



On the content of learning in interval timing: Representations or associations?



Jeremie Jozefowicz^{a,b,*}, Armando Machado^b

^a Laboratoire URÉCA, Université Lille Nord de France, Campus de Lille3, Villeneuve d'Ascq Cedex, France

^b Universidade do Minho, Braga, Portugal

ARTICLE INFO

Article history:

Received 5 December 2012

Received in revised form 15 February 2013

Accepted 15 February 2013

Keywords:

Interval timing

Linear timing

Log timing

Cognitive models

Associative models

ABSTRACT

Models of timing differ on two fundamental issues, the form of the representation and the content of learning. First, regarding the representation of time some models assume a linear encoding, others a logarithmic encoding. Second, regarding the content of learning cognitive models assume that the animal learns explicit representations of the intervals relevant to the task and that their behavior is based on a comparison of those representations, whereas associative models assume that the animal learns associations between its representations of time and responding, which then drive performance. In this paper, we show that some key empirical findings (timescale invariant psychometric curves, bisection point at the geometric mean of the trained durations in the bisection procedure, and location of the indifference point in the time-left procedure) seem to make these two issues interdependent. That is, cognitive models seem to entail a linear representation of time, and at least a certain class of associative models seem to entail a log representation of time. These interdependencies suggest new ways to compare and contrast timing models.

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

© 2013 Elsevier B.V. All rights reserved.

Interval timing is the ability of organisms to perceive intervals between two events ranging from seconds to minutes. In recent years, many computational models have been proposed to account for interval timing, including the Scalar Expectancy Theory (SET; Gibbon et al., 1984), the Behavioral Theory of Timing (BeT; Killeen and Fetterman, 1988), the Learning-to-Time model (LeT; Machado, 1997; Machado et al., 2011), the Multiple Timescale Model (MTS; Staddon and Higa, 1999), the packet theory of timing (Guilhardi et al., 2007; Kirkpatrick, 2002), and the Behavioral Economic Model (BEM; Jozefowicz et al., 2009a). To compare and contrast the models, we examine how each model answers two fundamental sets of questions about time perception: (a) How do animals perceive time? A corollary to this question is, What are the formal properties of the representation of time? and (b) What is the content of learning in an interval timing task? In other words, What exactly do animals learn in such tasks? The answer to the second set of questions includes the decision rule in timing, or how animals decide when to respond.

As far as the first set of questions is concerned, various answers have been proposed but the most common ones assume either a linear representation of time (subjective time is a linear function of real time), or a logarithmic representation of time (subjective time is a logarithmic function of real time). For instance, SET assumes that the representation of time is based on a pacemaker whose pulses accumulate in short-term memory (STM). The most recent version of LeT (Machado et al., 2009) assumes that the onset of the time-marker activates a set of states serially. The system stays in each state a fixed amount of time before moving to the next state. Though different on the surface, both models entail to linear representations of time with similar formal properties. More specifically, the representation of an interval t is a Gaussian random variable with mean at and standard deviation proportional to the mean. Henceforth, the term “linear timing” will cover these three assumptions. By contrast, models such as MTS and BEM both assume a logarithmic representation of time. MTS hypothesizes that the perception of time is grounded in the decaying STM trace of the time-marker, which induces a logarithmic-like function between real time and subjective time BEM assumes a series of time-dependent states similar to LeT, but the activation of the states differs such that the encoding of time is logarithmic instead of linear. In both of these models (MTS and BEM), the representation of an interval t is a Gaussian random variable with mean $\ln(t)$ and constant standard deviation. Henceforth, the term “log timing” will cover these three assumptions.

* Corresponding author at: Laboratoire URÉCA, Université Lille Nord de France, Campus de Lille3, Domaine Universitaire du Pont de Bois, B.P. 60149, 59653 Villeneuve d'Ascq Cedex, France. Tel.: +33 320 041 6866; fax: +33 320 041 6036.

E-mail address: jeremie.jozefowicz@univ-lille3.fr (J. Jozefowicz).

Regarding the content of learning, models fit in one of two categories, cognitive or associative. Cognitive models assume that the animal forms representations of the various relevant intervals, and decides to respond on the basis of an explicit comparison involving these representations (what we will call a cognitive decision rule). Associative models assume that the animal learns associations between its representations of time intervals and responses or stimuli. The strength of the associations change according to the general principles of associative learning as embedded in models such as Rescorla and Wagner (1972) or Bush and Mosteller (1955). Responding occurs when representations strongly associated with responding become active (what we will call an associative decision rule). Included in the cognitive category are models like SET and MTS; included in the associative category are models like LeT and BEM.

One could assume that the answers to the representation and learning/decision rule questions are independent. In this case, one could take the representation of time proposed by SET, for instance, and combine it with the associative learning assumptions proposed by the original LeT model to create a new hybrid model. This is exactly what Machado et al. (2009) did to create the new version of LeT. In this article, we argue that the two sets of questions may be intertwined: Due to strong empirical constraints, the answer to the first set of questions (How is time represented?) may constrain the answer to the second set of questions (What is the content of learning?) and reciprocally, the answer to the content of learning question may constrain the answer to the representation question. More precisely, a cognitive model may imply a linear representation of time while a certain class of associative models may imply a log representation of time. Our argument starts with one of the most common procedures for the study of timing in humans and animals, the bisection paradigm.

1. The bisection procedure

In the bisection procedure (Stubbs, 1968; Church and Deluty, 1977), the subject is reinforced for emitting response R1 following a stimulus lasting S sec, and response R2 following a stimulus lasting L sec ($S < L$). Then, on unreinforced probe trials, it is presented with stimuli ranging in duration from S to L . Fig. 1 shows the typical sigmoid psychometric curves obtained in this procedure. They display two fundamental quantitative properties that all timing models have to account for (Church, 2004): (a) the point at which the subject is indifferent between R1 and R2, the

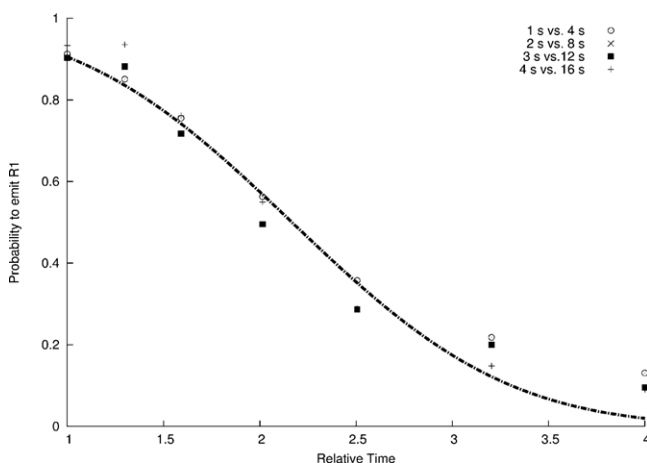


Fig. 1. Probability to emit R1, the response reinforced after the longest stimulus duration, as a function of the relative stimulus duration in the Church and Deluty (1977)'s study on rats.

Redrawn from Church and Deluty (1977).

bisection point, is located, at least for non-human animals, at or close to the geometric mean of S and L ; and (b) when plotted in relative time, the psychometric curves superimpose (timescale invariance).

How would a cognitive model explain the animal's performance in a bisection task? There are three important intervals in the task: t , the duration of the current stimulus, S , the duration of the stimulus associated with response R1, and L , the duration of the stimulus associated with response R2. In a cognitive model, the animal will have a short-term memory (STM) representation $f(t)$ of t , and a long-term memory (LTM) representations $X(L)$ and $X(S)$ of L and S , respectively. A comparison involving these three representations will determine whether the animal chooses R1 or R2.

If we assume that $f(t)$, $X(L)$, and $X(S)$ are all Gaussian random distributions, Gibbon (1981) showed that, in such a framework, only two encoding schemes predict timescale invariant psychometric curves with a bisection point at the geometric mean of the trained durations (we refer the readers to Gibbon's, 1981 original report for the mathematical derivation): linear representations with scalar variability coupled with a ratio decision rule, and logarithmic representations with constant variability coupled with a difference decision rule. In the linear scheme, $f(t)$, $X(S)$, and $X(L)$ are all Gaussian random variables with standard deviation proportional to the mean. The mean of each distribution is itself proportional to the interval it represents. This is the representation scheme favored by Gibbon as it is the one predicted by the pacemaker-accumulator mechanism postulated by SET. The subject computes the ratio $f(t)/X(S)$ and compares it with the ratio $X(L)/f(t)$. It emits R1 if the former ratio is smaller than the latter, R2 otherwise. In the logarithmic scheme, $f(t)$, $X(S)$, and $X(L)$ are also Gaussian random variables but their means are equal to the logarithm of the represented interval, and their standard deviations are constant and equal. The subject computes the differences $f(t) - X(S)$ and $X(L) - f(t)$. It emits R1 if the former difference is smaller than the latter, R2 otherwise.

Can similar conclusions be reached for associative models? In other words, which encoding schemes are consistent with timescale invariance and bisection at the geometric mean if an associative model is used? Let's assume that the organism represents only t , the duration of the current stimulus, through a random variable $f(t)$. In an associative model, each value of the representation is differentially associated with either R1 or R2. These associations are updated according to some learning rule. We consider one of the simplest cases, the Bush–Mosteller (1955) linear rule. Let $V_{R1}(x)$ and $V_{R2}(x)$ be the strengths of the associations between the value x of the representation and responses R1 and R2, respectively. When the representation takes the value x , the animal emits R1 with probability $P(R1|x)$ and R2 with probability $P(R2|x) = 1 - P(R1|x)$. When the animal emits R1, $V_{R1}(x)$ changes but $V_{R2}(x)$ does not. When the animal emits R2, $V_{R2}(x)$ changes, but $V_{R1}(x)$ does not. Assume the animal emitted R1 and received a reinforcer. Then, $V_{R1}(x)$ increases by the amount $\Delta V_{R1}(x) = \beta[\lambda_{R1} - V_{R1}(x)]$ where $\beta > 0$ is the learning rate following reinforcement, and λ_{R1} is the reward magnitude when R1 is reinforced. In the usual bisection task, R1 is reinforced if and only if stimulus S was presented. Therefore, the change mentioned above for $V_{R1}(x)$ occurs with probability $P(S|x)$, the probability that the stimulus duration was S -sec long given that the representation took the value x . On the other hand, if the animal emits R1 when the representation has the value x and it is not reinforced, $V_{R1}(x)$ decreases by $\Delta V_{R1}(x) = -\alpha V_{R1}(x)$ where α is the learning rate following non-reinforcement. This change occurs when the L stimulus was presented, which occurs with a probability $1 - P(S|x)$. Combining the effects of reinforcement and extinction yields the expected change in $V_{R1}(x)$, $E[\Delta V_{R1}(x)] = P(R1|x) \times \{P(S|x)\beta[\lambda_{R1} - V_{R1}(x)] - [1 - P(S|x)]\alpha V_{R1}(x)\}$.

At the asymptote of learning, the average value of $V_{R1}(x)$ is stable, which means that $E[\Delta V_{R1}(x)] = 0$. It follows that either

$P(R1|x)=0$ or $V_{R1}(x) = \frac{\beta(S|x)\lambda_{R1}}{(\beta-\alpha)P(S|x)+\alpha}$. If $\alpha = \beta$, the expression simplifies to $V_{R1}(x) = P(S|x)\lambda_{R1}$, which is the average payoff for emitting R1 when the representation has the value x . In the same vein, $V_{R2}(x)$, the strength of the association between x and R2, will converge on $P(L|x)\lambda_{R2}$, the average payoff for emitting R2 when the representation has the value x .

On any given trial, following a stimulus lasting t units of time, the animal emits whichever response is more strongly associated with its current time percept $x=f(t)$. Hence, it emits R1 if $V_{R1}[x] > V_{R2}[x]$, and it emits R2 otherwise. For relatively large values of x , $V_{R1}(x) < V_{R2}(x)$ because these values of x are more likely when the signal is L than when it is S ; for relatively small values of x , $V_{R1}(x) > V_{R2}(x)$ because these values of x are more likely when the signal is S than when it is L . Hence, there is a critical value x_c such that $V_{R1}(x_c) = V_{R2}(x_c)$. If $f(t) < x_c$, the animal emits R1; otherwise it emits R2. Hence, the associative model with a reward-following linear rule is mathematically equivalent to a signal detection model with the criterion set at x_c (Wickens, 2002). It follows that

$$V_{R1}(x) > V_{R2}(x) \Leftrightarrow P(S|x)\lambda_{R1} > P(L|x)\lambda_{R2}$$

Using Bayes' theorem and assuming that $\lambda_{R1} = \lambda_{R2} = \lambda$, we have

$$\frac{P(S)P(x|S)}{P(x)}\lambda > \frac{P(L)P(x|L)}{P(x)}\lambda \quad (1)$$

where $P(S)$ and $P(L)$ are the probabilities of trials with the S and L stimuli, respectively. In a standard bisection procedure, $P(S) = P(L)$. Hence, Eq. (1) becomes

$$\frac{P(S|x)}{P(L|x)} > 1$$

which is the decision rule for a signal detection model using likelihood ratio to decide how to respond (Wickens, 2002): R1 is emitted when one hypothesis about the state of the world ("The stimulus lasted S units of time") is more likely than the other ("The stimulus lasted L units of time") given the information available to the subject (the value of its representation of time).

This type of signal detection model was also studied by Gibbon (1981). He concluded that, if we assume that $f(t)$ is a Gaussian random variable, such models would produce timescale invariant psychometric curves with bisection points at the geometric mean of the training durations only when the representation of time $f(t)$ had mean $\ln(t)$ and constant standard deviation k . Notably, if $f(t)$ was a Gaussian random variable with mean proportional to t and standard deviation proportional to the mean, then the model would not be able to account for the bisection data. Although the psychometric curves would still be timescale invariant, their bisection points would not be located at the geometric mean of the two trained durations, but somewhere closer to the harmonic mean. The actual location would depend on ratio of the two training durations and on the standard deviation of the temporal representation, a result only rarely reported in the literature (see Machado et al., 2009).

Given the mathematical equivalence between likelihood ratio signal detection models and a class of associative models (the class representing t through a Gaussian random variable, using the Bush–Mosteller learning rule, and updating only the associations for the just-emitted response), at least in a simple discrimination task such as the bisection procedure, the conclusions that apply to signal detection models may apply as well to those associative models. This is illustrated in Fig. 2. The curves show the predictions of Jozefowicz et al.'s (2009a) BEM model¹ for a bisection procedure

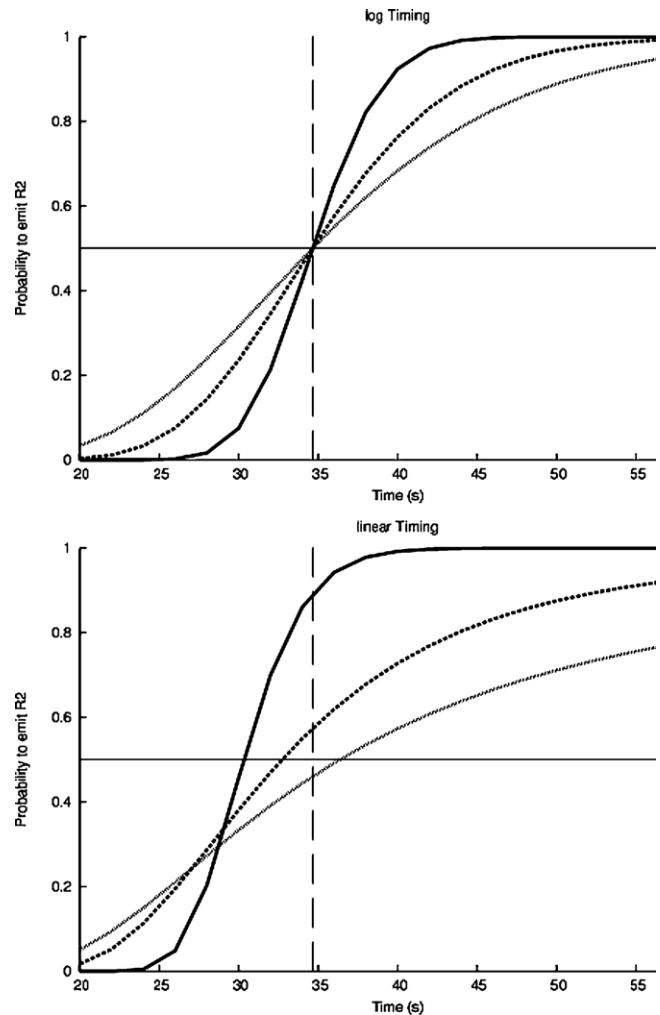


Fig. 2. Probability to emit R2, the response reinforced after a 60-s stimulus, in a simulation of a 20-s vs. 60-s bisection task by Jozefowicz et al.'s (2009a) BEM. The vertical line indicates the geometric mean between 20 and 60 s. Top panel: simulations where the model used a logarithmic encoding of time with constant variance. The difference between the three curves is the level of noise in the representation of time. Bottom panel: simulations when model used a linear encoding of time with standard deviation proportional to the mean. The difference between the three curves is the level of noise in the representation of time.

Redrawn from Jozefowicz et al. (2009a).

with $S = 20$ s and $L = 60$ s stimuli² (details for all the simulations can be found in the original report by Jozefowicz et al., 2009a,b). Usually, the model assumes a logarithmic representation of time with constant variance, a scheme that (correctly) predicts time scale invariant psychometric curves with bisection points at the geometric mean of the training stimuli. This prediction does not depend on the variance (the amount of noise in the temporal representation; see Fig. 2, top panel). In contrast, if the logarithmic representation is replaced by a linear representation with standard deviation proportional to the mean, the bisection point is no longer at the geometric mean and its location varies with the level of noise in the representation of time. Note that, although this is not shown in Fig. 2, the psychometric curves remain timescale invariant, no

¹ BEM is taken here only as an instance of the class of associative models for which we have shown the mathematical equivalence with a likelihood ratio signal detection models.

² A 20-s vs. 60-s discrimination is fairly unusual in a bisection procedure, but those were the values used by Jozefowicz et al. (2009a,b). In any case, because the predictions of BEM are timescale invariant, what matters is the ratio of the durations, not their absolute value. In relative time, the same bisection curves would have been generated for 1-s vs. 3-s or a 2-s vs. 6-s discrimination.

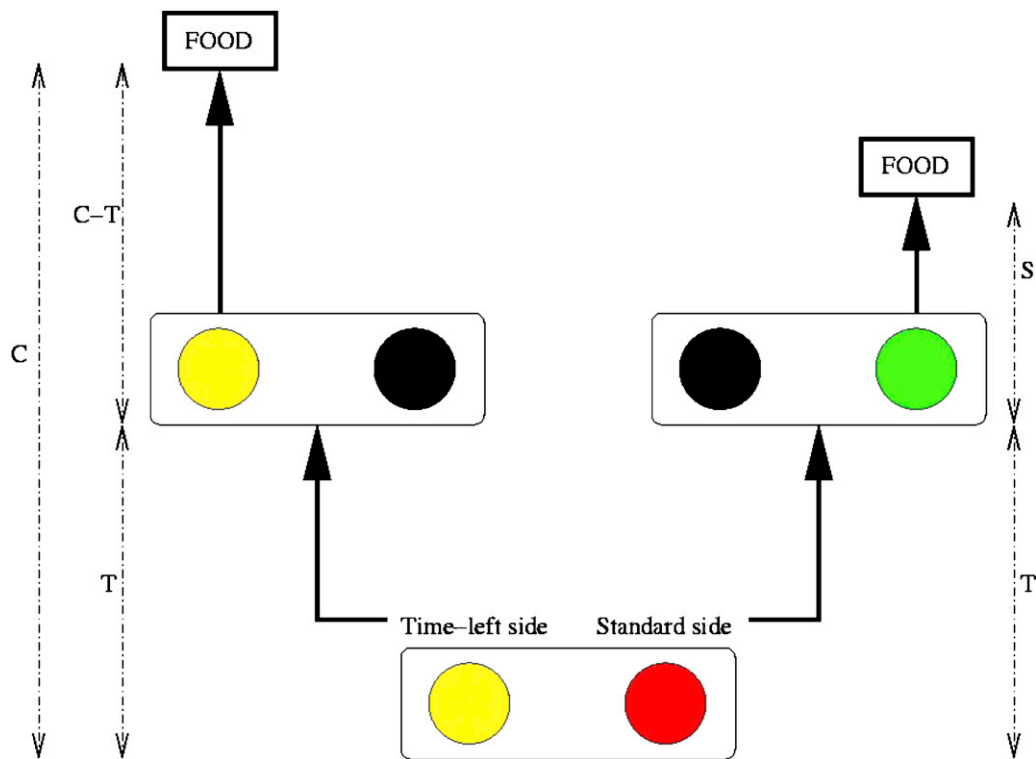


Fig. 3. Schematic representation of the time-left procedure.

matter whether the representation is logarithmic or linear. Similar to the linear version of BEM just described, Machado et al.'s (2009) new LeT model also combines an associative decision rule with a linear representation of time with standard deviation proportional to the mean. Even though it violates one of the assumption we made in order to reduce an associative model to a likelihood ratio signal detection model (only the associations for the response just emitted are updated on a given trial), the new LeT also has problems predicting a bisection point at the geometric mean (instead, it predicts that the bisection point approaches the harmonic mean as the ratio L/S increases),

Hence, while the data from the bisection procedure do not seem to impose any strong constraints on cognitive models on the issue of whether time is represented linearly or logarithmically, the case seems otherwise for the class of associative models that are mathematically equivalent to a likelihood ratio signal detection model in a bisection task. This will happen when the model assumes (a) that time is represented through a random variable, (b) that the associations are updated according to a Bush–Mosteller rule, and (c) that only the associations for the response just reinforced or non-reinforced are updated. For such models to account for the standard properties of psychometric curves in a bisection procedure (timescale invariance and bisection point at the geometric mean), they have to assume a logarithmic representation of time with constant variance. It follows that any argument against log timing is also an argument against this type of associative account of timing.

2. The time-left procedure

An argument against log timing is supposed to be provided by the results of a complex timing procedure designed by Gibbon and Church (1981), the time-left procedure. The procedure comes in two variants, but for our present purposes the most important one is the variant initially developed for pigeons (see Gibbon and

Church, 1981; also Machado and Vasconcelos, 2006). Fig. 3 shows the details: During the initial link of a concurrent-chain schedule the animal can choose between two keys: the “time-left” key (TL) and the “standard” key (Std). After a variable interval elapses, the first choice initiates the terminal link. If the animal chose the Std key, the Std key changes color, the TL key is turned off, and reinforcement is made available S sec later. If the animal chose the TL key, the TL key does *not* change color, the Std key is turned off, and reinforcement is made available $C-T$ sec later, where T is the time spent in the initial link.

According to Gibbon and Church (1981), animals in that situation form three representations, $f(C)$, the representation of the time-to-reinforcement on the TL key, $f(S)$, the representation of the time-to-reinforcement on the standard terminal link, and $f(t)$, the representation of the time elapsed in the initial link. They compute $f(C) - f(t)$, a representation of the time-left-to-reinforcement on the time-left key and compare it to $f(S)$, the time-to-reinforcement on the standard terminal link, and choose the option associated with the shorter time. It is expected that the animal initially will prefer the standard side but then, later in the trial, it will prefer the time-left side. More important for present purposes is that the point of indifference t^* predicted by the linear and logarithmic schemes will differ significantly. By definition, at the point of indifference, $f(C) - f(t^*) = f(S)$. If the representation of time is linear, $t^* = C - S$, but if the representation is logarithmic, $t^* = C/S$. Hence, if C and S change such that their ratio remains constant, the linear scheme predicts that t^* will change, but the logarithmic scheme predicts that it will not. An invariant indifference point would be extremely surprising, and indeed, it was not the result obtained by Gibbon and Church (1981). The results are shown in Fig. 4. The C/S ratio was held constant at 2 while the value of C varied between 15, 30, 60 and 90 s. As expected, the pigeons preferred the Std key early into the trial and the TL key late into the trial. In addition, consistent with the linear scheme, the indifference point changed when the absolute values of C and S were changed but their ratio remained constant. The vertical lines in Fig. 4 show where the indifference points

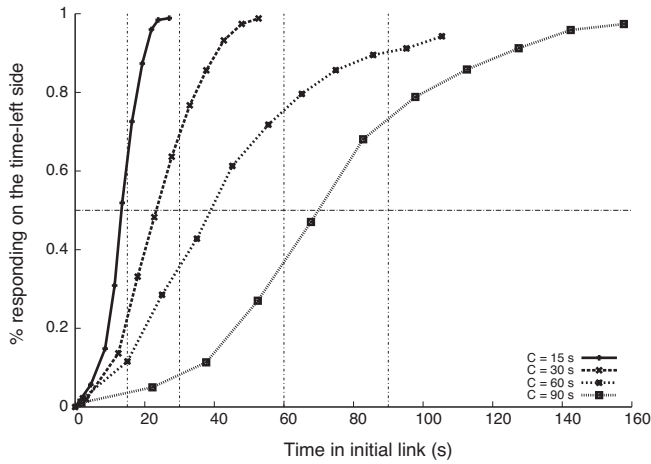


Fig. 4. Proportion of responding on the time-left side as a function of time in the initial link and the overall time-to-reinforcement on the time-left side (C) in Gibbon and Church's (1981) study with pigeons. The C/S ratio was held constant at 2 (S being the duration of the standard terminal link). The vertical line indicates where the indifference points should have been according to the cognitive model with linear timing. The cognitive model with log timing predicts that the animals should have switched to the time-left side 2 s in the initial link, no matter the actual value of C and S .

Redrawn from Gibbon and Church (1981).

should have been according to the linear model: As can be seen, the animals switch to the time-left side much earlier than expected by the linear model. Overall, given that $C/T=2$, the indifference point should have equaled $t^*=S$. Instead, the best-fitting regression lines had equation $t^*=0.74S+0.64$ ($R^2=0.97$, with no systematic deviation from a straight line). The strong bias toward the time-left side has been observed in subsequent replications of the concurrent-chain time-left variant (Cerutti and Staddon, 2002; Gibbon et al., 1988; Gibbon and Fairhurst, 1994; Preston, 1994), but no principled explanation has been provided for it in cognitive models.

Though the time-left experiment is considered the definitive argument against log timing in non-human animals, the entire argument advanced by Gibbon and Church assumes a cognitive model in the sense that the animal represents all the relevant durations and uses them to decide where to respond. Hence, the time-left procedure is an argument against log timing *within the context of a cognitive model*. For it to count against log timing in general, we need to show that even with an associative learning model still we cannot explain the standard finding of the time-left experiment (i.e. the indifference point is determined by the $C-S$ difference, not by the C/T ratio).

What would an associative account of the time-left procedure be like? Again, as in the previous section, we will assume a model that (a) represents time through a random variable, (b) updates the associations through a Bush–Mosteller learning rule, and (c) only updates the associations of the just-emitted response. These assumptions are met by BEM but might also be met by other past, present, and future associative accounts of timing. The traditional account of concurrent chain schedule performance is that choice during the initial link is determined by the conditioned reinforcement values of the terminal links, which are hyperbolic function of the delays to reinforcement signaled by the terminal link stimuli (see, for instance, Mazur, 2001 or Williams, 1988). To extend this account to the time-left experiment, we assume that the animal is timing the interval spent in the initial link, t , through the internal representation $f(t)=x$. Let $V_{TL}(x)$ be the strength of the association between the value x of the representation and the TL key. When the representation takes the value x , the animal chooses the TL side with a probability $P(TL|x)$ and the Std side

with probability $P(Std|x)=1-P(TL|x)$. If the animal chooses the Std key, nothing happens to $V_{TL}(x)$; if the animal chooses the TL key then $V_{TL}(x)$ changes, increasing if the animal entered the TL terminal link (conditioned reinforcement) and decreasing if it did not (extinction). With conditioned reinforcement $V_{TL}(x)$ increases by $\Delta V_{TL}(x)=\beta[(\lambda/C-t)-(V_{TL}(x))]$ where β is the learning rate following reinforcement, λ the amount of primary reinforcement the animal will collect at the end of the TL terminal link, and t is the actual time spent responding during the initial link. The equation is similar to the equation for the bisection procedure, except that the conditioned reinforcement value of the stimulus signaling the time-left terminal link substituted for the amount of primary reinforcement. The change in $V_{TL}(x)$ happens with a probability $P(t|x)P_R(t)$ where $P(t|x)$ is the probability that t sec have elapsed given that the representation took the value x , and $P_R(t)$ is the probability that the VI initial link has scheduled a transition to the terminal link t sec after the beginning of a trial. If the animal chooses the time-left side and does not receive the (conditioned) reinforced, $V_{TL}(x)$ decreases by $\Delta V_{TL}(x)=-\alpha V_{TL}(x)$ where α is the learning rate for non-reinforcement. This change occurs with probability $P(t|x)[1-P_R(t)]$. Assuming that the maximum duration of the initial link is M sec and that a transition to the terminal link can take place at any time during a trial, we have

$$E[\Delta V_{TL}(x)] = P(TL|x) \int_0^M P(t|x) \left\{ P_R(t) \beta \left[\frac{\lambda}{C-t} - V_{TL}(x) \right] - [1 - P_R(t)] \alpha V_{TL}(x) \right\} dt$$

Following the same reasoning for $V_{Std}(x)$, the associative strength between x and the Std key, yields

$$E[\Delta V_{Std}(x)] = P(Std|x) \int_0^M P(t|x) \left\{ P_R(t) \beta \left[\frac{\lambda}{S} - V_{Std}(x) \right] - [1 - P_R(t)] \alpha V_{Std}(x) \right\} dt$$

At the asymptote, the associative strengths are stable. Hence, $E[\Delta V_{TL}(x)] = E[\Delta V_{Std}(x)] = 0$. Assuming that both $P(TL|x)$ and $P(Std|x)$ are different from 0, it follows that

$$\int_0^M P(t|x) \left\{ P_R(t) \beta \left[\frac{\lambda}{C-t} - V_{TL}(x) \right] - [1 - P_R(t)] \alpha V_{TL}(x) \right\} dt = 0$$

and

$$\int_0^M P(t|x) \left\{ P_R(t) \beta \left[\frac{\lambda}{S} - V_{Std}(x) \right] - [1 - P_R(t)] \alpha V_{Std}(x) \right\} dt = 0$$

When $\alpha = \beta$, then

$$V_{TL}(x) = \frac{\int_0^M \frac{P(t|x)P_R(t)\lambda}{C-t} dt}{\int_0^M P(t|x) dt} \quad (2)$$

and

$$V_{Std}(x) = \frac{\int_0^M \frac{P(t|x)P_R(t)\lambda}{S} dt}{\int_0^M P(t|x) dt} \quad (3)$$

The animal picks the TL side if $V_{TL}(x) > V_{Std}(x)$. That is to say, from Eqs. (2) and (3), it picks the TL side if

$$\int_0^M P(t|x)P_R(t) \frac{\lambda}{C-t} dt > \int_0^M P(t|x)P_R(t) \frac{\lambda}{S} dt,$$

or, equivalently, whenever the average payoff for choosing the TL side when the representation takes the value x is higher than the

average payoff for choosing the Std side when the representation takes the value of x .

The indifference point is reached when $V_{TL}(x) = V_S(x)$. It is not easy to see when this happens from Eqs. (2) and (3), except for a few special cases, which might still help us to understand how the indifference point varies as a function of C and S . For instance, consider the case where the animal has a noiseless representation of time, so that, after t_x sec spent responding in the initial link, the representation always takes the value x . In this case, $P(t|x)=1$ if $t=t_x$, 0 otherwise. Then, $V_{TL}(x) = P_R(t_x) \frac{\lambda}{C-t_x}$ and $V_{Std}(x) = P_R(t_x) \frac{\lambda}{S}$. The animal is indifferent between the TL and the Std keys when the representation takes the value x^* , yielding $V_{TL}(x^*) = V_{Std}(x^*)$. Hence, indifference is reached at $t_{x^*} = C - S$, the same indifference point predicted by the linear cognitive model. Because we made no assumption concerning the encoding of time in our analysis, the noiseless, (conditioned) reinforcement-following associative model will learn to be indifferent at $t^* = C - S$, regardless of whether time is represented linearly or logarithmically.

It is likely that part of this result, at least the dependence of the indifference point on the $C - S$ difference, will transfer to a model with noise in the representation of time. Indeed, Fig. 5 shows the simulation results of Jozefowicz et al.'s (2009a) BEM model, an associative model with a logarithmic representation of time that meets the assumptions (a–c) above (details of the simulation can be found in Jozefowicz et al., 2009a,b). The model has only one free parameter, the level of noise in the representation of time, and this parameter was held constant across conditions. Just as in Gibbon and Church's (1981) original study, the C/T ratio was held constant at 2. As can be seen and as expected, the model correctly predicts that the indifference point is a function of the $C - T$ difference. Moreover, the vertical lines indicate where the indifference points should be located according to the equation $t^* = 0.74S + 0.64$ which fitted the data from Gibbon and Church (1981). The indifference points predicted by BEM are very close to the points predicted by the equation, which mean that the model anticipates the strong bias toward the time-left side observed in the data. This result was unexpected and it is not clear whether it is specific to BEM or whether it is a general prediction of the class of associative models of which BEM is a representative.³

Hence, the lesson of the time-left experiment is not the one commonly drawn, that it rules out log timing. Instead, it just rules out the combination of log timing with a cognitive model. More generally, as the bisection procedure tied log timing and (some) associative models together, the time-left experiment ties linear timing and cognitive models together. And just like any argument against log timing was an argument against a whole class of associative models, any argument against linear timing is an argument against a whole class of cognitive models. Such an argument may be given by an overlooked study by Yi (2008).

3. Log vs. linear timing in signal detection theory

Consider a bisection experiment in which an animal is reinforced for emitting response R1 following an S -sec stimulus and

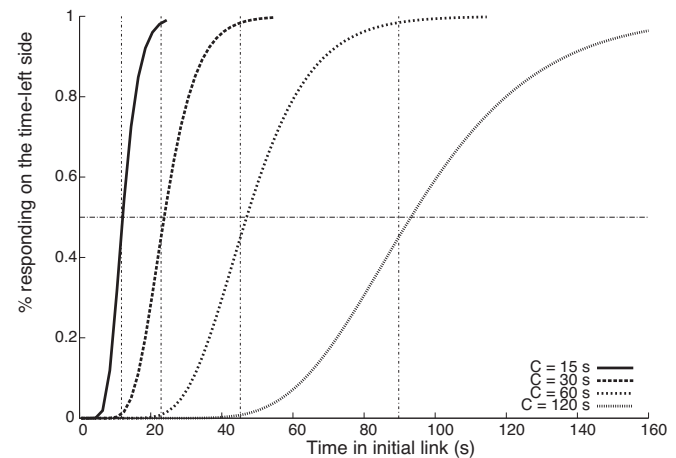


Fig. 5. Proportion of responding on the time-left side as a function of time in the initial link, in a simulation the time-left experiment by Jozefowicz et al.'s (2009a) BEM. As in the Gibbon and Church's (1981) study, the C/S ratio (C = overall time-to-reinforcement on the time-left side, S = time-to-reinforcement on the standard terminal link) was held constant at 2. The vertical lines indicate the location of the indifference points according to the equation $t^* = 0.74S + 0.64$ which predicted the indifference point with a R^2 of 0.97 in Gibbon and Church's original study.

Redrawn from Jozefowicz et al. (2009a).

response R2 following an L -sec stimulus ($S < L$). We examine the task from the perspective of signal detection theory (Wickens, 2002; Yi, 2008). Call stimulus L the signal and all other stimuli the noise. Hence, choosing R2 following stimulus L defines a hit and choosing R2 following any other stimulus t , with $S \leq t < L$ defines a false alarm. The representation of a t -sec stimulus, x , is a Gaussian random variable with mean μ_t and standard deviation σ_t , and similarly, the representation of an L -sec stimulus is a Gaussian random variable with mean μ_L and standard deviation σ_L . If x is above a criterion value x_c , the animal chooses R2; otherwise it chooses R1. Hence, the probability of a false alarm, P_F , is the probability that the representation of time, x , following a t -sec stimulus falls to the right of x_c , and the probability of a hit, P_H , is the probability that the representation of time, x , following a t -sec stimulus falls to the left of x_c . Both P_F and P_H are determined by x_c . By manipulating the value of x_c , one can draw a ROC curve relating P_H to P_F . If instead of P_H and P_F we use their associated z scores, the resulting curve is a z ROC curve. It can be shown that if the representations of time are Gaussian random variables, then (Wickens, 2002; Yi, 2008)

$$Z_H = \frac{\sigma_t}{\sigma_L} Z_F + \frac{\mu_L - \mu_t}{\sigma_L}$$

Hence, if the slope of z ROC curve is equal to 1, the standard deviations of the distributions of subjective time are equal.

Whether standard deviations are equal or not is what Yi (2008) tried to figure out. She trained rats in a 4s vs. 16s bisection. The animals were then tested on unreinforced probe trials with 7 durations ranging from 4 to 16s. This allowed her to plot 6 z -ROC curves (4s vs. 16s, 5.04s vs. 16s, 6.35s vs. 16s, etc.). To manipulate the criterion, Yi could have varied the probabilities of reinforcement for R1 and R2, for example, and thereby bias responding toward one or the other response. The problem with this method is that the experiment needs to be run several times, one for each reinforcement probability pair. As an alternative, Yi (2008) used the confidence rating method, which is a shortcut often used in signal detection theory to obtain ROC curves without having to run the experiment multiple times. Initially developed for humans (Wickens, 2002), the technique was adapted for animals by Blough (1967) and Wright (1972). In humans, the technique asks the participant to judge how confident he is in his

³ Indeed, the reasons why the bias is predicted remain unclear. The bias might be partially due to the log representation used by BEM (Jozefowicz et al., in press), although it might also be a consequence of the fact that the hyperbolic value functions for the time-left and standard sides (Eq. (2)) are extremely asymmetric around the indifference point $(C+S)/2$. The conditioned reinforcement value of the time-left side is below that of the standard side before the point of indifference and above it after the point of indifference. However, the difference between the functions is much larger past the point of indifference. As associative models are sensitive to differences in the amount of reinforcement (Jozefowicz et al., 2009a, in press), this difference may explain why they predict a bias toward the time-left side (see also Cerutti and Staddon, 2004, for a similar line of reasoning).

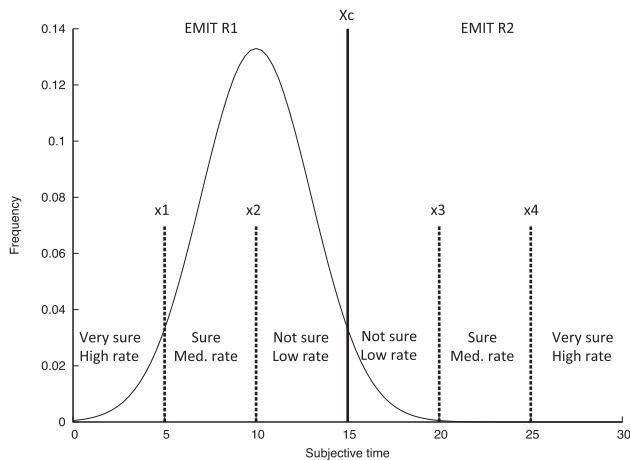


Fig. 6. Signal detection analysis of a bisection task. When presented with a stimulus, the perception of time x is drawn from a Gaussian distribution. If the response is above the threshold x_c , the animal emits R2, otherwise R1. The subthreshold x_1, x_2, \dots determines the response of a human subject to a confidence rating question ("Very sure", "Sure", "Not sure", ...) or the response rate (high, medium, low) in an animal.

categorization of the stimulus. The idea is that each confidence level corresponds to a different criterion. To illustrate, the criterion x_c determines whether the subject respond R1 or R2, but each of the secondary criteria x_1, x_2, \dots determines how confident he will be. For instance, in Fig. 6, if the representation falls left of x_1 , the subject will emit R1 and will be very confident in his choice. If the representation falls between x_3 and x_4 , the subject will emit R2 and will be mildly confident in his choice. Hence, at the end of the study, the experimenter not only knows the probability the subject will choose R1 and R2, but also the probability that he will choose R1 and be very confident about it, will choose R1 and be mildly confident about it, etc. The confidence rating method simulates a change in the criterion. For instance, if x_c had been at x_1 the probability of false alarm would have been the probability of choosing R2 plus the probability of choosing R1 and being sure or not sure. The probability of a hit would have been the probability of choosing R2 plus the probability of choosing R1 and being sure or not sure. The two values thus computed would give the first point of the ROC curve. To obtain the second point of the ROC curve, we assume that x_c equals x_2 . In this case, the probability of a false alarm would equal the probability of choosing R2 plus the probability of choosing R1 and not being sure, whereas the probability of a hit would have been the probability of choosing R2 plus the probability of choosing R1 and not being sure. The remaining points of the ROC curve would be computed in similar ways.

When applied to animals (Blough, 1967; Wright, 1972), the response rate of the animal is used as an indicator of its level of confidence. The implicit assumption is that the further away the representation is from x_c , the higher the response rate. This makes intuitive sense. For instance, in a bisection task, values of x further away to the right of x_c are sampled almost exclusively when a L -sec stimulus is shown. That means that they are strongly correlated with the reinforcement of R2 and hence should trigger a strong reward expectation expressed by a high response rate. In contrast, values close to x_c are sampled as often following an S -sec stimulus as an L -sec stimulus and hence are correlated with a strong level of error, and hence with a lower rate of reinforcement. A low expectation of reinforcement should be expressed by a low response rate (see Jozefowicz et al., 2009b for a similar analysis of experiments on animal metacognition). The response rate distribution is then discretized using percentiles in order to create the various confidence levels. The analysis used in humans and summarized above

is then applied to simulate the change in the criterion and to draw the ROC curve.

This is the technique that Yi (2008) used. In order to obtain response rate data, she modified the bisection paradigm slightly. Normally, the first choice following the sample stimulus either leads to reinforcement or non-reinforcement, but in Yi's (2008) study, the end of the trial (by reinforcement in case of a correct choice, by non-reinforcement in case of an incorrect choice or of a probe trial) was scheduled according to a random-interval 10-s schedule. As the rats did not switch from one lever to other once they made their first choice, Yi (2008) not only knew which lever they chose but also how fast they pressed that lever, allowing her to apply the confidence rating technique and draw the z -ROC curves.

The results are displayed in Fig. 7. The z -ROC curves are clearly parallel and have slope close to 1. Hence, Yi's data strongly support the equal variance model. To see what these results imply for cognitive and associative models of timing, we need to better understand the relationship between those models and signal detection theory. For associative models, this is straightforward. We have already shown that, in a simple bisection task, some associative models are mathematically equivalent to a signal detection model. Hence, these models can account for Yi's data only if they assume constant variance. The right panel of Fig. 7 shows that this type of model fits the data well. Following Gibbon (1981), we previously showed that, in order to account for the bisection point at the geometric mean and the timescale invariant psychometric function in the bisection procedure, this class of associative models had to assume log timing, which implies constant variance. We have reached the same conclusion by analyzing Yi's data: Associative models and log timing are closely intertwined.

What about cognitive models? As we mentioned previously, Gibbon (1981) had shown that, if we assume that temporal intervals are represented through Gaussian random variables, only two types of cognitive models could account for the bisection data: linear timing with standard deviation proportional to the mean coupled with a ratio decision rule, and log timing with constant variance coupled with a difference decision rule. The time-left experiment discarded the log timing model. Hence, we are left with the linear model. In this framework, the animal has a STM representation $f(t)$ of the duration of the current stimulus and it compares it to the LTM representations $X(S)$ and $X(L)$ of the two reinforced durations S and L in order to determine which option to choose. The three representations, $f(t)$, $X(S)$, and $X(L)$ are all random variables following a Gaussian distribution with mean proportional to the duration they represent (t, S , and L respectively) and standard deviation proportional to the mean. The animal chooses R2 provided

$$\frac{f(t)}{X(S)} > \theta \frac{X(L)}{f(t)}$$

where θ is a bias term (equal to 1 if the animal is not biased). This equation can be rewritten,

$$\frac{f(t)^2}{X(S)X(L)} > \theta \quad (4)$$

Now, the left term of Eq. (4) is a rather complex random variable: The ratio of the square of a Gaussian random variable and the product of two Gaussian variables. For this reason, most people, starting with Gibbon (1981), assume that there is no noise in LTM: $X(S)=S$, $X(L)=L$; only $f(t)$ is a random variable. When he compared this model with a version closer to Eq. (4), where both $X(S)$ and $X(L)$ were random variables, Gibbon (1981) concluded that the noise-only-STM model did a better job at fitting the data, though the differences between the two models were small. Research with animals (Rodríguez-Gironés and Kacelnik, 2001) and humans (Allan, 2002a) has suggested that the noise in STM is more important than

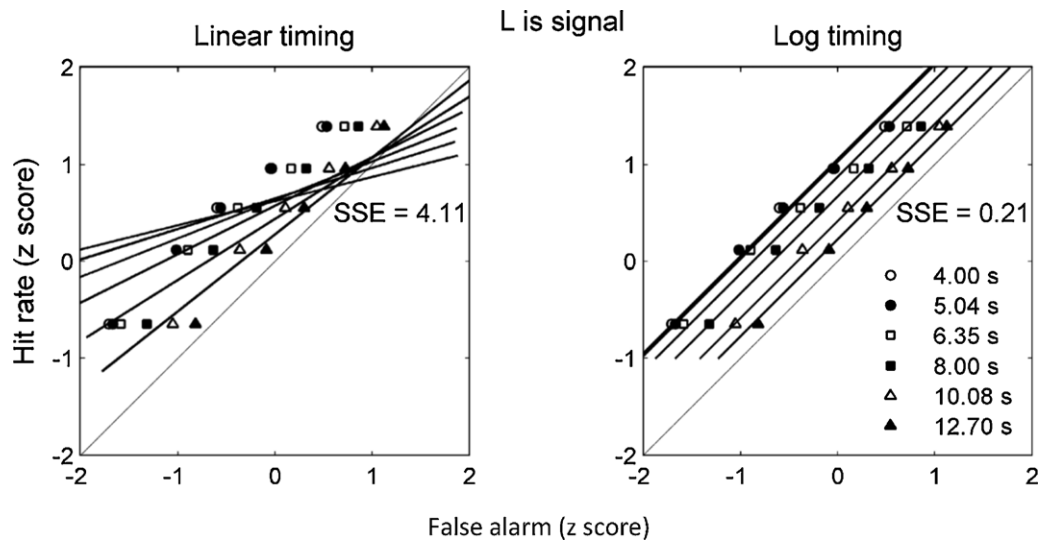


Fig. 7. z-Score of the probability of hit as a function of the z-score of the probability of false alarm in Yi's (2008) bisection study. The left panel shows the best fitting lines if we assume linear timing with standard deviation proportional to the mean. The right panel shows the best fitting lines if we assume log timing with constant variance. From Yi (2008).

the noise in LTM, justifying Gibbon's (1981) simplifying assumption. Under this assumption, Eq. (4) becomes

$$f(t) > \theta\sqrt{XS}$$

The preceding condition is equivalent to a signal detection model with the criterion set at $\theta\sqrt{XS}$: Biasing the performance (changing θ) is equivalent to a change in the criterion. Hence, it is unable to account for Yi's data since it assumes a non-constant variance. Indeed, the left panel of Fig. 7 shows how this type of model accounts for the data. It predicts a strong fanning-out of the zROC curves as the ratio between the noise and the signal is increased, which is nowhere to be seen in the data.

The only way to rescue the cognitive model is to go back to Eq. (4) and use the full model, with noise in both STM and LTM. Though the resulting random variable is complex, the calculations can be done numerically. The resulting z-ROC curves are shown in Fig. 8. As can be seen, the fanning-out of the z-ROC curves predicted by the model without noise in LTM is greatly reduced, to the point that the model cannot be rejected at once. But, as stated previously, nobody has used this model, mainly because of its complexity, but also because both theoretical (Gibbon, 1981) and empirical (Allan, 2002a,b; Rodríguez-Gironés and Kacelnik, 2001) arguments seem

to favor the model with noise in STM only. If this is the case, Yi's (2008) data provide a strong argument against linear timing and cognitive models, while making a case for log timing and associative models.

4. Conclusion

Various models of timing have been proposed in recent years. They differ with regard to the representation of time (linear vs. log timing) and to whether they are cognitive or associative models, that is to say, whether they assume that the animals learn the representations of the time of reinforcement and explicitly compare them, or whether they learn associations between their representation of time and responding and respond on the basis of these associations. In this article, we showed that these two types of questions might not be independent: Linear representations may imply a cognitive model; and log representations may imply an associative model.

Our argument relies on constraints imposed by several well-known empirical facts: The timescale invariance of the bisection curves, the location of the bisection point at the geometric mean between the two trained durations, and the results of the time-left procedure. The argument would be weakened if any of these findings proves to be unreliable. The time-left procedure has been strongly criticized in recent years (i.e. Cerutti and Staddon, 2004; Machado and Vasconcelos, 2006; Preston, 1995; Vieira do Castro and Machado, 2010). Researchers have challenged the cognitive account provided by Gibbon and Church (1981) and have pointed that several variables overlooked by Gibbon and Church are critical to explain performance in the task. Overall, these critics lower the probability that the cognitive account of the time-left procedure is correct, but the main results as far as our analysis are concerned (increased probability of choosing the time-left side with time spent in the initial link, indifference point being a function of the $C - T$ difference) have all been replicated.

Though exceptions might exist (i.e. Bizo et al., 2006; Zeiler and Powell, 1994), the generality of the timescale invariance property in timing procedures, including the bisection one, seems well established (i.e. Church, 2004; Wearden and Lejeune, 2006, 2008). The location of the bisection point at the geometric mean is also considered a canonical result (Church, 2004), even though some have argued that not enough pairs of S and L durations have been tested

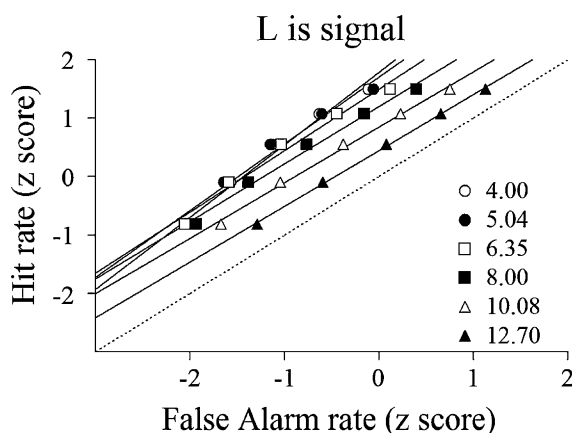


Fig. 8. Simulation of Yi's (2008) experiment by a cognitive model assuming linear timing with standard deviation proportional to the mean and noise in both short-term and long-term memory.

to determine whether or not the bisection point approaches the harmonic mean when the ratio L/S is large (see Machado et al., 2009). A reliable exception concerning the location of the bisection point is the set of studies with human subjects that have found bisection at the arithmetic mean of the trained durations (e.g. Allan, 2002b; Allan and Gerhardt, 2001; Droit-Volet and Wearden, 2001; Wearden, 1991; Wearden and Ferrara, 1995, 1996; but see Allan and Gibbon, 1991). It is not clear why such a discrepancy is observed between human and non-human animals. Still, a bisection point at the arithmetic mean is difficult to explain if one assumes log timing. Hence, this result, along with others (i.e. Wearden and Jones, 2007) might indicate that time is encoded linearly in humans. It may also rule out some associative accounts of human timing. It would be worth carrying more research on this topic, notably by investigating whether a signal detection analysis such as the one carried by Yi (2008) would yield in humans the same results that it did in animals, results that strongly support log timing and associative models.

Though we tried to characterize cognitive and associative models in terms as general as possible, we still had to make some assumptions regarding each class of models. The conclusions we have drawn might therefore be valid only when these assumptions are met. For cognitive models, we assumed that temporal intervals were represented by random variables. This seems general enough to encompass most cognitive models. For associative models, we also assumed that temporal intervals were represented by the mean of random variables. Moreover, we assumed that the Bush–Mosteller learning rule applied (this is, for instance, not the case of the timing/conditioning model developed by Ludvig, Sutton, & Kehoe, 2012, which uses the more sophisticated TD learning rule. Sutton and Barto, 1998) and that only the associations with the just-emitted response were updated (this is not the case in both versions of the LeT model, Machado, 1997; Machado et al., 2009). Moreover, we assumed Gaussian random variables (the first version of LeT, Machado, 1997, and BeT, Killeen and Fetterman, 1988, used Poisson distributions). In this paper, we have shown that those assumptions are sufficient for our conclusions to hold. It would be extremely valuable to know if they are also necessary. As already mentioned, the new version of LeT (Machado et al., 2009) violates our assumption concerning the fact that only the associations of the just-emitted response are updated. Yet it faces the same problem with the location of the bisection point as the linear version of BEM, an associative model that fits all our assumptions. Also, the reason why most models use Gaussian distributions is that non-Gaussian distributions do not fair very well. For instance, BeT and the first version of LeT, which both uses Poisson distributions, do not generate timescale invariant psychometric curve without the additional assumption linking arousal and time perception. Hence, we suspect that our conclusions may remain valid even if some of the assumptions we made are violated.

As a final note, let's point out that the controversy between associative and cognitive models of timing has implications for discussions about the relation between timing and associative learning. That the two are related in some way seems clear: Associative learning may be conceived of as the ability to perceive contingency relations between two events and interval timing as the ability to perceive temporal intervals between two events. Hence, it is not surprising to see in recent years an attempt to develop integrated models of timing and conditioning. A dichotomy which is starting to appear in those models is that between models that reduce associative learning to timing (i.e. Balsam et al., 2010; Balsam and Gallistel, 2009; Gallistel and Gibbon, 2000) and models that reduce timing to associative learning (i.e. Ludvig et al., 2012). This dichotomy parallels that between cognitive and associative models of timing, with cognitive models often used by researchers arguing that associative learning is nothing but timing, and

associative models often used by researchers arguing that timing simply extends the basic associative processes. Hence, clarifying the controversy between cognitive and associative models of timing might have implications well beyond the field of interval timing.

Acknowledgements

Jeremie Jozefowicz and Armando Machado acknowledge support from the Fundação para a Ciência e a Tecnologia as well as from the European project COST ISCH Action TD0904 "Time in Mental activity" (www.timely-cost.eu).

References

- Allan, L.G., 2002a. Are the referents remembered in temporal bisection? *Learn. Motiv.* 33, 10–31.
- Allan, L.G., 2002b. The location and interpretation of the bisection point. *Q. J. Exp. Psychol.* 55B, 43–60.
- Allan, L.G., Gerhardt, K., 2001. Temporal bisection with trial referents. *Percept. Psychophys.* 63, 524–540.
- Allan, L.G., Gibbon, J., 1991. Human bisection at the geometric mean. *Learn. Motiv.* 22, 39–58.
- Balsam, P.D., Drew, M.R., Gallistel, C.R., 2010. Time and associative learning. *Comp. Cogn. Behav. Rev.* 5, 1–22.
- Balsam, P.D., Gallistel, C.R., 2009. Temporal maps and informativeness in associative learning. *Trends Neurosci.* 32, 73–78.
- Bizo, L.A., Chu, J.Y.M., Sanabria, F., Killeen, P.R., 2006. The failure of Weber's law in time perception and production. *Behav. Processes* 71, 201–210.
- Blough, D.S., 1967. Stimulus generalization as a signal detection in pigeons. *Science* 158, 940–941.
- Bush, R.R., Mosteller, F., 1955. *Stochastic Models for Learning*. Wiley, New York.
- Cerutti, D.T., Staddon, J.E.R., 2004. Immediacy versus anticipated delay in the time-left experiment: a test of the cognitive hypothesis. *J. Exp. Psychol. Anim. Behav. Process.* 30, 45–57.
- Church, R.M., 2004. Temporal learning. In: Pashler, H., Gallistel, C.R. (Eds.), *Stevens' Handbook of Experimental Psychology: Learning, Motivation and Emotion*, vol. 3, 3rd edition. Wiley, New York, pp. 365–393.
- Church, R.M., Deluty, M.Z., 1977. Bisection of temporal intervals. *J. Exp. Psychol. Anim. Behav. Process.* 3, 216–228.
- Droit-Volet, S., Wearden, J.H., 2001. Temporal bisection in children. *J. Exp. Child Psychol.* 80, 142–159.
- Gallistel, C.R., Gibbon, J., 2000. Time, rate, and conditioning. *Psychol. Rev.* 107, 289–344.
- Gibbon, J., 1981. On the form and location of the psychometric bisection function for time. *J. Math. Psychol.* 24, 58–87.
- Gibbon, J., Church, R.M., 1981. Time left: linear versus logarithmic subjective time. *J. Exp. Psychol. Anim. Behav. Process.* 22, 3–28.
- Gibbon, J., Church, R.M., Meck, W.H., 1984. Scalar timing in memory. *Ann. N.Y. Acad. Sci.* 423, 52–77.
- Gibbon, J., Church, R.M., Fairhurst, S., Kacelnik, A., 1988. Scalar expectancy theory and choice between delayed rewards. *Psychol. Rev.* 95, 102–114.
- Gibbon, J., Fairhurst, S., 1994. Ratio versus difference comparators in choice. *J. Exp. Anal. Behav.* 62, 409–434.
- Guilhardi, P., Yi, L., Church, R.M., 2007. A modular theory of learning and performance. *Psychon. Bull. Rev.* 14, 543–559.
- Jozefowicz, J., Staddon, J.E.R., Cerutti, D.T., 2009a. The behavioral economics of choice and interval timing. *Psychol. Rev.* 116, 519–539.
- Jozefowicz, J., Staddon, J.E.R., Cerutti, D.T., 2009b. Animal metacognition: how do we know that they know? *Comp. Cogn. Behav. Rev.* 4, 29–39 (retrieved from <http://psyc.queensu.ca/ccbr/index.html?doi:103819/ccbr.2009.4003>).
- Jozefowicz, J., Machado, A., Staddon, J.E.R. Cognitive versus associative decision rules in timing. In: Arstila, V., Lloyd, D. (Eds.), *Subjective Time*. Cambridge, MA: MIT Press, in press.
- Killeen, P.R., Fetterman, J.G., 1988. A behavioral theory of timing. *Psychol. Rev.* 95, 274–295.
- Ludvig, E.A., Sutton, R.S., Kehoe, E.J., 2012. Evaluating the TD model of classical conditioning. *Learn. Behav.* 40, 305–319.
- Machado, A., 1997. Learning the temporal dynamics of behavior. *Psychol. Rev.* 72, 195–206.
- Machado, A., Vasconcelos, M., 2006. Acquisition versus steady state in the time-left experiment. *Behav. Process.* 71, 172–187.
- Machado, A., Malheiro, M.T., Erlhagen, W., 2009. Learning to time: a perspective. *J. Exp. Anal. Behav.* 92, 423–458.
- Mazur, J.E., 2001. Hyperbolic value addition and general models of animal choice. *Psychol. Rev.* 108, 96–112.
- Preston, R.A., 1994. Choice in the time-left experiment and in concurrent chains with a time-left terminal link. *J. Exp. Anal. Behav.* 61, 349–373.
- Rodríguez-Gironés, M.A., Kacelnik, A., 2001. Relative importance of perceptual and mnemonic variance in human temporal bisection. *Q. J. Exp. Psychol.* 54, 527–546.
- Rescorla, R.A., Wagner, A.W., 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, A.H.,

- Prokasy, W.F. (Eds.), *Classical Conditioning II*. Appleton Century Croft, New York, pp. 64–99.
- Stubbs, D.A., 1968. The discrimination of stimulus duration by pigeons. *J. Exp. Anal. Behav.* 11, 223–238.
- Sutton, R.S., Barto, A.G., 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.
- Vieira do Castro, A.C., Machado, A., 2010. Prospective timing in pigeons: isolating temporal perception in the time-left procedure. *Behav. Processes* 84, 490–499.
- Wearden, J.H., 1991. Human performance on an analogue of an interval bisection task. *Q. J. Exp. Psychol.* 43B, 59–81.
- Wearden, J.H., Jones, L., 2007. Is the growth of subjective time in humans a linear or nonlinear function of real time? *Q. J. Exp. Psychol.* 60, 1289–1302.
- Wearden, J.H., Lejeune, H., 2006. Scalar properties in animal timing: conformity and violations. *Q. J. Exp. Psychol.* 59, 1875–1908.
- Wearden, J.H., Lejeune, H., 2008. Scalar properties in human timing: conformity and violations. *Q. J. Exp. Psychol.* 61, 569–587.
- Wearden, J.H., Ferrara, A., 1995. Stimulus spacing effects in temporal bisection by humans. *Q. J. Exp. Psychol.* 48B, 289–310.
- Wearden, J.H., Ferrara, A., 1996. Stimulus range effects in temporal bisection by humans. *Q. J. Exp. Psychol.* 49B, 24–44.
- Williams, B.A., 1988. Reinforcement, choice, and response strength. In: Atkinson, R.C., Herrnstein, R.J., Lindzey, G., Luce, R.D. (Eds.), *Stevens' Handbook of Experimental Psychology: Vol 2: Learning and Cognition*. , 2nd edition. Wiley, New York, pp. 167–244.
- Wright, A.A., 1972. Psychometric and psychophysical hue discrimination functions for the pigeon. *Vision Res.* 12, 1447–1464.
- Wickens, T.D., 2002. *Elementary Signal Detection Theory*. Oxford University Press, Oxford.
- Yi, L., 2008. Do rats represent time logarithmically or linearly? *Behav. Processes* 81, 274–279.
- Zeiler, M.D., Powell, D.G., 1994. Temporal control in fixed-interval schedules. *J. Exp. Anal. Behav.* 61, 1–9.