior in the presence of that cue. For example, when a pigeon receives food every minute for pecking a key, the pigeon learns to pause from pecking immediately after food and then peck at a sustained or increasing rate until the next food; double the interfood interval, and the pigeon will double its initial pause. Similarly, when a rat receives food for pressing a lever after a 4 s signal, but not after shorter or longer signals, it learns to respond mostly after signal durations equal or close to 4 s; double the critical duration and the rat will come to respond mostly at 8 s (for summaries see, e.g., Catania, 1970; Church & Gibbon, 1982; Dews, 1970; Richelle & Lejeune, 1980; Roberts, 1998; Shettleworth, 1998).

The ability to change behavior according to the temporal attributes of stimuli is often interpreted as the outcome of an underlying timing process. To understand this process, researchers have developed and tested quantitative models of it. The present study continues a series of experiments that tested the predictions of two of these models, the Scalar Expectancy Theory (SET), an information-processing theory developed by Gibbon and collaborators (e.g., Gibbon, 1977, 1991), and Learning to Time (LeT), a behavioral model developed by Machado and collaborators (e.g., Machado, 1997; Machado & Cevik, 1998) on the basis of previous work by Killeen and Fetterman (1988). To contrast the two models, we have designed a new temporal discrimination task which consists of two temporal bisection tasks combined (see Machado & Arantes, 2006; Machado & Keen, 1999; Machado & Pata, 2005). Because this double bisection task is relatively new, in what follows we describe the task and then derive the models' predictions for it.

A *simple* temporal bisection task is a conditional discrimination task in which two sample stimuli differing only in duration are mapped to two comparison stimuli. A pigeon sees a houselight lit for either 1 s or for 4 s and then chooses between a red key and a green key. The choice of Red is rewarded after the 1 s sample and the choice of Green is rewarded after the 4 s sample. This conditional discrimination may be represented by a mapping between the stimulus pair (S_1, S_4) and the response pair (Red, Green),

$$\{S_1, S_4\} \rightarrow \{\text{Red, Green}\},$$

where the subscripts remind us of the sample durations involved and the arrow means that the first response is rewarded following the first sample and the second response is rewarded following the second sample.

A *double* temporal bisection task comprises two of these simple discriminations. To follow the preceding example, suppose a pigeon initially learns the mapping $\{S_1, S_4\} \rightarrow$ {Red, Green} and then learns a second mapping, namely,

$$\{S_4, S_{16}\} \rightarrow \{\text{Blue, Yellow}\},$$

that is, it learns to choose a blue key following a 4 s sample and a yellow key following a 16 s sample. Finally the two mappings are integrated in the same session such that half of the trials are of the first type, $\{S_1, S_4\} \rightarrow$ {Red, Green}, and the other half are of the second type, $\{S_4, S_{16}\} \rightarrow$ {Blue, Yellow}.

When the two discriminations are well learned, the experimenter may introduce test trials to examine stimulus generalization or stimulus-response generalization. To examine stimulus generalization, the experimenter varies the sample duration from 1 to 4 s and then gives the pigeon a choice between Red and Green, or varies the sample duration from 4 to 16 s and then gives the pigeon a choice between Blue and Yellow. These stimulus generalization tests may be represented by the following notation,

$\{S_{1..4}\}$:{Red, Green}
$\{S_{4..16}\}$:{Blue, Yellow},

where $\{S_{1..4}\}$ and $\{S_{4..16}\}$ mean that the samples range from 1 to 4 s and from 4 to 16 s, respectively; a colon replaces the arrow to indicate that no choice is rewarded during these test trials. The test results consist of two psychometric functions, one showing how preference for Red (over Green) changes with sample duration and the other showing how preference for Blue (over Yellow) changes with sample duration. One question that may be asked is, 'Are the two psychometric functions scale transforms of each other? Or, equivalently, do they superpose when plotted against relative stimulus duration?' The superposition question is important because, as we shall see, the SET and LeT models answer it differently.

To examine stimulus-response generalization, the experimenter varies both the sample durations and the choice sets. Specifically, the experimenter presents samples ranging in duration from 1 to 16 s and then gives the pigeon a choice between two keys that it had not seen together before (e.g., Green and Blue). The entire set of test trials is represented by the following notation

$\{S_{1..16}\}$:{Red, Blue}
:{Red, Yellow}
:{Green, Blue}
:{Green, Yellow},

where the samples range in duration from 1 to 16 s, and the new choice sets include always one element from the {Red, Green} training set and one element from the {Blue, Yellow} training set. The test results consist of four psychometric functions, each showing the effect of sample duration on the preference for Red or Green. One question that may be asked is, 'What are the shapes of the four psychometric functions?' In the present study we limit our attention to two of them, the functions involving the {Red, Yellow} and the {Green, Blue} sets. The latter in particular is crucial because, as we shall see below, SET and LeT predict substantially different shapes for that function. We derive each model's predictions for the double bisection task next.

Scalar Expectancy Theory

The top left panel of Fig. 1 shows the model's structure. A pacemaker generates pulses at a high and variable rate, an accumulator counts the pulses emitted during the sample, and two or more long-term memory stores save the number that is in the accumulator when a choice response is rewarded. In a $\{S_1, S_4\} \rightarrow$ {Red, Green} task, the numbers that are in the accumulator when a choice of Red is rewarded are saved into one memory store

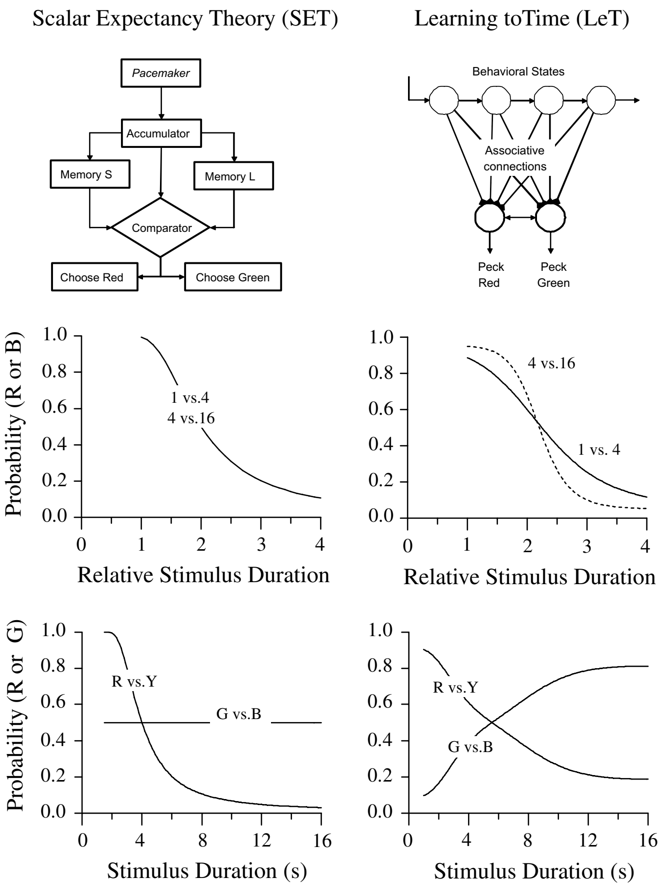

Fig. 1. Top: Structure of the two models, Scalar Expectancy Theory (SET) and the Learning to Time (LeT). Middle: Predictions of SET and LeT for the stimulus generalization test phase, $\{S_1..S_4\}$:{Red, Green} and $\{S_4..S_{16}\}$:{Blue, Yellow}; the two psychometric functions overlap for SET but not for LeT. Bottom: Predictions for the stimulus-response generalization test phase, $\{S_1..S_{16}\}$:{Red, Yellow} and $\{S_1..S_{16}\}$:{Green, Blue}; in the critical test, only LeT predicts that preference for Green over Blue increases with sample duration.

and the numbers that are in the accumulator when a choice of Green is rewarded are saved into another memory store. We may identify the two memory stores by M_{Red} and M_{Green}, respectively, and note that their contents represent the animal's learning history. According to SET, the choice after a sample with duration T depends on a comparison involving three numbers: X_T, the number of pulses in the accumulator at the end of the sample; X_S, a number extracted from M_{Red} and representing the short stimulus; and X_L, a number extracted from and M_{Green} and representing the long stimulus. Specifically, if $(X_T/X_S) < (X_L/X_T)$, then the organism is more likely to choose the Red or "Short" response, but if $(X_T/X_S) > (X_L/X_T)$, then the organism is more likely to choose the Green or "Long" response. SET predicts indifference (the point of subjective equality) when $(X_T/X_S) = (X_L/X_T)$, which is equivalent to, $X_T = \sqrt{}(X_S{*}X_L)$, that is, when X_T equals the geometric mean of the training durations.

The foregoing account for the simple bisection task extends straightforwardly to the double bisection task (for details see Machado & Arantes, 2006, and Machado & Pata, 2005). During training, the animal forms four memory stores, M_{Red} and M_{Green}, which represent 1 and 4 s samples, respectively, and M_{Blue} and M_{Yellow}, which represent 4 and 16 s samples, respectively. On a test trial, the sample duration represented by X_T is compared with samples extracted from the memory stores associated with the choice keys presented on the trial.

The middle and bottom left panels of Fig. 1 illustrate SET's predictions for the double bisection task. First, because the durations of the two training sets, $\{S_1, S_4\}$ and $\{S_4, S_{16}\}$, are in the same 1-to-4 ratio, the psychometric functions obtained during the stimulus generalization tests will superpose when plotted on a common scale. Moreover, the point of subjective equality will be located at the geometric mean of the two training durations (see Gibbon, 1981, 1991). This prediction is illustrated in the middle left panel. Second, because the choices of the Green and Blue keys are rewarded following samples of 4 s, the memory stores M_{Green} and M_{Blue} will be statistically identical, that is, they will contain the same distribution of counts. Hence, during the stimulus-response generalization tests in which the sample duration varies from 1 to 16 s and the choice keys are Green and Blue, preference for one of the colors, say, Green, will not vary with sample duration and the psychometric function $\{S_{1..16}\}$:{Green, Blue} will be horizontal. This prediction is illustrated in the bottom left panel of Fig. 1. Third, in contrast, preference for Red when the choice is between Red and Yellow will decrease with the sample duration, and the psychometric function $\{S_{1..16}\}$:{Red, Yellow} will be monotonically decreasing with a point of subjective equality at 4 s, the geometric mean of 1 and 16 s. This prediction also is illustrated in the bottom left panel of Fig. 1.

Learning-to-Time model

The top right panel of Fig. 1 shows LeT's structure. The model assumes a serial organization of behavioral states which function as cues for the choice response. These states embody our concepts of elicited, induced, adjunctive, interim, and terminal classes of behavior (see Killeen & Fetterman, 1988; Staddon, 1977; Staddon & Simmelhag, 1971; Timberlake & Lucas, 1985). According to LeT, they underlie the sequential and temporal organization of behavior. In the bisection task, the onset of the sample activates the first state in the series but, as time elapses, the activation of each state flows to the next state. The speed of the activation flow varies directly with the overall reinforcement rate, perhaps via changes in the animal's arousal level (e.g., Beam, Killeen, Bizo, & Fetterman, 1998). Note that the behavioral states in LeT are analogous to the pulses in SET, and the speed of the activation flow across the states in LeT is analogous to the speed of the pacemaker in SET.

LeT also assumes a matrix of associative links connecting the behavioral states to the choice responses. The strength of the link connecting state s with response r, w(s,r), can vary from 0 to 1. When a response is reinforced, the links between the active behavioral states and that response increase, whereas the links between those states and the other response decrease. Conversely, when a response is extinguished, the links between the active behavioral states and that response decrease, whereas the links between those states and the other response increase. At the steady state, the matrix of associative links

represents the animal's learning history. The associative links in LeT are analogous to the memory stores in SET.

On each trial, choice depends on which states are most active at the end of the sample and on the strength of the links between those states and the two responses. To illustrate, in a $\{S_1, S_4\} \rightarrow \{\text{Red, Green}\}$ task, after the 1-s sample only the first states (call them "Early") are active and because of the reinforcement contingencies, their link with Red will be strong whereas their link with Green will be weak (i.e., $w(\text{Early, Red}) \approx 1$ and $w(\text{Early, Green}) \approx 0$)—hence the preference for Red after short samples. However, after the 4-s samples, later states (call them "Middle" for reasons that will become clear below) will be the most active and because of the reinforcement contingencies $w(\text{Middle, Red}) \approx 0$ and $w(\text{Middle, Green}) \approx 1$—hence the preference for Green after the long samples. In summary, LeT predicts that preference for Red decreases as sample duration ranges from 1 to 4 s. Moreover, it can be shown (see Machado, 1997) that LeT predicts a point of subjective equality close to, but slightly greater than, the geometric mean of the training stimuli.

For the double bisection task, a straightforward extension of LeT makes the following predictions (see the middle and bottom right panels of Fig. 1). First, the two stimulus generalization curves will *not* superpose. This prediction stems from LeT's assumption that the activation flow across the behavioral states follows a Poisson process (for quantitative details and supporting evidence see Killeen & Fetterman, 1988, and Machado, 1997). As a result, the ratio of the standard deviation to the mean (i.e., the coefficient of variation) decreases with the interval to be timed, which means that the relative accuracy of a "Poisson clock" increases with sample duration. According to Let, then, the psychometric function for the $\{S_4..S_{16}\}$:{Blue, Yellow} discrimination should be steeper than the psychometric function for the $\{S_1..S_4\}$:{Red, Green} discrimination. This prediction is illustrated in the middle panel of Fig. 1 (see also Machado & Keen, 1999).

The foregoing prediction may be understood in a different way. The LeT model predicts superposition of the psychometric functions only when the two discriminations $\{S_1..S_4\}$:{Red, Green} and $\{S_4..S_{16}\}$:{Blue, Yellow} are trained under different reinforcement rate contexts (e.g., the two discriminations are trained in different sessions or within the same session but signaled by distinctive cues; see Bizo & White, 1994, 1995a,b; Fetterman & Killeen, 1991; Killeen & Fetterman, 1988; Machado, 1997; Morgan, Killeen, & Fetterman, 1993). For in this case the difference in the reinforcement rates on the two types of trials—a fourfold difference favoring the $\{S_1..S_4\}$:{Red, Green} discrimination—will occasion a fourfold difference in the speed of the activation flow across the states (Killeen & Fetterman, 1988; Machado, 1997), which will then cancel the fourfold differences in sample durations. In other words, if the activation flows four times faster during the relatively short discrimination, but the sample durations are four times shorter on that discrimination, the two effects cancel and the functions superpose. Because in the double bisection task the reinforcement context is the same for both types of trials, the activation flow also will be the same, and therefore superposition is not predicted.

Second, as the bottom right panel of Fig. 1 shows, when the choice is between the Green and Blue keys, both associated with 4-s samples, LeT predicts that the birds' preference for Green will increase monotonically with the sample duration. This is LeT's critical prediction for the stimulus-response generalization trials, the prediction that most distinguishes it from SET. To understand it intuitively (see Machado & Pata, 2005, for quantitative

details), divide the behavioral states into three sets, Early (most active at 1 s), Middle (most active at 4 s), and Late (most active at 16 s). Initially all states are associated equally with the four responses (i.e., $w = 0.5$ for all links), but these associations will change during training. Following the reasoning above, at the end of training with the mapping $\{S_1, S_4\} \rightarrow$ {Red, Green}, the states in the Early set become strongly associated with Red but, more importantly, weakly associated with Green (i.e., w(Early, Green) ≈ 0). Similarly, after training with the mapping $\{S_4, S_{16}\} \rightarrow$ {Blue, Yellow}, the states in the Late set become strongly associated with Yellow but, more importantly, weakly associated with Blue (i.e., w(Late, Blue) ≈ 0). It follows that, during test trials, after 1-s samples, the Early states are the most active, and because their links with Green have decreased to 0 whereas their links with Blue have remained at 0.5, the animal prefers Blue. Conversely, after 16 s samples, the Late states are the most active, and because their links with Blue have decreased to 0 whereas their links with Green have remained at 0.5, the animal prefers Green. Hence, LeT predicts that preference for Green increases with sample duration (see Fig. 1, bottom right panel). Third, when the choice is from Red and Yellow, LeT predicts a decreasing preference for Red as the sample duration increases.

The results of three studies (Machado & Arantes, 2006; Machado & Keen, 1999; Machado & Pata, 2005) were generally closer to LeT's predictions than to SET's predictions. In particular, Machado and Keen (1999) did not obtain superposition of the stimulus generalization functions and, in the critical stimulus-response generalization tests with the Green and Blue keys, all three studies revealed an increasing preference for Green with the sample duration. The results also showed that, as SET and LeT predicted, preference for Red over Yellow decreased with the sample duration.

At the present, we do not know the generality of the preceding results, whether they hold with different ranges of sample durations, different choice responses, different species, or even different arrangements of the basic double bisection procedure. Concerning this last category, consider that in all three studies with the double bisection procedure, the sample stimulus (a white light) remained the same on all trials, which means that the pigeon could anticipate neither the relative trial duration nor the set of comparison keys. In other words, on the basis of the sample cue the pigeon could not anticipate whether the trial would be relatively short (1 or 4 s long) or relatively long (4 or 16 s long), nor could it anticipate whether the choice keys at the end of the sample would be Red and Green or Blue and Yellow. The main purpose of the present study was to determine whether the standard results reported above are maintained when, during training, each trial type is signaled by a different sample cue and new sample durations are used. If the results are maintained, then LeT's account is strengthened considerably; if they are not, then the two models must be modified or rejected and a new model developed to account for the data.

To explore the potential consequences of different sample cues, suppose that two bars are used as sample cues, a horizontal bar during the relatively short trials (henceforth referred to as "Short set" trials) and a vertical bar during the relatively long trials ("Long set" trials). Fig. 2 illustrates the procedure with the durations used in the experiment: During training the sample durations on the "Short set" trials were 1.5 and 6 s, and on the "Long set" trials they were 6 and 24 s. Letting H and V stand for "Horizontal" and "Vertical" bars, respectively, the two training mappings are represented as follows,

$\{H_{1.5}, H_6\} \rightarrow$ {Red, Green}
$\{V_6, V_{24}\} \rightarrow$ {Blue, Yellow}.

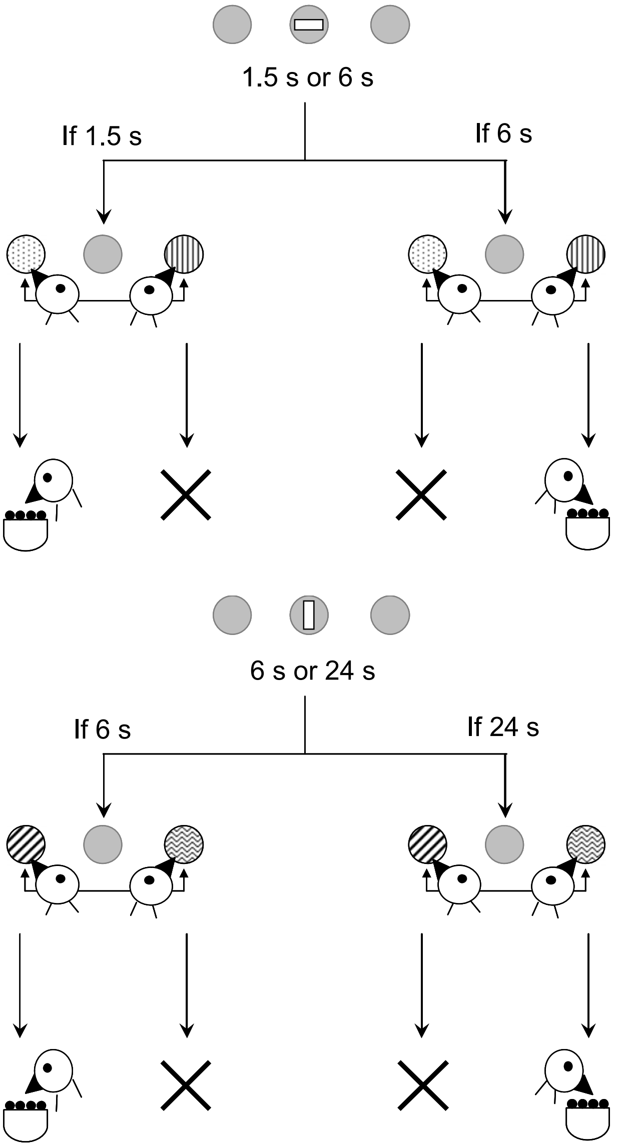

Fig. 2. Structure of the two types of regular trials. During the relatively short trials, the sample cue, a horizontal bar, is presented for 1.5 or 6 s. A red-key choice is reinforced if the sample was 1.5 s long, and a green-key choice is reinforced if it was 6 s long. During the relatively long trials, the sample cue, a vertical bar, is presented for either 6 or 24 s. After the former, the choice of blue is reinforced; after the latter, the choice of yellow is reinforced.

How could the two bars affect performance? One possibility is that because they signal two different local reinforcement rates, they could set the occasion for two different speeds of the pacemaker (SET) or of the transition across the behavioral states (LeT; for evidence consistent with this hypothesis, see Bizo & White, 1994, 1995a,b; Fetterman & Killeen,

1991; Killeen & Fetterman, 1988; Morgan et al., 1993). Changing the pacemaker or state transition speed could therefore change the psychometric functions obtained during the stimulus generalization trials, the stimulus-response generalization trials, or both. A second possibility is that the two cues would control different collateral behaviors, which some studies suggest may mediate the animals' choices in simple bisection tasks (Fetterman, Killeen, & Hall, 1998; Killeen & Fetterman, 1988; Richelle & Lejeune, 1980) and in double bisection tasks (Machado & Arantes, 2006; Machado & Keen, 1999; Machado & Pata, 2005). Therefore, changing collateral behaviors could also change the psychometric functions. A third possibility is that the two bars would bias the four choice responses differently because in their presence only some of these choice responses would have been reinforced. More specifically, because the horizontal bar would set the occasion for the reinforcement of Green but not of Blue, whereas the vertical bar would set the occasion for the reinforcement of Blue but not Green, given a choice between Green and Blue, preference for Green would be higher when the sample was the horizontal bar than when it was the vertical bar. This prediction would hold for all sample durations.

The experiment reported below was not designed to contrast these three possibilities, but rather to determine whether signaling the two types of training trials with distinctive cues and using new duration ranges changes the generalization functions obtained with a single cue. The empirical clarification of this issue must precede any test of the hypotheses concerning the underlying processes responsible for the potential sample cue effect. To that end, we trained pigeons in a double bisection task with two distinctive cues and then conducted stimulus-response generalization tests with each sample cue,

$\{H_{1.5}..H_{24}\}$:{Green, Blue}
$\{H_{1.5}..H_{24}\}$:{Red, Yellow},

and

$\{V_{1.5}..V_{24}\}$:{Green, Blue}
$\{V_{1.5}..V_{24}\}$:{Red, Yellow}.

By comparing the four psychometric functions we expected to determine the effects of sample duration and sample cue on choice performance. Finally, we conducted also stimulus generalization tests to determine whether the two psychometric functions, $\{S_{1.5}..S_{6}\}$:{Red, Green} and $\{S_{6}..S_{24}\}$:{Blue, Yellow}, superposed when plotted on a common scale. The results of these generalization tests will pave the way for better models of temporal discrimination.

Methods

Subjects

Nine pigeons (*Columba livia*) maintained at 80% of their free-feeding weights participated in the experiment. Eight of these pigeons were experimentally naïve and one (P555) had experience with a concurrent timing task. Water and grit were continuously available in their home cages, and a 14:10 h light/dark cycle (lights on at 7:00 a.m.) was in effect in the pigeon colony.

Apparatus

Four identical LehighValley® experimental chambers for pigeons were used. The front panel of each chamber contained three keys centered on the wall, 2.5 cm in diameter, 8 cm apart center to center, and 22 cm above the wire mesh floor. The side keys could be illuminated from behind with red, green, blue, or yellow light. When illuminated, the center key displayed a horizontal or a vertical white bar over a dark background. A 6×5 cm hopper opening centered on the wall directly below the center key gave access to mixed grain when it was raised and illuminated with a 7.5-W white light. Another 7.5-W white light, situated on the back wall of the chamber, provided general illumination. Each panel was enclosed by an external box equipped with a fan to provide ventilation and mask extraneous noises. Four personal computers programmed in C++ controlled the experimental events and recorded the data.

Procedure

The experiment was divided into four phases: separate training on the two basic discriminations, combined training on the two discriminations, stimulus-response generalization testing, and stimulus generalization testing. Table 1 shows the procedural details for each bird.

Separate training

After the birds learned to peck the keys using an autoshaping procedure, they were exposed to one of the two simple discrimination tasks, one involving the relatively short durations of 1.5 and 6 s, and the other involving the relatively long durations of 6 and 24 s. Four birds started with one task and five started with the other task. The sample cue (horizontal or vertical) assigned to each task was counterbalanced across birds. The assignment of keylight colors to signal durations also was counterbalanced, but with two restrictions: The choice sets were always {Red, Green} and {Blue, Yellow}, and the green and blue colors were associated always with the two 6 s samples. However, for clarity, the procedure and the experimental results are described as though all birds had the following assignment: $\{H_{1.5}, H_6\} \rightarrow$ {Red, Green} and $\{V_6, V_{24}\} \rightarrow$ {Blue, Yellow} (see Fig. 2). The two colors presented during each trial type appeared always the same number of times on the left and right keys.

Sessions consisted of 60 trials and during each trial the following sequence of events occurred: The houselight was illuminated with white light and the center key was illuminated with a white bar for the duration of the sample. When the sample duration elapsed (e.g., 6 s), the center key was turned off and the side keys were illuminated with different colors (e.g., red and green). A peck at a choice key turned the keylights and the houselight off. If the choice was correct, the hopper was raised for 2 to 4 s, according to the pigeon. The hopper duration for each pigeon was adjusted during the first sessions to maintain body weight and minimize post-session feedings. Rewards were followed by a 30 s intertrial interval (ITI) during which all lights were off. If the subject made an incorrect choice, the ITI was initiated immediately and the trial was repeated (correction method). If the bird made three consecutive errors, then only the correct key was illuminated after the sample stimulus.

Table 1
Structure of the training and test trials

Pigeon	Training trials			Stimulus-response generalization trials				Stimulus generalization trials	
	Short	Long	First	$S^+ = 0$	$S^+ = .5$	$S^+ = .5$	$S^+ = 0$	Short	Long
205	$\{V_{1.5}, V_6\} \rightarrow \{R, G\}$	$\{H_6, H_{24}\} \rightarrow \{B, Y\}$	Long	V	H	V	H	$\{V_{1.5}..V_6\}:\{R, G\}$	$\{H_6..H_{24}\}:\{B, Y\}$
438			Short	H	V	H	V		
364	$\{H_{1.5}, H_6\} \rightarrow \{R, G\}$	$\{V_6, V_{24}\} \rightarrow \{B, Y\}$	Short	H	V	H	V	$\{H_{1.5}..H_6\}:\{R, G\}$	$\{V_6..V_{24}\}:\{B, Y\}$
812			Long	V	H	V	H		
G13	$\{V_{1.5}, V_6\} \rightarrow \{Y, B\}$	$\{H_6, H_{24}\} \rightarrow \{G, R\}$	Long	H	V	H	V	$\{V_{1.5}..V_6\}:\{Y, B\}$	$\{H_6..H_{24}\}:\{G, R\}$
G14			Short	V	H	V	H		
539	$\{H_{1.5}, H_6\} \rightarrow \{Y, B\}$	$\{V_6, V_{24}\} \rightarrow \{G, R\}$	Short	V	H	V	H	$\{H_{1.5}..H_6\}:\{Y, B\}$	$\{V_6..V_{24}\}:\{G, R\}$
499			Long	H	V	H	V		
555			Long	H	V	H	V		

The sample consisted of a vertical (V) or horizontal (H) bar for the duration identified by the subscript. The two choice keys were illuminated with red (R) and green (G), or with blue (B) and yellow (Y) lights. The mapping $\{V_{1.5}, V_6\} \rightarrow \{R, G\}$ means that the R and G choices were correct and reinforced after the 1.5 and 6 s samples, respectively. "First" refers to the basic discrimination trained first. During the Stimulus-Response generalization trials, the V and H sample cues alternated across blocks, the sample duration ranged from 1.5 to 24 s, the choice sets were {R, Y} and {G, B}, and the probability of reinforcement, S^+, was either 0 or 0.5 (non-differential reinforcement). During the Stimulus generalization trials, the mapping $\{H_6..H_{24}\}:\{B, Y\}$ means that the H sample cue ranged in duration from 6 to 24 s but only correct choices following the anchor durations of 6 and 24 s were reinforced.

Two pigeons (P499 and P539) consistently stood by and pecked one of the (dark) side keys during the sample, thereby making a large number of errors. To reduce these errors and increase the salience of the sample cue, we cancelled the trial and initiated the ITI immediately if the pigeons pecked the side keys during the sample, and we lengthened the ITI to 45 s. These changes reduced the number of errors considerably.

Once the birds learned the first basic discrimination (at least 80% correct choices, excluding repeated trials, for four consecutive sessions), they were exposed to the other discrimination task. After the second discrimination was learned, the two types of sessions, one with the "Short set" and the other with the "Long set" trials, alternated across days. This training phase lasted from 24 to 61 sessions (mean = 39).

Combined training

Each session comprised 32 "Short set" trials (i.e., $\{S_{1.5}, S_6\}$) and 32 "Long set" trials (i.e., $\{S_6, S_{24}\}$). Of each of these 32 trials, 16 presented the shorter sample of the set ($S_{1.5}$ or S_6) and 16 the longer (S_6 or S_{24}). The session ended after 64 reinforcers were collected. During the first five to seven sessions (mean = 5.4), correct choices were reinforced and incorrect choices led to trial repetition. These trials are called "regular trials".

During the next five to nine sessions (mean = 5.8), extinction trials were introduced to adapt the birds to the lower rate of food that would occur during the subsequent testing phases. Besides not ending with food, even after a correct choice, extinction trials were not repeated if the choice was incorrect. Sessions comprised 48 regular trials and 24 extinction trials. Training lasted until the birds achieved an average of at least 80% correct choices over five sessions, excluding repeated trials, on each of the four sample durations.

Stimulus-response generalization testing

This phase lasted for 20 sessions, and each session comprised 48 regular trials and 20 test trials. During the test trials, the sample cue, either the horizontal or the vertical bar (see Table 1), lasted for 1.5, 3, 6, 12, or 24 s. Furthermore, each sample was followed by a choice between the new color pairs Red-Yellow or Green-Blue. Five sample durations and two choice sets yield a total of ten distinct trials. Because each color was presented once on the left and once on the right key, there were 20 test trials overall.

A major difficulty in obtaining the stimulus-response generalization curves is the fact that the novel key pairings {Red, Yellow} and {Green, Blue} signal extinction and therefore as the birds experience them they may begin to pause and even stop; control of choice by the sample is necessarily reduced. It follows that the number of test trials should be as small as possible. However, the two generalization curves (one for each sample cue), each with 5 durations, cannot be estimated reliably if the number of test trials is excessively small. To reduce this difficulty, the 20 sessions were divided into four blocks of five sessions each. In each block, only one of the bars appeared during the test trials and the type of bar alternated across blocks, with order counterbalanced across birds (Table 1 shows the details). During the first and last blocks, the test trials were conducted in extinction, but, during the second and third blocks, 50% of the test trials, randomly selected, were reinforced regardless of the pigeon's choices. Finally, to reduce the effects of testing on the two basic discriminations, after each block the pigeons received three to six sessions with regular and extinction trials only (as in the Combined Training phase).

Stimulus generalization testing

This phase consisted of five sessions, each comprising 48 regular trials and 24 stimulus generalization test trials. For the test trials, two sets of logarithmically spaced durations were used: 2.1, 3.0, and 4.2 s for the 1.5-to-6 s range, and 8.5, 12.0, and 17.0 s for the 6-24 s range. In each set, the middle duration corresponds to the geometric mean of the training durations. Each of the six samples occurred four times per session, two for each left-right color assignment. Test trials were not reinforced.

Results

All subjects learned the basic discriminations. In the last five sessions of the Combined Training phase, proportion of correct responses averaged 0.92 across birds, ranging from 0.81 to 1.0 across the four sample durations. These results are consistent with the results obtained by Machado and Arantes (2006), Machado and Keen (1999), and Machado and Pata (2005).

Stimulus-response generalization tests

The main purpose of this test was to determine whether and how the psychometric functions obtained with the {Red, Yellow} and {Green, Blue} choice sets varied with sample duration and sample cue. Fig. 3 shows the individual results for the first test block (test trials in extinction). Each panel shows two psychometric functions, one corresponding to the preference for Red and the other to the preference for Green. For the panels on the left, the sample cue was the bar associated with the "Short set" trials; for the panels on the right, it was the bar associated with the "Long set" trials.

When the pigeons chose between Red and Yellow, the preference for Red decreased with sample duration. On the basis of visual inspection, the sample cue seemed to have no major effect on choice. A between-within ANOVA assessing the effects of sample duration (within-subjects factor with five levels) and sample cue (between-subjects factor with two levels) revealed a significant effect of sample duration [$F(4,28) = 43.3$, $p < .001$], but not of sample cue [$F(1,7) = 0.46$] or of the interaction between the two factors [$F(4,28) = 0.64$].

When the pigeons chose between Green and Blue, the preference for Green increased with sample duration. Again, visual inspection suggested no major effect of sample cue. The ANOVA revealed a significant effect of sample duration [$F(4,28) = 19.1$, $p < .001$], but not of sample cue [$F(1,7) = 2.04$, $p = .20$] or of their interaction [$F(4,28) = 0.82$]. Despite the low number of trials (10) used to estimate each point of the psychometric functions, the individual data show clearly the influence of sample duration on choice performance.

The analysis of the individual data from the remaining three blocks yielded similar results, which are summarized in Fig. 4. The left panels show the average preference for Red when the choice was between Red and Yellow in blocks 1 (top) to 4 (bottom). The empty and filled circles show the results with the bars associated with the "Short set" and "Long set" trials, respectively. In all blocks and for both sample cues, preference for Red decreased with sample duration. The relative position of the curves for the two sample cues did not change consistently across the blocks. The between-within ANOVAs

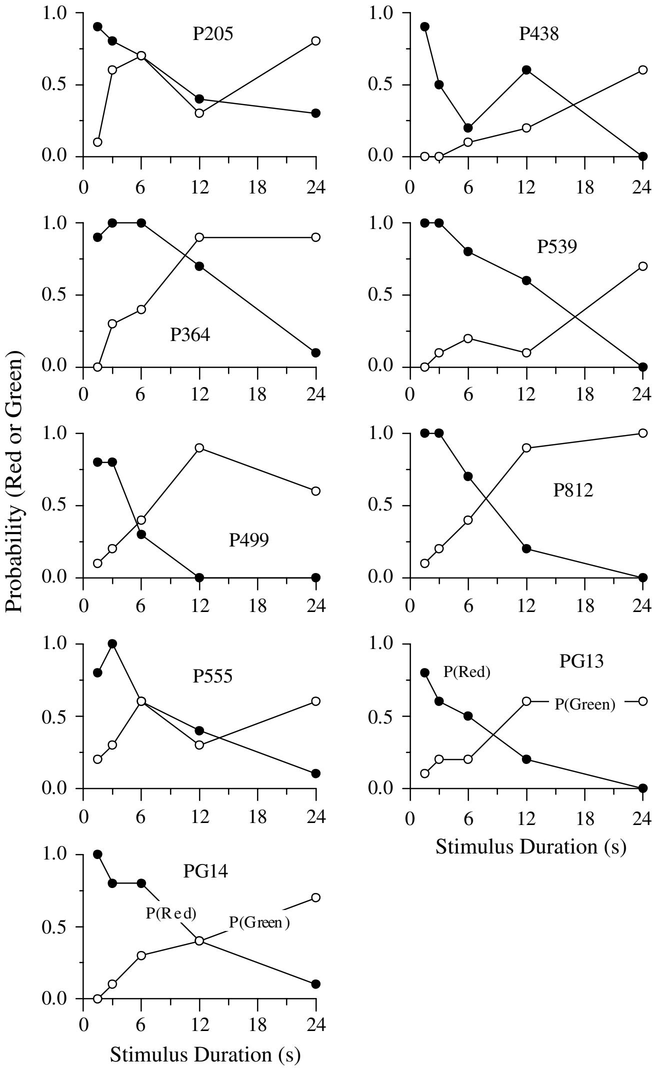

Fig. 3. Individual results from the first block of the Stimulus-Response generalization test phase, $\{S_{1.5}..S_{24}\}$:{Red, Yellow} and $\{S_{1.5}..S_{24}\}$:{Green, Blue}. Proportion of Red (filled circles) and Green (empty circles) responses are plotted against stimulus duration. The left panels show the data from the pigeons tested with the bar associated with the "Short set" trials, and the right panels show the data from the pigeons tested with the bar associated with the "Long set" trials.

yielded always a strong and significant effect of sample duration, but no effect of sample cue; the interaction was significant only during Block 2 [$F(4,28) = 4.06$, $p = .01$].

The right panels show the average preference for Green when the choice was between Green and Blue. In all blocks and for both sample cues, preference for Green increased with sample duration. The curve for the "Short set" bar tended to be close to or above the curve for the "Long set" bar. The ANOVAs yielded always a strong and significant

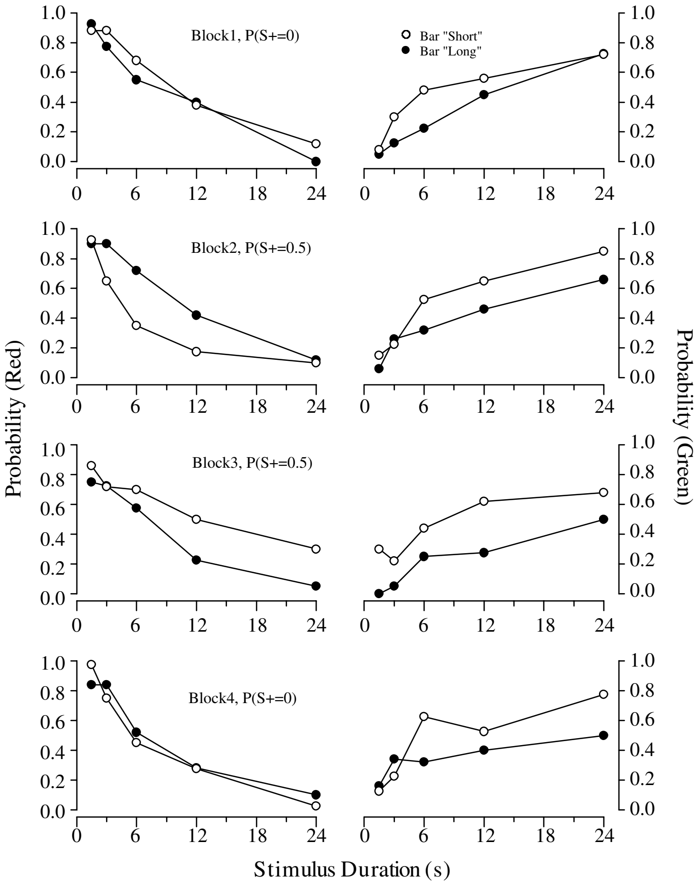

Fig. 4. Average data for each of the four blocks of the Stimulus-Response generalization test phase, $\{S_{1.5}..S_{24}\}$:{Red, Yellow} (left panels), and $\{S_{1.5}..S_{24}\}$:{Green, Blue} (right panels). $P(S^{+})$ refers to the probability of reinforcement after each choice. The two curves on each panel show the data obtained with the bar associated with the "Short set" and "Long set" trials.

effect of sample duration. The effect of sample cue was significant only during Block 3 [$F(1,7) = 7.69$, $p = .03$] and the interaction effect never reached statistical significance.

The preceding findings reveal that the sample duration affects choice markedly and reliably both when the choice set includes the two "end" elements, Red and Yellow, and when it includes the two "middle" elements, Green and Blue. They also show that the sample cue has no reliable effect when the choice set is {Red, Yellow}, but it may have a small effect when the choice set is {Green, Blue}.

To clarify the effect of the sample cue, we performed a more sensitive statistical test, a repeated measures ANOVA with sample duration and sample cue as factors, and the average preference for Red or Green as the dependent variable. In other words, instead of assessing the effect of the sample cue between subjects, as we had done before, we assessed it within subjects. For each pigeon, we aggregated the data from the two blocks with the "Short set" bar and then compared it with the aggregated data from the two blocks with

the "Long set" bar. The results revealed a strong effect of sample duration [Red: $F(4, 32) = 97.0$, $p < .001$; Green: $F(4, 32) = 32.8$, $p < .001$], no effect of sample cue when the choice was from Red and Yellow [$F(1, 8) = 0.07$], but an effect when the choice was between Green and Blue [$F(1, 8) = 13.1$, $p = .007$], and no effect of the interaction between duration and cue on either choice set. These results are summarized in Fig. 5. The top curves show the strong effect of sample duration on the preference for Red and the bottom curves show the strong effect of sample duration and the weaker effect of the sample cue on the preference for Green.

Stimulus generalization tests

The main purpose of this test was to determine whether the psychometric functions obtained with two sets of durations with the same ratio would superpose when plotted in relative time. Fig. 6 shows the individual data and the group mean (bottom right panel). The filled and empty circles correspond to the preference for Red given the choice set {Red, Green} and the preference for Blue given the choice set {Blue, Yellow}, respectively. To plot the two data sets on the same scale, all stimulus durations from the range $\{S_6..S_{24}\}$ were divided by 4. The curves through the data points show the best-fitting, two-parameter logistic functions with equation $P(t) = 1/(1 + \exp(\lambda(t - \mu)))$, where μ is the point of subjective equality (PSE) and λ is the slope of the function at the point $t = \mu$.

Although there was substantial variability both within and between subjects, there was no clear tendency for one curve to be steeper than the other (consistent differences in the slope parameter), or for the curves to be in different locations (consistent differences in the PSE parameter). Two t-tests for related samples comparing the slopes and the locations of the two curves yielded non-significant differences [slope: $t(8) = 1.16$, $p = .28$; PSE: $t(8) = 0.01$, $p = .98$]. We conclude that the two psychometric functions did not differ systematically.

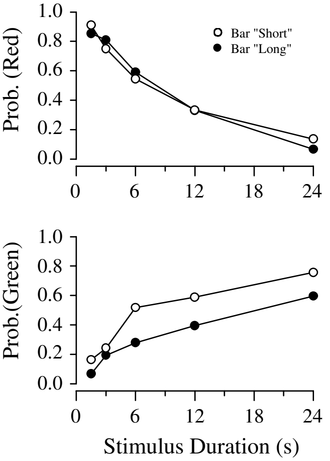

Fig. 5. Average data from all Stimulus-Response generalization test trials, $\{S_{1.5}..S_{24}\}$:{Red, Yellow} (top panel), and $\{S_{1.5}..S_{24}\}$:{Green, Blue} (bottom panel). The two curves on each panel show the data obtained with the bar associated with the "Short set" and "Long set" trials.

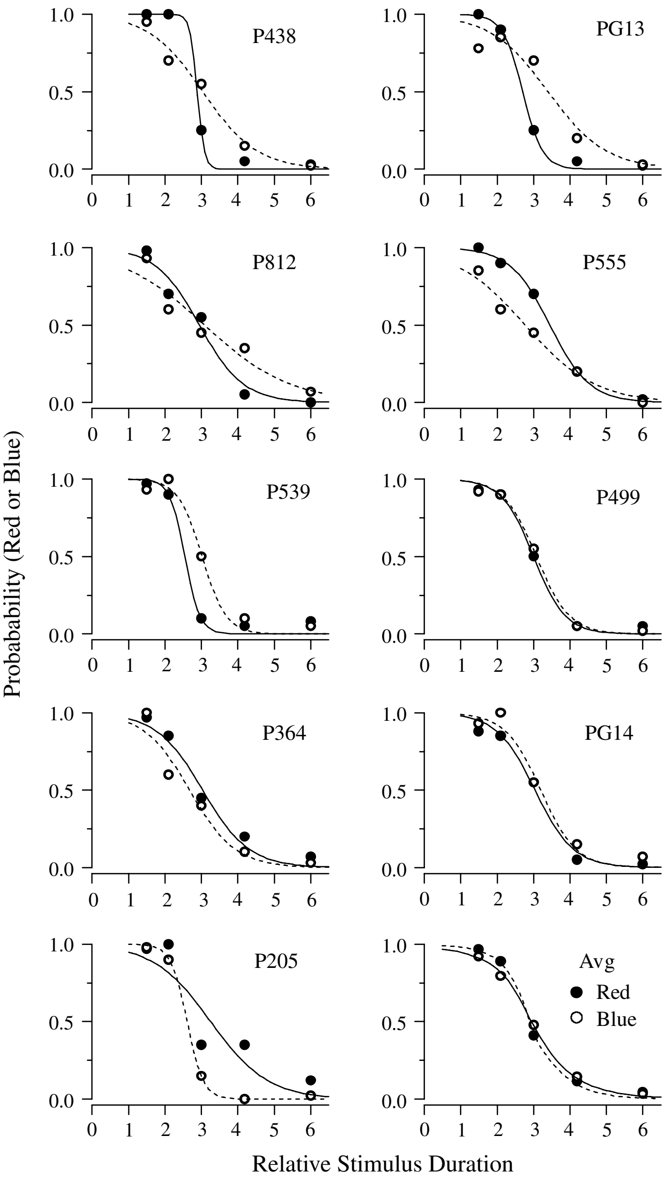

Fig. 6. The symbols show the individual results from the stimulus generalization test phase, $\{S_{1.5}..S_6\}$:{Red, Green} and $\{S_6..S_{24}\}$:{Blue, Yellow}, the latter rescaled. The lines are the best-fitting two parameter logistic functions, with equation $y(t) = 1/(1 + \exp(-\lambda\ (t - \mu)))$. The bottom right panel shows the average data and the average of the individual fits.

Discussion

Both the SET and LeT models of timing account equally well for the typical results obtained with the simple bisection procedure (e.g., $\{S_1, S_4\} \rightarrow$ {Red, Green}). For exam-

ple, both predict that the psychometric function plotting preference for Red against sample duration decreases monotonically, has the PSE close to the geometric mean of the two sample durations, and is a scale transform of other psychometric functions obtained with different durations in the same ratio (see Machado, 1997; for the properties of the psychometric function see, e.g., Catania, 1970; Church & Deluty, 1977; Platt & Davis, 1983; Stubbs, 1968). Therefore, to contrast the two models we needed to design a new procedure, more specifically a procedure that exploited the models' different conceptualizations of the learning process. The double bisection procedure satisfies the requirement. Pigeons learn two simple discriminations $\{S_{1.5}, S_6\} \rightarrow$ {Red, Green} and $\{S_6, S_{24}\} \rightarrow$ {Blue, Yellow} and then, in the critical test, they are exposed to stimulus-response generalization tests in which sample durations ranging from 1.5 to 24 s are followed by a choice between the Green and Blue keys, the keys associated with the same 6-s duration. Whereas LeT predicts that preference for Green should increase with sample duration, SET predicts that preference for Green should not vary with sample duration. The results of previous studies supported LeT (Machado & Arantes, 2006; Machado & Keen, 1999; Machado & Pata, 2005).

The present study examined the generality of this critical effect, that is, the conditions under which preference for Green increases with sample duration. Could the effect be limited to training procedures in which the sample stimulus remained the same, thus preventing the animal from "anticipating" the sample's relative duration and the choice alternatives that follow the sample? In the affirmative, signaling the two types of trials with distinct cues, say, the "Short set" trials $\{S_{1.5}, S_6\} \rightarrow$ {Red, Green} with a horizontal bar and the "Long set" trials $\{S_6, S_{24}\} \rightarrow$ {Blue, Yellow} with a vertical bar, would eliminate the critical effect. To answer the question, nine pigeons learned the two basic temporal discriminations and then were exposed to the test trials $\{S_{1.5}..S_{24}\}$:{Green, Blue} and $\{S_{1.5}..S_{24}\}$:{Red, Yellow} in the presence of either the horizontal or the vertical bars. In the last phase of the experiment, the pigeons were exposed also to stimulus generalization tests.

The results showed that the critical effect was maintained—preference for Green increased monotonically with sample duration. The effect was observed in all pigeons, for both sample cues, and regardless of whether the test trials were conducted in extinction or under non-differential reinforcement (Figs. 3 and 4). On the test trials involving the {Red, Yellow} choice set, preference for Red decreased with sample duration and the PSE was slightly above 6 s, the geometric mean of 1.5 s and 24 s. This result also was observed with all pigeons, with both sample cues, and regardless of the reinforcement contingency. Together, the two sets of results reproduce the results obtained in previous studies (Machado & Arantes, 2006; Machado & Keen, 1999; Machado & Pata, 2005) and extend their generality. In addition, they provide further support for the LeT model and cast strong doubts on SET's conceptualization of timing in the bisection task.

The effect of the sample cue was relatively weak on the critical tests with Green and Blue, and non-existent on the tests with Red and Yellow. The weak effect was that the preference for Green was generally greater when the sample cue was the bar associated with the "Short set" trials than when it was the bar associated with the "Long set" trials (see Fig. 4, right panels, and Fig. 5, bottom panel). Moreover, the sample cue did not seem to interact with the sample duration.

To better appreciate the effects of the sample cue and duration, Fig. 7 plots the average results of the present study together with the average results from two previous studies,

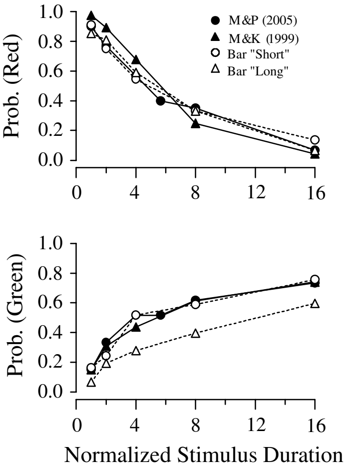

Fig. 7. Average data from the Stimulus-Response generalization test phase, {$S_{1.5}$..S_{24}}:{Red, Yellow} (top) and {$S_{1.5}$..S_{24}}:{Green, Blue} (bottom), both data sets rescaled, and the corresponding data from Machado and Keen (1999) and Machado and Pata (2005), {S_1..S_{16}}:{Red, Yellow} (top) and {S_1..S_{16}}:{Green, Blue} (bottom). In the present study there were two sample cues, the bar associated with the "Short set" trials and the bar associated with the "Long set" trials; in the other studies there was only one sample cue.

Machado and Keen (1999) and Machado and Pata (2005). These previous studies used only one sample cue (a white keylight) and the sample durations were 1 and 4 s on the "Short set" trials, and 4 and 16 s on the "Long set" trials. The top panel shows the results with the {Red, Yellow} choice set. The four curves overlapped considerably, which illustrates the absence of the sample cue effect when the choice was between Red and Yellow. The bottom panel shows the results with the {Green, Blue} set. Three curves overlapped considerably and were clearly above the curve obtained with the "Long set" bar. That is, the psychometric function obtained with the bar associated during training with the Green choice ("Short set" or horizontal bar) was similar to the psychometric functions obtained in previous studies using only one sample cue, but the psychometric function obtained with the bar associated during training with the Blue choice ("Long set" or vertical bar) lowered the preference for Green. We return to this finding below.

On the stimulus generalization tests, the results with two distinct sample cues did not reproduce the results obtained by Machado and Keen (1999) with one sample cue. In that study, the psychometric functions {S_1..S_4}:{Red, Green} and {S_4..S_{16}}:{Blue, Yellow} did not superpose for any of the eight pigeons. The fits of the two parameter logistic function revealed that for six pigeons the function for the "Long set" trials was steeper than the function for the "Short set" trials; for the other two birds, the opposite was the case. Although not reported in the original article, a *t* test for related samples showed that the mean slope of the curve for the "Long set" generalization trials was greater than the mean slope of the curve for the "Short set" generalization trials; the PSEs did not differ significantly. In contrast, in the present study, some pigeons produced a steeper curve during the "Short set" generalization trials, other pigeons produced a steeper curve during the "Long set" generalization trials, and still other pigeons produced two curves with identical slopes (see Fig. 6). The average results for all nine subjects show that neither the slopes nor the PSEs differed significantly between the two curves.

Fig. 8 summarizes the difference between the two studies in the stimulus generalization test trials. The symbols show the average of the individual data, and the curves show the average of the individual fits. Whereas in Machado and Keen (1999) the average function for the "Long set" trials is clearly steeper than the average function for the "Short set" trials, in the present study the two functions overlap significantly.

The results from the present experiment have some implications for the three hypotheses mentioned in the Introduction regarding the effect of the sample cue: the hypothesis of changes in pacemaker speed (SET), or of its equivalent in LeT, the rate of transition across behavioral states; the hypothesis of changes in mediating behaviors; and the hypothesis of differential associative connections between the sample cues and the choice alternatives. In what follows, we consider each hypothesis in the light of the preceding stimulus generalization and stimulus-response generalization findings.

The first hypothesis stated that the two sample cues could set the occasion for different mean speeds of the pacemaker, or different mean rates of transition across the behavioral states. Whether this hypothesis is consistent with the data depends on the details of its instantiation, which in turn depend on the model. According to SET, provided the coefficient of variation of pacemaker speed remains constant, changes in mean speed will still yield superposition of the two stimulus generalization curves. Thus SET has no difficulty accounting for the results displayed in Fig. 6. However, it has difficulties explaining the lack of superposition of the psychometric functions when only one sample cue is used, as in Machado and Keen (1999).

According to LeT, superposition requires that the rate of transition across the behavioral states be proportional to the rate of reinforcement. In the present study, this means that superposition is predicted only if the rate of transition in the presence of the horizontal bar is four times greater than the rate of transition in the presence of the vertical bar. Previous studies support changes in rate of transition with changes in reinforcement rate, but they do not support strict proportionality (i.e., a fourfold increase in reinforcement rate yields a less than fourfold increase in rate of transition; Bizo & White, 1994; Fetterman & Killeen, 1991; Killeen & Fetterman, 1988). In the absence of strict proportionality, LeT predicts a curve for "Long set" trials steeper than the curve for "Short set" trials, as observed in Machado and Keen (1999) study, but not in the present experiment.

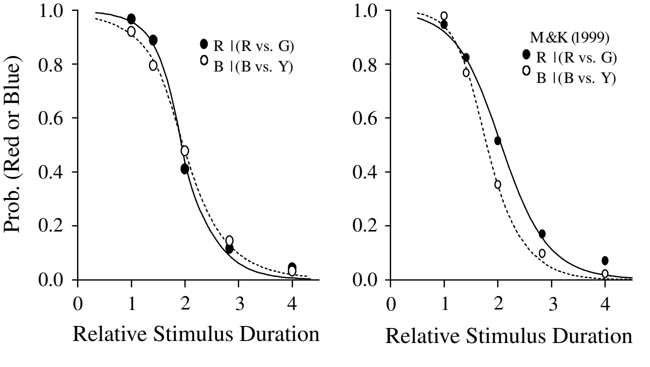

Fig. 8. Average (rescaled) data from the stimulus generalization test phase of the present study, $\{S_1..S_6\}$:{Red, Green} and $\{S_6..S_{24}\}$:{Blue, Yellow} (left panel), and from Machado and Keen (1999), $\{S_1..S_4\}$:{Red, Green} and $\{S_4..S_{16}\}$:{Blue, Yellow} (right panel).

Thus, in contrast with SET, LeT has no difficulties explaining a steeper curve on the "Long set" trials, but has difficulties explaining superposition.

To find out whether the hypothesis that pacemaker mean speed varies with the sample cue can account for the stimulus response generalization data, we derive specific predictions for SET, but a similar reasoning shows that they hold also for LeT. Fig. 9 illustrates the details for the double bisection task used during the experiment. Assume that mean pacemaker speed is higher in the presence of the horizontal than the vertical bar. In the top panel of the figure, these mean speeds were arbitrarily set at $\lambda_H = 1.0$ and $\lambda_V = 0.5$, respectively. Assume also that the coefficient of variation of pacemaker speed remains constant at 0.3. Then the distributions of the number of pulses saved in the four memory stores will be Gaussian with the following means and standard deviations: M_{Red} (1.5, 0.45), M_{Green} (6.0, 1.8), M_{Blue} (3.0, 0.9), and M_{Yellow} (12.0, 3.6). To predict choice between Green and Blue, note that the distribution in the M_{Green} store is located to the right of the distribution in the M_{Blue} store. It follows that, according to SET's decision rule, at the end of a stimulus with duration T, Green will be chosen provided the value in the accumulator, X_T , is greater than $\sqrt{(X_B * X_G)}$, where X_B and X_G are samples extracted from the two memory distributions. In other words, the probability of choosing Green is given by

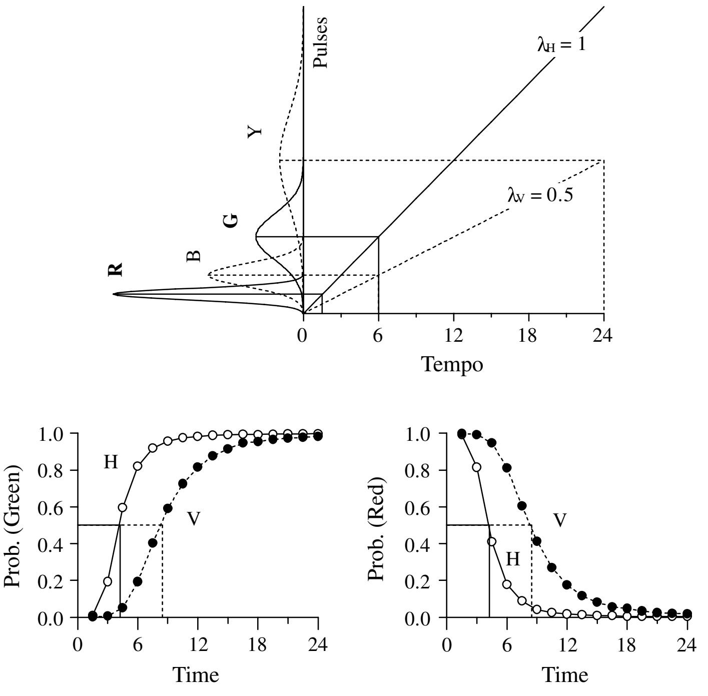

Fig. 9. Predictions of SET for the double bisection task assuming differences in pacemaker mean speed, λ, with reinforcement context (horizontal, H, and vertical, V, bar). The coefficient of variation of pacemaker speed equaled 0.3. (Top) The distributions of the number of pulses on the left represent the contents of the M_{Red}, M_{Green}, M_{Blue} and M_{Yellow} memory stores at the steady state. (Bottom left) Preference for Green over Blue increases with sample duration, but twice as fast in the presence of the H bar than the V bar. The curves are scale transforms, $y_H(x) = y_V(2x)$. (Bottom right) Preference for Red over Yellow decreases with sample duration, but twice as fast in the presence of the H bar than the V bar.

$$P(\text{Green}|T) = P\left(X_T > \sqrt{X_B X_G}\right).$$

But for all sample durations T, the inequality $X_T > \sqrt{(X_B * X_G)}$ is more likely to hold in the presence of the faster pacemaker, that is, of the horizontal bar. Therefore, SET predicts that preference for Green increases with sample duration, but faster in the presence of the horizontal bar.

The bottom left panel of Fig. 9 shows the bar effect. The PSE is close to 4.2 s in the presence of the horizontal bar and 8.4 s in the presence of the vertical bar. In fact, the psychometric functions obtained with the two bars are scale transforms: If the abscissas for the V curve are divided by two, the ratio of the pacemaker mean speeds, then that curve will superpose the H curve. The bottom right panel shows the functions predicted given a choice between Red and Yellow. Preference for Red decreases with sample duration but faster in the presence of the horizontal bar. The two curves are obtained from the curves on the left panel by subtracting the values of the latter from 1. Hence the PSEs are the same, and the two curves also are scale transforms.

The results from the stimulus-response generalization tests seem inconsistent with the hypothesis of changes in pacemaker speed for two reasons. First, the cue effect was visible when the choice involved the Green and Blue keys, but not when it involved the Red and Yellow keys. This asymmetry is at odds with the hypothesis. Second, the relatively weak effect obtained with the {Green, Blue} set seems more consistent with a shift along the vertical axis (an additive effect) than with a scale transformation along the horizontal axis (compare the bottom panels of Figs. 5 and 9).

A second hypothesis to account for the sample cue effects referred to collateral behaviors, which may mediate choice in temporal discrimination tasks (e.g., Fetterman et al., 1998; Killeen & Fetterman, 1988; Richelle & Lejeune, 1980). Collateral behaviors also occur in the double bisection task. In Machado and Keen (1999) study, during the longest, 16-s samples, most pigeons exhibited a chain of three behaviors, say, A followed by B followed by C. The transitions between these behaviors correlated with changes in the pigeon's choices. Thus, if the opportunity to choose occurred while the pigeon was engaged in A, then it tended to choose Red; if it occurred while it was engaged in B, then it tended to choose Green or Blue; and if it occurred while it was engaged in C, it tended to choose Yellow (see also Machado & Pata, 2005). If this serial organization of collateral behaviors is altered when one sample cue is replaced by two distinct sample cues, then choice could also change. To illustrate, suppose a pigeon acquires the behavioral chain 'A followed by B' during the "Short set" trials with the horizontal bar, perhaps switching from A to B at 3 s into the trial, and the chain 'C followed by D' during the "Long set" trials with the vertical bar, switching from C to D at 12 s into the trial. Then, its performance on the test trials would depend on the sample cue. In the presence of the horizontal bar the bird would perform behavior A until 3 s into the trial. Given a choice between Green and Blue while performing A, the bird would tend to avoid Green because during training choices of Green while performing A would have been extinguished. However, after 3 s the pigeon would switch from A to B, and given a choice between Green and Blue while engaging in B, it would prefer Green because during training choices of Green would have been reinforced while engaging in B. In summary, after samples shorter than 3 s the bird would prefer Blue and after samples longer than

3 s the bird would prefer Green. However, in the presence of the vertical bar the pigeon would perform behavior C until 12 s into the trial (and while engaging in C it would prefer Blue) and then it would switch to D (and then prefer Green because behavior C would have been associated with extinction of Blue choices). The function obtained with the vertical bar would be shifted to the right of the function obtained with the horizontal bar. Similar reasoning shows that the psychometric function obtained with the {Red, Yellow} choice set should shift to the right when the bar changed from horizontal to vertical.

Although it was not possible to observe and record the pigeons' behaviors in this experiment, the present findings seem inconsistent with the hypothesis of changes in collateral behavior. The reasons are the same as above, namely, the hypothesis cannot account for the fact that the sample cue had an effect when the choice set was {Green, Blue} but not when it was {Red, Yellow}, and the fact that the effect obtained with the {Green, Blue} set is more consistent with a displacement along the vertical axis than with a change in scale along the horizontal axis.

A third hypothesis stated that the sample cue effect could be due to the associations learned during training between the bars and the keylight colors. The explanation is similar in kind to LeT's account of the main sample duration effect: Preference for Red over Yellow decreases with sample duration because during training choices of Red are *reinforced* after the shortest samples and choices of Yellow are *reinforced* after the longest samples; preference for Green over Blue increases with sample duration because during training choices of Green are *extinguished* after the shortest samples and choice of Blue are *extinguished* after the longest samples. The first account stresses the effect of reinforcement, the second the effect of extinction. Add to this account two reasonable assumptions: First, each bar will bias the animal towards the color reinforced in the presence of that bar (e.g., because in the presence of the horizontal bar the choice of Green is reinforced but that of Blue is not, the animal will be biased towards Green). Second, as is well known (see, e.g., Roberts, 1998), the stimulus control function of a sample cue increases with its duration. Hence, all rest equal, the vertical bar will exert a significantly greater biasing effect on choice than the horizontal bar because during training it lasts four times longer. From the two assumptions some consequences follow. Because the horizontal bar has small effects given its relatively short duration, the curves obtained with it should be similar to the reference curves obtained with only one sample cue. Fig. 7 shows this was the case. Because the vertical bar has larger effects, the curve obtained with it when the choice is between Green and Blue should be below the {Green, Blue} reference curve. Again, Fig. 7 shows this was the case. However, the hypothesis does not explain why the curve obtained with the vertical bar when the choice was from Red and Yellow is not below the {Red, Yellow} reference curve. None of the three hypotheses is consistent with the full pattern of results.

In conclusion, the present set of findings show that when a sample cue announces, as it were, the sample's relative duration and the set of choice alternatives that will follow it, the critical effect obtained in the double bisection task is maintained: Preference for Green over Blue increases with sample duration. This finding is predicted by LeT, but it is inconsistent with SET, the dominant model in the field. However, neither model can account for the full pattern of results (stimulus generalization and stimulus-response generalization) obtained with the double bisection procedure.

References

Beam, J. J., Killeen, P., Bizo, L., & Fetterman, J. G. (1998). How reinforcement context affects temporal production and categorization. *Animal Learning & Behavior, 26*, 388–396.

Bizo, L. A., & White, K. G. (1994). The behavioral theory of timing: Reinforcer rate determines pacemaker rate. *Journal of the Experimental Analysis of Behavior, 61*, 19–33.

Bizo, L. A., & White, K. G. (1995a). Biasing the pacemaker in the behavioral theory of timing. *Journal of the Experimental Analysis of Behavior, 64*, 225–235.

Bizo, L. A., & White, K. G. (1995b). Reinforcement context and pacemaker rate in the behavioral theory of timing. *Animal Learning & Behavior, 23*, 376–382.

Catania, A. C. (1970). Reinforcement schedules and the psychophysical judgments: A study of some temporal properties of behavior. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 1–42). New York: Appleton-Century-Crofts.

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavioural Processes, 3*, 216–228.

Church, R. M., & Gibbon, J. (1982). Temporal generalization. *Journal of Experimental Psychology: Animal Behavior Processes, 8*, 165–186.

Dews, P. (1970). The theory of fixed-interval responding. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 43–61). New York: Appleton-Century-Crofts.

Fetterman, J. G., & Killeen, P. R. (1991). Adjusting the pacemaker. *Learning and Motivation, 22*, 226–252.

Fetterman, J. G., Killeen, P. R., & Hall, S. (1998). Watching the clock. *Behavioural Processes, 44*, 211–224.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*, 279–325.

Gibbon, J. (1981). On the form and location of the psychometric bisection function for time. *Journal of Mathematical Psychology, 24*, 58–87.

Gibbon, J. (1991). Origins of scalar timing theory. *Learning and Motivation, 22*, 3–38.

Killeen, P., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review, 95*, 274–285.

Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological Review, 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. *Behavioural Processes, 72*, 195–206.

Machado, A., & Cevik, M. (1998). Acquisition and extinction under periodic reinforcement. *Behavioural Processes, 44*, 237–262.

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. *Learning and Behavior, 33*, 111–122.

Morgan, L., Killeen, P. R., & Fetterman, J. G. (1993). Changing rates of reinforcement perturbs the flow of time. *Behavioural Processes, 30*, 259–272.

Platt, J. R., & Davis, E. R. (1983). Bisection of temporal intervals by pigeons. *Journal of Experimental Psychology: Animal Behavioural Processes, 9*, 160–170.

Richelle, M., & Lejeune, H. (1980). *Time in Animal Behavior*. Oxford, UK: Pergamon Press.

Roberts, W. A. (1998). *Principles of animal cognition*. New York: McGraw-Hill.

Shettleworth, S. J. (1998). *Cognition, evolution, and behavior*. NY: Oxford University Press.

Staddon, J. E. R. (1977). Schedule-induced behavior. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 125–152). Englewood Cliffs, NJ: Prentice-Hall.

Staddon, I. E. R., & Simmelhag, V. L. (1971). The "superstition" experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review, 78*, 3–43.

Stubbs, D. A. (1968). The discrimination of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior, 11*, 223–238.

Timberlake, W., & Lucas, G. A. (1985). The basis of superstitious behavior: Response contingency, stimulus substitution, or appetitive behavior? *Journal of the Experimental Analysis of Behavior, 44*, 279–299.