GAMIT – A Fading-Gaussian Activation Model of Interval-Timing: Unifying Prospective and Retrospective Time Estimation

Robert M. French[1], Caspar Addyman[2], Denis Mareschal[2], and Elizabeth Thomas[3]

[1]*Laboratory for Research on Learning and Development (LEAD-CNRS), Université de Bourgogne, Dijon, France*
[2]*Centre for Brain and Cognitive Development, Birkbeck, University of London, London, UK*
[3]*Institut National de la Santé et de la Recherche Médicale (INSERM), Université de Bourgogne, Dijon, France*

Two recent findings constitute a serious challenge for all existing models of interval timing. First, Hass and Hermann (2012) have shown that only variance-based processes will lead to the scalar growth of error that is characteristic of human time judgments. Secondly, a major meta-review of over one hundred studies of participants' judgments of interval duration (Block et al., 2010) reveals a striking interaction between the way in which temporal judgments are queried (i.e., retrospectively or prospectively) and cognitive load. For retrospective time judgments, estimates under high cognitive load are longer than under low cognitive load. For prospective judgments, the reverse pattern holds, with increased cognitive load leading to shorter estimates. We describe GAMIT, a Gaussian spreading activation model of interval timing, in which the decay and sampling rate of an activation trace are differentially affected by cognitive load. The model unifies prospective and retrospective time estimation, normally considered separately, by relating them to the same underlying process. The scalar property of time estimation arises naturally from the model dynamics and the model shows the appropriate interaction between mode of query and cognitive load.

Keywords: Timing and time perception, attention, memory, computational models

1   Introduction

Time perception is central to cognition in humans and other animals (for extended reviews see Buhusi & Meck, 2005; Gibbon & Allan, 1984; Grondin, 2008, 2010; Merchant et al., 2013). It may even be central to explaining conditioned learning (Gallistel & Gibbon, 2000). Estimates of time intervals on the order of half a second to several minutes are affected by many factors including level of attention, the intensity of the stimuli, and whether the judgments were made prospectively or retrospectively. Existing models of interval timing have been successful at explaining some of these effects but to date no single model captures them all. Moreover, models have focused on either prospective *or* retrospective time estimates, but rarely both at the same time.

In this paper we will focus on the modeling of three key phenomena. Firstly, extensive empirical evidence (Gibbon, 1977; Gibbon & Allan, 1984; Matell & Meck, 2000; Meck, 2003, 2005) suggests that time-estimation errors in interval times grow approximately linearly with the size of the estimate. Known as the *scalar property* of time estimation, this fact sets a hard constraint on the nature of the underlying processes involved in time estimation (Hass & Herrmann, 2012). Even though there is some disagreement over the validity of the scalar property for interval timing (e.g., Bangert et al., 2011; Burr et al., 2013; Grondin, 2012; Lewis & Miall, 2009), it has proved to be a

formidable obstacle for a number of existing models of interval-time judgments (Shi et al., 2013). In the GAMIT model, it falls directly out of the (biologically plausible) manner in which activation spreads.

Secondly, in a careful meta-analysis of well over one hundred studies, Block et al. (2010) established that human adults' perception of the passage of time differs according to whether they are forewarned that they will need to make a timing judgment, and are therefore actively attending to its passage (*prospective* time estimation), or whether they are required to make an unexpected, after-the-fact judgment of the passage of time (*retrospective* time estimation). And thirdly, this difference is heavily modulated by cognitive load, showing a classic cross-over interaction in which either prospective or retrospective judgments are longer depending on whether the participant experiences high or low cognitive load (Fig. 1).

These three findings — namely, the scalar property of time interval judgments, differential prospective and retrospective judgments, and the mediating effect of cognitive load —, taken together, pose significant challenges for existing computational models of interval-time judgments. While current models may be able to explain some of these phenomena, they need to appeal to secondary mechanisms to account for all of them. We will show that a model of time perception based on the idea of sampling a fading-Gaussian activation trace, GAMIT, naturally captures all three of these critical properties of interval time estimations. The key intuition is this: retrospective time judgments are based on the *amount* the Gaussian has faded; prospective time judgments additionally incorporate the *rate at which* the Gaussian is fading. This mechanism will allow us to explain the puzzling interaction of retrospective and prospective time estimation with high and low cognitive load recently reported by Block et al. (2010), thereby providing a single unifying explanation for both retrospective and prospective time estimates. In addition, we will show that the scalar property of time estimation arises naturally from the proposed Gaussian decay mechanism.

In the present article we implement the fading-Gaussian model of interval time estimation, using the classic equation of spreading activation as an approximation to the underlying stochastic processes involved in the spread of information in the distributed cognitive system. We show that such a model has linear growth in error (and, therefore, captures the scalar property), then turn to show that it also accounts for both the prospective and retrospective time judgment data presented in Block et al.'s (2010) meta-analysis.

However, before discussing GAMIT, we will begin with a brief discussion of existing models and their limitations.

1.1 An Overview of Existing Models of Interval Timing

There are three major paradigms for interval-time judgments: (1) pacemaker-accumulator models, (2) multiple oscillator-coincidence detector models (also sometimes called timestamp models), and (3) memory or neural process models. The first class of models relies on an internal pacemaker that emits regular, short pulses that are counted by an accumulator. The number of pulses stored in the accumulator gives the measure of the time that has passed (e.g., Gibbon et al., 1984; Taatgen et al., 2007; Wearden, 1991, 2001). A second class of models relies on multiple neuronal oscillators with coincidence detectors associating particular patterns of firing with given time intervals, effectively time-stamping when an event occurs (e.g., Church & Broadbent, 1990; Matell & Meck, 2000; Miall, 1989). An alternative type of oscillator-based timing model (e.g., Brown et al., 2000) assumes that some representation of the state of an already-running set of oscillators (started, say, at the birth of the individual), is associated with each event in memory, in essence, as one of the features of the event. The third class of models involves recovering the passage of time from a neural process that is decaying (Lewis & Miall, 2006; Staddon & Higa, 1999) or increasing (Reutimann et al., 2004). Here, the current state or change in state of the activation trace allows the system to recover the passage of time.

1.2 Interval Timing and the Scalar Property

Interval timing operates in the range from half a second to several minutes. Here humans and other animals show very similar abilities. If human adults or rats are required to reproduce a given time interval, their responses will have an approximately normal (right skewed) distribution peaked at the target value (e.g., Church et al., 1994; Lejeune et al., 1997; Rakitin et al., 1998). The *scalar property*, also referred to as time-scale invariance (Gibbon, 1977), states that the width of this distribution is directly proportional to the length of the interval. So, for example, the standard deviation for a distribution of estimates of an interval of 2X seconds will be (approximately) twice that for an interval of X seconds. This is equivalent to saying that time perception obeys Weber's Law (equal relative increments in a stimulus produce equal increments in sensation). This effect is very widely replicated with humans, pigeons, and rodents (see Buhusi et al., 2009; Gibbon et al., 1984, 1997; Penney et al., 2008). Similar behavioral responses to time

scales can even be found in rate-dependent habituation in *C. elegans* (Staddon & Higa, 1999).

No model that we are aware of accounts for the scalar property as an unavoidable consequence of the way the timing mechanism works (Hass & Hermann, 2012; Hass et al., 2008). For example, models based on repetitive clock-like processes have *less* intrinsic variability than predicted by the scalar property and have to introduce ad-hoc assumptions as to why the cognitive system cannot use these more precise quantities. Hass and Hermann (2012) use information theoretic arguments to show how the scalar property places several important restrictions on the nature of any interval timing mechanism. They show that, in order to display scalar error profiles, the neural process underlying time perception must be based on a measure of growing variance in the system. Power law decay functions found in memory-decay models would give rise to more than linear growth in error while the errors in accumulators and oscillators grow too slowly. Accumulator models base their estimates on mean number of accumulated ticks or oscillations. However, according to the *Central Limit Theorem,* such estimates have errors that grow with the square root of the total. Only with logarithmic decay does a constant error around activation values convert to a scalar error in magnitude.

Accumulator models cannot account for the scalar property of time without positing a secondary process that modifies the shape of the error distribution (Hass & Herrmann, 2012). Gibbon (1977) acknowledges this problem for the original Scalar Expectancy Theory (SET) pacemaker-accumulator model. In SET, the pacemaker is a Poisson process and variance in a cumulative Poisson process grows according to the square root. Gibbon et al. (1997, 1984) get round this by attributing the error primarily to a multiplicative factor associated with the comparison of accumulated estimates and their counterparts in memory, relying on a mathematical argument by Gibbon (1992). Decisions as to whether the clock has reached a given value are performed by seeing if the ratio of the accumulated value and the valued stored in memory is within a certain threshold. This ratio induces the scalar property and is central to permitting SET to fit the empirical data. However, it is unclear why this calculation has to be done using a ratio when comparisons of the absolute accumulator magnitudes are possible and would permit the cognitive system to make temporal judgments of greater accuracy. As Staddon and Higa (1999) observe, the assumptions behind SET are far from parsimonious and the neural mechanisms have not yet been described that can explain the full range of phenomena observed with interval timing.

In multiple oscillator models (Church & Broadbent, 1990; Miall, 1989) timing is measured by a large array of neuronal oscillators of different frequencies. An event starts all the oscillators simultaneously and at the end of the interval a coincidence detection mechanism learns which oscillators are in phase with each other. On future trials this same subset of the oscillators will also be in phase after the same amount of time has passed, allowing this signal to be used as timing mechanism. However, in general, this signal does not show the necessary scalar properties. In Miall's (1989) Beat-Frequency model, the distribution of firing was not normally distributed, having a sharp peak at the target time and smaller peaks at the major harmonics of the fundamental interval. In addition, the width of the peak was not proportional to the length of the interval.

Matell and Meck's (2004) Striatal Beat-Frequency model tried to address these problems. They made a sequence of modifications to Miall's model that induced the scalar property. This required globally varying oscillators to retain a significant degree of correlation with each other even as they drift out of synchrony. Recent work by Buhusi and Oprisan (e.g., Buhusi & Oprisan 2013; Oprisan & Buhusi, 2011) addresses this issue using more realistic, noisy neural oscillators and validates the initial approach of Matell and Meck (2004) which as subsequently been extended to include an unified account of duration-based and beat-based timing mechanisms (e.g., Allman et al., 2014; Teki et al., 2012).

A third class of model is based on memory decay and neural activation. Activation decay and growth processes are ubiquitous and well understood and can account for evidence that timing and memory use the same cognitive resources (Fortin, 1999; Fortin & Rousseau, 1998) and both recruit the dorso-lateral prefrontal cortex (Genovesio et al., 2006; Lustig et al., 2005; Wager & Smith, 2003). However, derivation of the scalar property is not always straightforward in these models. In the Multiple Time Scales model (MTS, Staddon & Higa, 1999), a series of leaky integrators with power law decay must be carefully chained together to approximate the required logarithmic function. The Temporal Context Model (TCM, Shankar & Howard, 2010) is built from many leaky integrators using complex dynamics.

By contrast, Reutimann et al. (2004) use a single climbing neuronal trace that reaches a threshold at the expected end of an interval. Single cell recordings in the inferotemporal cortex of monkeys have found neurons with the appropriate time-dependent firing rates (Kojima & Goldman-Rakic, 1982; Komura et al., 2001; Leon & Shadlen, 2003) – but see Kononowicz and Van Rijn (2014)

and Van Rijn et al. (2011) for a different interpretation of climbing neural activity. Learning of new intervals occurs via Hebbian learning within the adaptation process, such that neuronal firing reaches threshold at an earlier or later time. This threshold varies according to a normal distribution around a constant level. The interaction of the linearly increasing trace and the threshold gives rise to the scalar property. Advantages of this are that it is built on a single mechanism using well-understood principles of synaptic plasticity and the decision rule is built into the model itself. But the scalar property derives primarily from the Gaussian nature of the threshold which appears to be an arbitrary choice to fit the data. Recent work by Simen et al. (2011) extends this idea.

1.3   Retrospective and Prospective Time Estimation

Time judgments can be made with or without prior notice. In *retrospective time estimation* an individual is asked to estimate how long ago an event occurred without prior warning that they would have to do so. By contrast, in *prospective time estimation* the individual knows in advance that they will be asked to estimate the time that has elapsed from a particular event. Historically, these have been studied as separate phenomena. We believe that prospective and retrospective time estimation are intimately related and should not be considered as distinct phenomena. While we would not deny the fact that retrospective timing (especially for events that occurred long ago) relies in part on aspects of episodic memory and the reconstruction or reactivation of traces in memory, it seems unreasonable to suggest that once the memory has been retrieved, the timing mechanism that estimates how long ago the event occurred is completely different from the one that operates in other contexts.

In pacemaker-accumulator models the assumption that an accumulator is started when a particular event occurs makes this mechanism unsuitable for retrospective timing. All events are potential candidates for retrospective time judgments, which would require a separate accumulator for every event in memory that could be recalled and about which a time judgment might be requested.

Multiple oscillator accounts have the same "reset difficulty" that besets the pacemaker-accumulator models; namely, they imply a separate set of oscillators must be initialized for each potential event about which a time judgment will be made. An alternative approach would be to have a set of oscillators that is set in motion at the beginning of life, and that all events are time stamped with the values of these oscillators when the event occurs and that this set of oscillator values is an intrinsic part of their semantic representation. A model of serial order recall

(Brown et al., 2000) has some of these properties. However, it was not developed as a model of interval timing and only uses the timestamps to reconstruct the order of events in memory. Moreover scalar error growth and effects of cognitive load discussed below are not easily explained in this framework.

The problem of a system reset for each potential event that might require a time judgment could be avoided by suggesting that time perception depends on a memory trace (Staddon, 2005). There is, in general, no reset problem here because all world events encoded by the cognitive system automatically result in representations that are governed by the same trace dynamics. That said, most activation-trace systems posit a specialist timing mechanism that is only recruited when timing is required (e.g., Reutimann et al., 2004; Staddon & Higa, 1999) and models of this type can only address prospective timing. The Temporal Context Model (Shankar & Howard, 2010), developed from a model of episodic memory, can potentially perform both retrospective and prospective timing having been. This model is very complex but its authors claim it demonstrates the scalar property, can account for timing effects in simple Pavlovian conditioning (Drew et al., 2005) and for temporal responding found in some hippocampal cells (Pastalkova et al., 2008). It is an overlooked feature of episodic memory that it deals with temporal stimuli, both when recalling the past and making predictions about the future (see MacDonald et al., 2014; Schacter et al., 2007). To our knowledge, Shankar and Howard (2010) is the first attempt to use features of memory directly as a mechanism for interval timing. GAMIT has similar motivations but is much simpler than TCM. In addition, GAMIT accounts for the surprising effects of cognitive load.

1.4   Cognitive Load in Existing Models

Our estimates of time passing can be affected by whether or not we are actively attending to the passage of time and by the amount of additional cognitive load we face. Block et al. (2010) analyzed the results from over one hundred interval-timing studies and summarized their results in the graph shown in Fig. 1. They found a striking interaction between the type of time judgment requested and cognitive load. High cognitive load *increases* your estimates in the case of retrospective timing, whereas high cognitive load *decreases* your estimates in the case of prospective timing. This strong interaction is puzzling for two reasons. First, as discussed above, the mere fact that there is a difference between prospective and retrospective time is a challenge to clock and timestamp models. There is no a priori reason to expect a difference

between these two conditions. Secondly, the interaction with cognitive load suggests that cognitive load is not just an additive factor (e.g., damping responses across the board), but somehow has opposite effects on timing in retrospective and prospective contexts. This is a challenge for all existing models of interval timing.
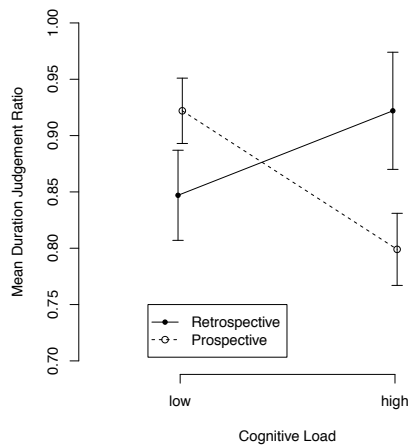

Figure 1. The effects of cognitive load on interval timing based on a meta-analysis of 82 prospective and 31 retrospective tasks. Duration judgment ratio is the ratio between subjective estimates of time and the actual objective time that has passed. Error bars show standard errors. Reproduced from Block et al. (2010). Permission pending.

The original SET (Gibbon, 1977; Gibbon et al., 1984) was a model of animal timing and so did not consider attention effects. It has subsequently been argued that this scalar timing model can accommodate attentional effects if attention is allowed to affect the switch mechanism controlling the flow of pulses into the accumulator (Lejeune, 1998, 2000; Meck, 1984; Penney et al., 1996). Under high cognitive load the switch would 'flicker' on and off letting few ticks accumulate and leading to *decreased* prospective estimates. The attentional gate hypothesis (AGH - Zakay & Block, 1995; Block & Zakay, 1996) is an alternative approach that has functionally a very similar outcome. In this case, a gating mechanism added next to the switch that attenuates the flow of ticks from the pacemaker in proportion to the amount of attention allocated to the timing aspect of the task. Alternatively, Taatgen, van Rijn and Anderson (2007) argue that this attentional effect on prospective timing would arise as a result of resource sharing if a clock module is integrated into a wider and fully developed cognitive architecture. This model makes the important point that timing can be affected by the allocation of limited attentional resources as a result of other processes operating in parallel elsewhere within the cognitive system (see also Buhusi & Meck, 2009).

Less work has been done on modeling the effect of attention on time perception in other frameworks. Many models do not consider attentional effects, in particular, the effect of cognitive load on attention (Reutimann et al., 2004; Shankar & Howard, 2010; Staddon & Higa, 1999). Matell and Meck (2004) propose that attention might modulate clock speed directly. If decreased attention to timing causes the organism's internal clock to beat slower, then it will tend to underestimate the length of intervals. This idea is developed further in the time-sharing model developed by Buhusi and Meck (2006, 2009). Working memory, timing and attention all depend on dopaminergic pathways (Cools & D'Esposito, 2011; Lake & Meck, 2013; Lustig & Meck, 2005; Meck et al., 2012a). The changes observed in interval timing estimates following pharmacological interventions that modulate clock speed (Coull et al., 2011; Meck, 1996) have been modeled by letting dopamine levels affect oscillator frequency (e.g., Allman & Meck, 2012; Buhusi & Oprisan, 2013; Oprisan & Buhusi, 2011). Nevertheless, none of these models can account for the *increase* in retrospective estimates under high cognitive load.

Far fewer models attempt to explain retrospective timing, in part because retrospective timing does not have an equivalent in animal behavior. A common theme behind all approaches to retrospective timing is that intervals are estimated by reconstructing a sequence of remembered events. Cognitive load could affect this by changing the memorability or numerosity of events. For example, in the contextual change model (Block, 2003) and the segmentation model (Poynter, 1983) information processing demands *per se* will not affect remembered duration but manipulations that involve greater interruption to the task or more switching between different kinds of activity will. This is consistent with findings of Block et al. (2010) for retrospective timing but cannot simultaneously account for the interaction.

In summary, existing models of interval-time judgments rely either on a central clock and accumulator, a time stamp mechanism, or a transient activation trace to explain participants' judgments of time intervals. However, it is currently uncertain which (if any) of these models offers the most parsimonious account of scalar timing. First of all, in order to do so, they must posit secondary mechanisms. In addition, it is unclear how the radical differences between retrospective and prospective time judgments would emerge from these timing models, nor how this difference would be modulated by cognitive load.

In what follows, we will suggest that time is measured by estimating the extent to which and the rate at which a

fading-Gaussian activation trace has decayed. The scalar property arises because this activation follows a broadly logarithmic decay. The differences that appear under cognitive load, between retrospective and prospective estimates, are due to the interaction of two factors: activation decay and sampling rate. In retrospective timing the cognitive system does not know ahead of time that it will be asked to make a time judgment, the trace is "sampled" only once at the end of the interval. In prospective timing, the cognitive system randomly samples the trace throughout the interval. Higher cognitive load will have two antagonistic effects. In both cases (prospective and retrospective) it will cause the activation trace to decay faster leading to apparently longer estimates. However, in the prospective case, higher cognitive load will also lead to less frequent sampling. This leads to a perception that time is actually passing more quickly. The system makes an over-correction giving an estimate that is actually shorter than before. In what follows we first begin by describing the fading-Gaussian mechanism itself and then go on to show how the scalar property and the Block et al. (2010) interaction arise within this model.

2   GAMIT: A Fading-Gaussian Activation-Trace Model of Interval Timing

In this section we describe GAMIT, our model of interval timing (The MATLAB® code is available on request.). The fading-Gaussian activation model is built on the assumption that our sense of time is learned through our experience of changes in the world around us. Small changes in the activation trace mean little time has passed; large changes mean a lot of time has passed. These changes allow us to interpret a fading-Gaussian activation trace associated with a particular event as the passage of time.

A number of further assumptions underlie the model. The first assumption is that the original activation trace generated by an event fades over time in a statistically predictable manner (this is discussed further in Section 3.2). The second assumption is that this decay is affected by cognitive load. These assumptions alone account for retrospective timing. For prospective timing we further assume that the cognitive system is sensitive to not only this pattern of fading activation, but also to how this activation pattern changes over time. Since we are aware that we will be required to make a time judgment with respect to a particular event, we "sample" the decaying trace during the interval. Finally, this sampling occurs less

frequently when the system is under cognitive load. The crucial cue for prospective interval-time estimation will not only be the final level of the activation trace generated by the initial event, but also *the rate at which* the activation of this trace is changing.

To implement the GAMIT model, we begin with a cluster of cortical columns. The activation in the central column corresponds to an event in the world that is registered in memory. This event may be the passage of a car on a street, the dropping of a dish, or the sound of a beep in a laboratory time perception experiment. Activation then spreads across the cortical columns as follows.

If we designate the activation of the ith column at time step t by $A_i(t)$, its activation at time $t+1$ is determined by the following equation:

$$A_i(t+1) = \alpha A_i(t) + \beta(A_{i-1}(t) + A_{i+1}(t)) + \xi$$

where α is the fraction of activation that remains in column i on each time step (i.e., α = 1-leakage); β is the fraction of activation spread from each immediate neighbor of i on each time step; ξ is a noise parameter. The values of α and β must be chosen so that the total activity of the system neither rapidly decreases to zero nor increases exponentially. Unless otherwise stated, we used values of α = 0.7, β = 0.14952 and ξ = 0.000025.

We start by assuming that the default initial activation of the central column (i.e., i = 0), is 1 at t = 0, which produces a narrow "activation spike" on that column. On each time step, activation in the central column decreases and spreads to neighboring columns. This is illustrated by the Gaussians in Fig. 2.
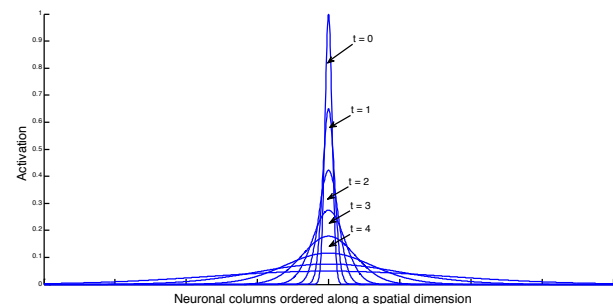

Figure 2. The activation of the initial Gaussian above the central column fades and spreads with time.

There is ample neurobiological evidence for this type of spreading-activation mechanism (e.g., Amari, 1977, 1980; Capaday et al., 2011; Grinvald et al., 1994; Grossberg, 1980; Herman et al., 1993; Koch & Segev, 1998). It is associated with memory-like processes in the cortex and provides a time-dependent signal with suitable statistical properties

for interval timing. In particular, when activation spreads, the amount of change in the overall activation of the trace allows us to estimate the length of the time interval since the onset of the stimulus event.

In the simulations reported below, we argue that the cognitive system is sensitive to both the height of the fading Gaussian at time t (designated by $H(t)$) and the total activation of the fading Gaussian at time t (i.e., the sum of the activations of all columns to which activation has spread, a value which we will call $A(t)$). The first variable, $H(t)$, is a measure of how salient the Gaussian is with respect to background noise. The second variable, $A(t)$, is a measure of how much the overall signal has faded. In our model, the activation values on which time estimates are based is the sum of these two values: $S(t) = H(t) + A(t)$. Not only is this quantity easy to compute, it provides a stable estimate of the long-term average of the underlying stochastic diffusion process. Henceforth, when we refer to activation as it relates to time estimation in our model, we will mean this combined spreading-activation value, $S(t)$.

It is this value $S(t)$ that gives the estimate for how much time has passed since the initial event. In the retrospective case this is straightforward. At the end of the interval, the cognitive system makes a single estimate of how far the activation has decayed and compares this value to a reference curve built up from a lifetime of experience. Although the underlying process may be stochastic it is statistically predictable. Therefore, a certain amount of decay corresponds to a certain amount of time and the greater the decay the longer the interval. The noise in the system and the error when estimating the decay leads to the error in the estimates. We will show that it has the scalar property. In the prospective case additional information is available, as the cognitive system will have made a series of 'sampling' estimates during the interval.

3 Time Estimates and the Scalar Property

In this section we focus on the scalar property (Gibbon, 1977; Gibbon et al., 1984) of interval-time estimation – namely, that time-estimation error is a linear function of the amount of time estimated. We will begin by discussing certain mathematical and statistical considerations that buttress the plausibility of a spreading-activation model. We then discuss how time estimates are generated in the GAMIT model and the potential sources of error this introduces. We conclude this section by presenting a simulation that shows that our fading-Gaussian model

does, indeed, implement the scalar property for interval-time estimation.

Hass and Hermann (2012) show mathematically that the neural process underlying time perception must be based on a measure of growing variance in the system. We suggest that timing is based on statistical estimates of a stochastic process whose long-term expected value is equivalent to a spreading of activation. Such processes are common in nature. For example, variance grows linearly with time in both diffusion processes (Einstein, 1905) and maximum entropy distributions (Jaynes, 1957). Our model is based on the idea that estimating the extent of the activation spread lets us estimate the length of the interval. This is an inherently noisy and stochastic process but we know from work in perceptual psychophysics (Ernst & Banks, 2002; Knill & Richards, 1996) and computational neuroscience (Paninski et al., 2004; Pouget et al., 2003) that the brain can make near optimal use of inherently noisy signals.

The central idea of the GAMIT model is that intervals are estimated by determining how much a stochastic activation trace has faded and comparing this to a reference curve (white line in Figure 3) built from a lifetime of experience (Fig. 3).
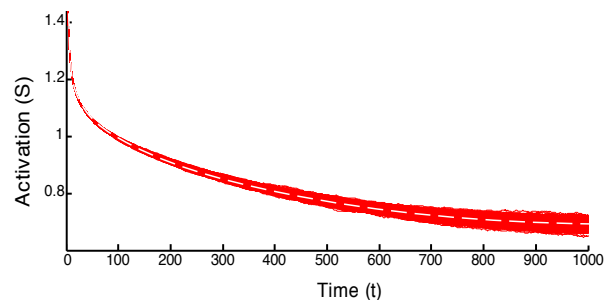

Figure 3. One thousand individual activation-decay curves (in red), running for 1000 time steps, are averaged to create the Reference Curve (white).

This process has several steps, each of which introduces a source of error. Firstly, each individual curve will evolve somewhat differently leading to a different activation level at the target time. When compared to the reference curve, these activations lead to different estimates for the same target interval (Fig. 4).

Secondly, the activation of a single curve cannot be known precisely, leading to an additional sampling error, illustrated by the error distribution on the vertical axis of Fig. 4. The reference curve will be the average of many previous individual curves and would, therefore, have an associated uncertainty. However, in the present paper we do not address how the lifetime curve might be generated

(but see Addyman et al., 2011), and, therefore, we have not included this in the present discussion.

The crucial scalar property of interval-time estimation says that time-estimation error will increase linearly with the time to be estimated. To test the scalar property, we assume that the current activation-decay curve is decaying as shown by the red curve in Fig. 4. (Note this curve differs slightly from the (blue) Reference activation-decay curve, the latter being the average of a lifetime of "current activation-decay curves.")
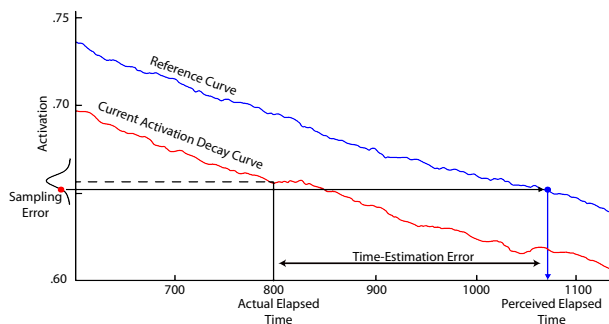

Figure 4. Estimating the activation of a single curve and comparing it to the lifetime (or "typical") Reference activation-decay curve introduces two sources of error. One is associated with the evolution of the decay curve and one comes from the sampling process.

At a particular time (here t = 800 time units), a time judgment is requested. As shown in Fig. 4, an "actual elapsed time" (i.e., t = 800) corresponds to an activation level of the current activation curve of S = 0.652. There is a Gaussian sampling error distribution around S with σ = 0.05. We sample from this distribution and get a value of S = 0.656. The corresponding time value, t, for S = 0.656 is "read off" the Reference activation-decay curve. This value, the "perceived elapsed time" is t = 1075. Thus, the time-estimation error (E) is the difference between the perceived elapsed time and the actual elapsed time, which, in this case, is 275 time units.

We considered all time values, t, between t = 1 and t = 750 and calculated the time-estimation error, E, as calculated above, for each of these time values. Averaged over 250 runs of the program, we obtain a linear fit (E = 0.23t) to the data with an r of 0.99 (Fig. 5). The activation decay in GAMIT satisfies Hass and Herrmann's (2012) variance requirements for scalar growth in time estimation error. In other words, the scalar property falls directly out of the core neurocomputational principles of GAMIT, and does not require the positing of any further secondary mechanisms.

4 Modeling Retrospective and Prospective Time Judgments under Cognitive Load

A key feature of our model is that the activation decay and the sampling rate of the activation trace are differentially affected by cognitive loads. To begin with, we assume that greater cognitive load causes more rapid decay of the trace activation, due primarily to global inhibition from other concurrent tasks that require encoding and storing information in memory (Fig. 6). This alone allows us to explain differences in retrospective time estimation (Fig. 7). However, in the case of prospective time estimation, when it is known ahead of time that a time judgment will be required, we further propose that the state of the activation trace will be repeatedly *sampled.* Sampling can be thought of in our model as "attentional saccades" to the event trace. Just as visual saccades involve a switch of visual attention, we suggest that mental saccades involve a switch of focus of attention to the trace. Just as the rate of visual saccading is interfered with by increased cognitive load (Halliday & Carpenter, 2010; Stuyven et al., 2000), we suggest that the same is true of attentional saccading. In other words, attentional resources are limited and must be distributed among the currently active tasks in working memory. Similarly, as cognitive load increases and more tasks must be processed with limited attentional resources, fewer resources (attentional saccades) are allocated to attending to the activation trace of the event about which a time judgment will be made (see also the time-sharing hypothesis of Buhusi & Meck, 2006)[1].
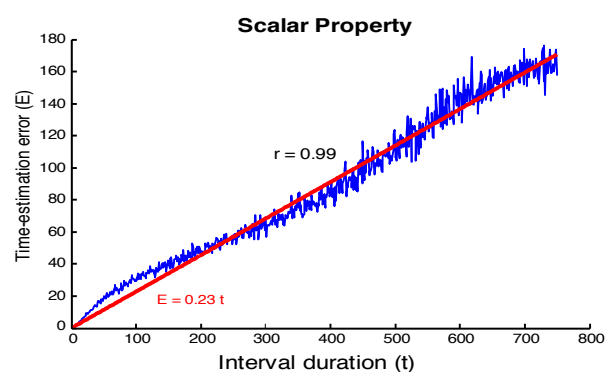

Figure 5. Time-estimation error (E) grows linearly with the duration of the time interval estimated. Results were averaged over 250 runs.

---

[1] One helpful analogy is to compare this to time-sharing in a computer CPU. Tasks, both in the computer and in the cognitive system, must be processed simultaneously. As the load on the CPU increases, fewer time slices can be allocated to each task that needs to be performed.

Over time the cognitive system learns a very simple association – namely, the more the activation of a trace has changed since it was last sampled, the more time that has elapsed. In other words, small changes in activation correspond to small changes in time; large changes in activation correspond to large changes in time. This is one of the key insights to understanding our fading-Gaussian model of interval-time estimation.

4.1   Retrospective Time Estimation

As we said at the start of Section 3 above, our explanation of both retrospective and prospective time estimation relies on the assumption that we have, over time, learned a typical or "average" activation-decay curve for activation traces of events and that this curve serves as a reference curve for time estimation (Fig. 3). This curve can be thought of as a running average of a lifetime of experiences. We further assume that under higher-than-normal cognitive load, there will be somewhat less activation, $S(t)$, in the trace because it is being actively inhibited by the other concurrent tasks in working memory ("High cognitive load" curve, Fig. 6). This means that the activation curve falls somewhat more rapidly than normal under high cognitive load.
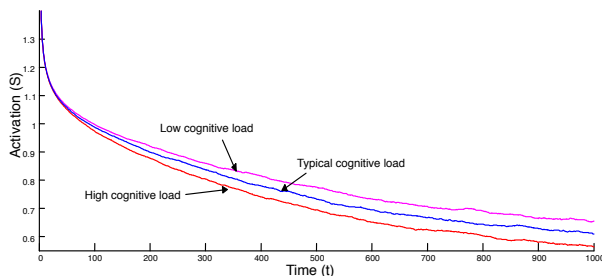

Figure 6. The typical activation-decay curve, $S(t)$, learned by experience is shown in blue. Under high cognitive load, activation falls off more quickly than under typical cognitive load. Under (very) low cognitive load (curve in pink) activation falls off more slowly than typical load conditions.

In retrospective time judgment there is no prior announcement that a time judgment will have to be made. This means that there is *no sampling of the activation trace* prior to the moment when the time judgment must be made, and corresponds to what Zakay and Block (2004) refer to as "*remembered duration,*" since there is no on-going experience of the time interval between the moment of the stimulus event and the time when a time judgment must be made. Thus, the only cue to the amount of time that has passed is the total activation of the memory trace.

4.1.1   Retrospective time estimation under high cognitive load

Under high cognitive load we assume that there are a greater-than-normal number of distracting tasks interfering with the activation trace of the stimulus event. This means that the activation trace will "squash and spread" more rapidly than normal, meaning that activation will fall more rapidly than normal (red curve in Fig. 7).

Suppose that under high cognitive load, at the moment of a time judgment (e.g., t = 600), the activation value is approximately 0.66. But, in memory, there is only a stored representation of the activation curve under typical cognitive load. Based on this reference activation-decay curve (i.e., the blue curve in Figs. 3, 4 and 6), an activation level of 0.66 occurs, not at t = 600, but rather, at t = 710. Thus, when asked for a (retrospective) time judgment at t = 600, we reply 710. In other words, we overestimate the amount of time that has passed under high cognitive load. This concurs with Block et al.'s (2010) finding.
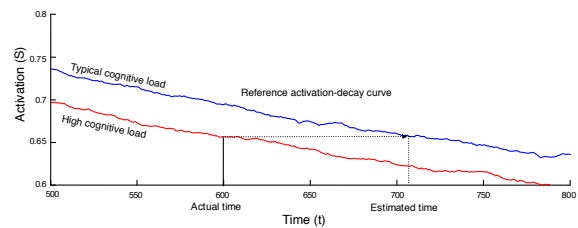

Figure 7. A run of GAMIT retrospectively estimating time under high cognitive load.

4.2   Prospective Time Estimation

In prospective time estimation the participant knows ahead of time that a time judgment will be required about a particular stimulus event at some point in the future. This implies an on-going monitoring of the activation trace, a process that engenders what Zakay and Block (2004) refer to as "*experiencing time*". In other words, the activation trace associated with that event will be sampled more or less frequently until the time estimation has been made. The frequency of this sampling -- what we refer to as "attentional saccading" — depends on cognitive load. In GAMIT, this attentional sampling is what provides information about the rate-of-change of total activation of the trace. Crucially, we assume that there is a "typical" sampling profile, not necessarily uniform, that defines when and how often sampling of the activation trace occurs under typical cognitive load in the context of prospective time estimation. We assume that, under high cognitive load, there is approximately the same sampling distribution as under typical cognitive load, the only

difference being that this "attentional saccading" to the activation trace occurs less frequently. The lower sampling frequency is due to the fact that sampling requires cognitive resources and that, under high cognitive load, some of those limited resources are being diverted to additional mental activities.

We argue that our perception of the passage of time is intimately related to the rate of this sampling of the activation trace. To understand this, an analogy is helpful. Consider some event that unfolds over approximately 30 seconds, say, a woman walking down a street. Now, assume that we have two cameras at the scene: the first films this event at a "typical" speed of 20 frames/sec.; the second camera must film, not only this event, but simultaneously, some other event nearby. The latter camera shoots one frame of the walking woman, followed by one frame of the other event, then one frame of the woman, etc. The film of the woman walking is put together in the cutting room and, of course, contains only 10 frames/second of the woman walking. We show both films to people and ask them in which of the two films time seems to pass most quickly. They will reply that the flow of time is faster in the latter film (Eagleman, 2004). We suggest that the reason is because between each image in the 20 frames/sec case very little changes, whereas there is a much greater change between each image in the 10 frames/sec case. Thus, since we have learned over the course of our lifetime that, on average, small changes in the world correspond to small changes in time and large changes in the world correspond to large changes in time, we perceive time as passing more quickly in the latter film.

We suggest that sampling of the activation trace for the stimulus event corresponds precisely to the above analogy. Just as the first camera is sampling the event consisting of the women walking at the "typical" rate of one frame every 0.05 seconds and the second camera, because it must also film another event simultaneously, is only sampling the woman walking at the rate of one frame every 0.1 seconds, the sampling of the activation trace associated with the event about which a time judgment is to be made occurs at different frequencies depending on cognitive load. Just as viewers' perception of the passage of time in a film varies depending on the speed at which the event is filmed, we suggest that the same holds for prospective time estimation: as the sampling rate slows down, the subjective experience of the passage of time is compressed (i.e., speeds up).

4.2.1 Prospective time estimation under high cognitive load

We claim that prospective time estimations rely on, not only the amount of activation decay, but also the *rate at which the activation is decaying, as measured by activation change between attentional saccades.* To calculate the approximate rate of change of the decreasing activation function, a small number of recent activation changes between successive samplings of the activation curve are kept in memory. (In our simulations, the 7 most recent activation changes were stored.) The average of these values provides an estimate of the rate at which the activation curve is falling. Over time, the cognitive system under normal cognitive load learns how much the activation trace associated with an event typically changes between attentional saccades. This value, stored in memory, we call $\Delta_{\text{TYPICAL-LOAD}}$. Under high cognitive load we sample the curve less often because some of the resources devoted to attention saccading are devoted to the other tasks (Fig. 8). Thus, the amount of activation change between each attentional saccade, $\Delta_{\text{HIGH-LOAD}}$, is greater. (For low cognitive load, $\Delta_{\text{LOW-LOAD}}$ is calculated in the same way, except there is more than average sampling of the activation trace.) This is the intuition behind the definition of a time-compression (or time-dilation) factor $\Phi = \dfrac{\Delta_{TYPICAL-LOAD}}{\Delta_{CURRENT-LOAD}}$ for prospective time judgments. A prospective time estimate, P, is the retrospective time estimate, R, adjusted by the multiplicative time-compression/dilation factor, Φ. In other words, P = RΦ.
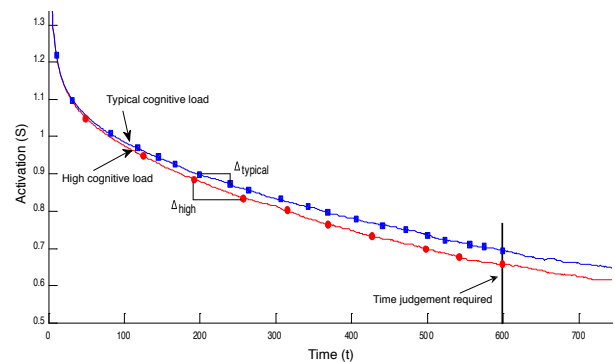

Figure 8. Prospective time estimation under high cognitive load. The blue curve shows the evolution of activation under typical cognitive load. For prospective time estimation under typical cognitive load the frequency of sampling of this curve is shown by the blue square markers.

In the example in Fig. 8, under typical cognitive load, Δtypical-load is approximately 0.214. Because under high cognitive load sampling occurs half as often as in the typical-load condition, the activation changes between samplings is greater than in the typical-load condition and, as a result, Δhigh-load is approximately 0.321. Thus,

we can calculate Φ to be $0.214/0.321 = 0.667$. This means that under high cognitive load we would experience a time compression of 0.667 of the time estimate we would have had under retrospective time estimation.

Recall that under retrospective timing conditions, where there is no sampling (Fig. 7), we overestimated time. At $t = 600$ (the real interval time), we estimated it to be $t_R = 710$. Now, to calculate the prospective time estimate, we start with t_R, the retrospective time estimate and adjust this value by the multiplicative time-compression/time-dilation factor, Φ (i.e., P = RΦ). In other words, the prospective time estimate, t_P, of the actual time of $t = 600$ would be: $t_P = 710 * 0.667 = 467$. In other words, in the prospective timing condition, we perceive $t = 600$ to be $t = 467$.

We calculate Φ in an identical fashion for the low cognitive load condition, in which there is more than average sampling of the activation trace.

4.3   Simulation of Retrospective and Prospective Time Estimation Together

In order to simulate changes in cognitive load in GAMIT, we varied the amount of activation spread to neighboring columns, thereby causing the activation curve to fall more or less rapidly[2]. To simulate high cognitive load conditions we decreased the value of β to 0.14946. For lower-than-typical cognitive load, the value of β was increased to 0.14955. The other two parameters remained unchanged. We ran the program 20 times under both high and low load conditions and averaged the results of these simulations.

Figure 9. Performance of the GAMIT model on prospective/retrospective time judgments under high and low cognitive load. Results averaged over 20 runs of the program. SD error bars are also plotted.

We assume that a time judgment for both the retrospective and prospective conditions would be made at $t = 600$. There was no sampling of the activation trace in the retrospective time-estimation condition. In the prospective time-estimation condition under typical cognitive load, we assumed that by $t = 600$ there would have been 20 samplings of the activation trace. Under high cognitive load, this was decreased by 50% to 10. Under lower-than-normal cognitive load, we increased this amount of sampling by 10%. This asymmetry is intended to reflect the fact that cognitive load can only be decreased slightly with respect to the typical cognitive load condition, whereas it can be increased substantially. The results are shown in Fig. 9. As in Block et al. (2010), an ANOVA showed that there is no significant main effect of either Cognitive Load or Retrospective-Prospective time estimation. However, again, as in Block et al., there was a highly significant interaction between the two main variables ($F_{(1, 76)} = 19.3$, $p < .0001$, $\eta^2 = 0.2$). In other words, GAMIT qualitatively reproduces the interaction between cognitive load and mean time-judgment duration ratio reported in Block et al. (2010).

4.4   Summary of Prospective and Retrospective Time Estimation in GAMIT: P = RΦ

A central claim of the GAMIT model is that Prospective time estimation (P) is equivalent to Retrospective time estimation (R) adjusted by a time-compression/time-dilation factor (Φ) that is a function of cognitive load, i.e. P = RΦ.

---

[2] People tend to underestimate (or over-produce) objective durations leading to mean judgment ratios less than 1. GAMIT includes a bias parameter to model variations in risk judgments. In all current simulations, the parameter does not play a functional role, and is fixed at 0.87 which is the global average of those values reported in Block et al. (2010).
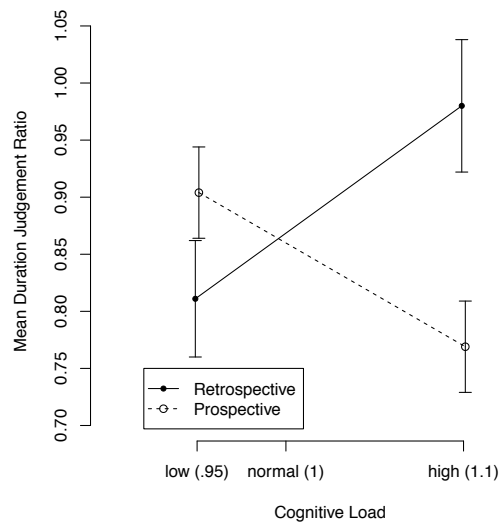
In retrospective time estimation, there is no sampling of the activation trace since participants do not know ahead of time that a time judgment will be required. At the moment when they are asked to perform a time judgment about how long ago a particular event happened, they rely on the value of the total activation of the activation trace at that moment in time. However, in prospective time estimation participants know ahead of time that they will be asked for a time judgment about a particular event. They, therefore, devote attentional resources specifically to the event in question. The amount of attention paid to the activation trace determines how much their perception of the passage of time with respect to that event is speeded up or slowed down. Moreover, the amount of attention paid to the activation trace varies as a function of cognitive load. We suggest that in prospective time estimations we adjust our baseline memory of the event based on this perceived speeding up or slowing down of time. In short, prospective time estimation is simply retrospective time estimation "adjusted" by the perceived rate of change of the passage of time and this perception of how fast time is passing depends on cognitive load.

Finally, note that the scalar property demonstrated for retrospective time judgments (Section 3) also holds for prospective time estimation. Indeed, prospective time estimation is determined by means of a multiplicative factor of the activation curve for a retrospective time judgment (i.e., $P = R\Phi$). Multiplying this linear relationship by a constant will not affect the linearity of the relationship. Thus, for both retrospective and prospective time estimation, we have a linear relationship between the actual amount of time (t) and the time-estimation error (E) of t.

5   General Discussion

We have described a "fading-Gaussian" model of interval-time estimation, GAMIT, which is based on the classic equation of spreading activation as an approximation to the underlying stochastic processes involved in the spread of information in a distributed cognitive system. The model naturally produces scale-invariant time estimates and captures participants' prospective and retrospective time estimates, in addition to capturing the interaction of these estimates with changes in cognitive load. All three of these phenomena remain important challenges for existing accumulator-based, oscillator-based, and memory-based models. GAMIT is based on the processes of activation decay and activation-trace sampling, both of which are ubiquitous within the brain and cognitive system. In addition, it does not suffer from the reset

problem that plagues accumulator and oscillator-based models of interval timing. Finally, because a core component of the timing mechanisms (the reference trace) is based on a lifetime of past experiences of observing change in the world, GAMIT provides an approach to timing in which information about the passage of time is embedded in the world rather than constructed in an abstract cognitive module.

The model takes seriously recent analyses of the scalar property of time showing that mechanisms underlying time estimation should be noisy, stochastic and based on the spread of information (Buhusi & Oprisan, 2013; Hass & Hermann, 2012). As far as we are aware, this is the first model that is actually built directly from this neural information-processing assumption showing how very common, cognitive-level phenomena are rooted in the dynamics of neural information processing. As such, the model constitutes a strong bridge between the cognitive and neural-timing communities that have traditionally worked separately.

A second contribution is that the GAMIT model provides a unified account of retrospective and prospective time judgments. These have traditionally been studied as separate phenomena (Block et al., 2010; Zakay & Block, 2004; but see Brown & Stubbs, 1992) largely because (1) it is hard for timestamp and accumulator-based models to provide any explanation of retrospective timing, and (2) they are differentially affected by attention and cognitive load, suggesting that different underlying processes may be at play. In contrast to these views, our model suggests that the same underlying mechanisms (namely, decay and sampling) operate in both contexts. What varies is the amount of sampling that occurs between the two contexts: no sampling occurs on retrospective time estimates, whereas repeated sampling occurs in the prospective time estimates. The effect of cognitive load on interval-time perception is explained by how cognitive load affects the rate of sampling of the activation trace. Thus, the GAMIT model provides a parsimonious account of interval-time judgments.

In addition, we suggest that the estimation of time is tuned across the lifetime. The reference curve that is used to estimate the passage of time by sampling the decaying trace is the result of multiple experiences of the passage of time. Although the current model does not address this issue directly, it also suggests a possible mechanism for explaining developmental differences in interval timing (see Addyman et al., 2011 for further discussion)

GAMIT bears some resemblance to other trace-based models of item memory (e.g., Barrouillet et al., 2004; Brown et al., 2000; Lépine et al., 2005) in that they posit

decay mechanisms and sampling of decaying traces. However, these models are concerned with the retrieval of items from memory, and possibly order effects in memory retrieval and, in general, return rank order information of elements presented in a sequence. However, they do not produce duration estimates. Nor, of course, do they address the scalar property of time or the effects of cognitive load on time judgments. The fact that there is some processing similarity between our model and these memory models underscores the generality and parsimony of our approach.

5.1 Model Predictions

A number of predictions fall out of this differential-sampling hypothesis. For example, if cognitive load is increased to the point that no sampling of the activation trace is possible, then a dramatic flip to the retrospective time estimate should occur. Further, there is a ceiling effect in the direction of low cognitive load compared to typical cognitive load. Whereas it is always possible to increase cognitive load, it can only be decreased to a limited extent below typical cognitive load. This is the reason that we will never get very large values of Φ in low cognitive-load situations, which would imply that in this condition we do not vastly overestimate the passage of time.

Because the reference trace reflects a lifetime of accumulated experience, individuals who have lived for a prolonged period in higher cognitive-load environments (e.g., air-traffic controllers) will have a different reference curve from that of individuals having experienced a lesser cognitive-load on average. Or alternately, they will have two separate context-dependent reference curves, one for their high cognitive-load job, the other for their normal cognitive-load existence. The model, therefore, predicts that these individuals will experience the passage of time in a high cognitive load situation differently than the rest of us. In short, they will be able to significantly more accurate in their time-estimates than people who have no high cognitive-load reference curve. For example, their prospective time estimates should be less affected by time dilation than for those people who have experienced relatively low cognitive loads during their lifetimes.

A further prediction is that younger children should show greater between-participant variability than older children and adults in their estimates of time. As children get older, their sample of experiences increases, and their average reference trace will converge to a mean representation comparable to the rest of the participants.

5.2 Model Limitations

In all models of time estimation based on a fading activation trace, including GAMIT, there is a potential issue of confusing stimulus recency with stimulus repetition (McCormack & Russell, 1997; McLaren, 1994; Shaw & Aggleton, 1994). The idea is that, upon seeing the initial stimulus, the corresponding activation would assume its maximum value. Thereafter, that activation would decay and the amount of decay would serve as the basis for time estimation. But what if the initial stimulus were presented again before the time estimate was made? Would this not return the activation trace associated with the stimulus to its maximum value, thereby preventing us from making a time judgment about how long ago the stimulus had first been presented? The answer, we believe, is provided by the Greek philosopher Heraclitus, famous for his claim that "No one ever steps in the same river twice." The point is that the second presentation of the stimulus is *not* the same as the first presentation for the simple reason that part of the internal representation of the second stimulus *contains the first presentation of the stimulus*. In other words, while it might have some influence on the initial activation trace, the second presentation of the stimulus would be perceived as a different event (unless, of course, short-term memory deficits meant that there was no recollection of the first presentation) and would be associated with a different trace.

In its current form, the GAMIT model captures the pattern of meta-data reported in Block et al. (2010). It does not, however, model data from individual experiments. However, although input-output representations would need to vary according to the task constraints, the core timing mechanisms would remain the same for all of these timing studies.

How might this model work with filled intervals? We speculate that the stimulus event is perceived as a "change in the state of the world". So, in a laboratory setting, a very short beep would represent a change in the world, (actually two, virtually simultaneous changes, from no sound to the beep and back to no sound). In the case of a prolonged beep (i.e., a filled interval), the first change in the world would occur at the onset of the beep (i.e., no sound to sound) and the second would occur when the beep stopped (i.e., sound to no sound). In both cases, it is the change in the world that triggers the activation trace, rather than the sound itself.

Two central assumptions we make when considering the scalar property of time is that the underlying information processing is stochastic and that the activation trace decays logarithmically. However, for

implementational reasons, we have used a continuous diffusion process as a time-averaged approximation of this stochastic spread of information. This is analogous to the use of continuous activation functions to approximate the stochastic updating of neurons in large neural networks (Hertz et al., 1991). These continuous approximations can be shown to be equivalent to the long-term time-averaged profile of the stochastic process. Nevertheless, in future implementations of the model the actual stochastic processes should be constructed.

6   Conclusion

In summary, we have described a parsimonious model of interval timing that is based on ubiquitous neural and cognitive processes. It provides a unifying account of retrospective and prospective timing, captures the modulating effects of cognitive load on both prospective and retrospective timing, and in contrast to other models, intrinsically captures the scalar property of time judgments.

7   Authors' Note

8   References

Addyman, C., French, R. M., Mareschal, D., & Thomas, E. (2011). Learning to perceive time: A connectionist, memory-decay model of the development of interval timing in infants. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proc. 33rd Annu. Conf. Cogn. Sci. Soc.* (pp. 354-359). Austin, TX: Cognitive Science Society.

Allman, M. J., & Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain, 135,* 656-677.

Allman, M. J., Teki, S. Griffiths, T. D., & Meck, W. H. (2014). Properties of the internal clock: First- and second-order principles of subjective time. *Annu. Rev. Psychol., 65,* 743-771.

Amari, S. I. (1977). Neural theory of association and concept formation. *Biol. Cybern., 26,* 175-185.

Amari, S. I. (1980). Topographic organization of nerve fields. *Bull. Math. Biol., 42,* 339-364.

Bangert, A. S., Reuter-Lorenz, P. A., & Seidler, R. D. (2011). Dissecting the clock: Understanding the mechanisms of timing across tasks and temporal intervals. *Acta Psychol. (Amst.), 136,* 20-34.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *J. Exp. Psychol. Gen., 133,* 83-100.

Block, R. A. (2003). Psychological timing without a timer: The roles of attention and memory. In H. Helfrich (Ed.), *Time and mind II: Information-processing perspectives* (pp. 41-59). Seattle, WA: Hogrefe & Huber.

Block, R. A., & Zakay, D. (1996). Models of psychological time revisited. In H. Helfrich (Ed.), *Time and mind* (pp. 171-195). Kirkland, WA: Hogrefe & Huber.

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychol. (Amst.), 134,* 330-343.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychol. Rev., 107,* 127-181.

Brown, S. W., & Stubbs, D. A. (1992). Attention and interference in prospective and retrospective timing. *Perception, 21,* 545-557.

Buhusi, C. V., Aziz, D., Winslow, D., Carter, R. E., Swearingen, J. E., & Buhusi, M. C. (2009). Interval timing accuracy and scalar timing in C57BL/6 mice. *Behav. Neurosci., 123,* 1102-1113.

Buhusi, C. V. & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci., 6,* 755-765.

Buhusi, C. V., & Meck, W. H. (2006). Interval timing with gaps and distractors: Evaluation of the ambiguity, switch and time-sharing hypotheses. *J. Exp. Psychol. Anim. Behav. Process., 32,* 329-338.

Buhusi, C. V., & Meck, W. H. (2009). Relative time sharing: New findings and an extension of the resource allocation model of temporal processing. *Philos. Trans. R. Soc. B., 364,* 1875-1885.

Buhusi, C. V., & Oprisan, S. A. (2013). Time-scale invariance as an emergent property in a perceptron with realistic, noisy neurons. *Behav. Process., 95,* 60-70.

Burr, D., Della Rocca, E., & Morrone, M. C. (2013). Contextual effects in interval-duration judgements in vision, audition and touch. *Exp. Brain Res., 230,* 87-98.

Capaday, C., van Vreeswijk, C., Ethier, C., Ferkinghoff-Borg, J., & Weber, D. (2011). Neural mechanism of activity spread in the cat motor cortex and its relation to the intrinsic connectivity. *J. Physiol., 589,* 2515-2528.

Church, R., & Broadbent, H. (1990). Alternative representations of time, number and rate. *Cognition, 37,* 55-81.

Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *J. Exp. Psychol. Anim. Behav. Process.*, *20*, 135-155.

Cools, R., & D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biol. Psychiatry, 69* (12):e113-125.

Coull, J. T., Cheng, R. K., & Meck, W. H. (2011). Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology, 36*, 3-25.

Drew, M. R., Couvillon, P. A., Zupan, B., Cooke, A., & Balsam, P. D. (2005). Temporal control of conditioned responding in goldfish. *J. Exp. Psychol. Anim. Behav. Process.*, *31*, 31-39.

Eagleman, D. M. (2004). Time perception is distorted during slow motion sequences in movies. *J. Vision, 4*, 491.

Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik, 17*, 549–560.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*, 429-433.

Fortin, C. (1999). Short-term memory in time interval production. *Int. J. Psychol., 34*, 308-316.

Fortin, C., & Rousseau, R. (1998). Interference from short-term memory processing on encoding and reproducing brief durations. *Psychol. Res., 61*, 269-276.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychol. Rev., 107*, 289-344.

Genovesio, A., Tsujimoto, S., & Wise, S. P. (2006). Neuronal activity related to elapsed time in prefrontal cortex. *J. Neurophysiol.*, *95*, 3281-3285.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev., 84*, 279-325.

Gibbon, J. (1992). Ubiquity of scalar timing with a Poisson clock. *J. Math. Psychol., 36*, 283-293.

Gibbon, J., & Allan, L. (1984). *Timing and time perception.* New York, NY: The New York Academy of Sciences.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Ann. N.Y. Acad. Sci., 423*, 52-77.

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Curr. Opin. Neurobiol.*, *7*, 170-184.

Grinvald, A., Lieke, E., Frostig, R., & Hildesheim, R. (1994). Cortical point spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *J. Neurosci., 14*, 2545-2568.

Grondin, S. (2008). *Psychology of time.* Bingley, UK: Emerald Publishing.

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Atten. Percept. Psychophys., 72*, 561-582.

Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: Evidence from interval discrimination, reproduction, and categorization. *J. Exp. Psychol. Hum. Percept. Perform., 38*, 880-890.

Grossberg, S. (1980). How does the brain build a cognitive code? *Psychol. Rev., 87*, 1-51.

Halliday, J., & Carpenter, R. H. S. (2010). The effect of cognitive distraction on saccadic latency. *Perception, 39*, 41-50.

Hass, J., Blaschke, S., Rammsayer, T., & Herrmann, J. M. (2008). A neurocomputational model for optimal temporal processing. *J. Comput. Neurosci., 25*, 449-464.

Hass, J., & Herrmann, J. M. (2012). The neural representation of time: An information-theoretic perspective. *Neural Comput., 24*, 1519-1552.

Herman, M., Ruppin, E., & Usher, M. (1993). A neural model of the dynamic activation of memory. *Biol. Cybern., 68*, 455-463.

Hertz, J., Krogh, A., & Palmer, R. G. (1991). *Introduction to the theory of neural computation.* Redwood City, CA: Addison-Wesley.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Phys. Rev., 106*, 620-630.

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference.* Cambridge, MA: Cambridge University Press.

Koch, E., & Segev, I. (1998). *Methods in neuronal modeling: From ions to networks.* Cambridge, MA: MIT Press.

Kojima, S., & Goldman-Rakic, P. S. (1982). Delay-related activity of prefrontal neurons in rhesus monkeys performing delayed response. *Brain Res., 248*, 43-49.

Komura, Y., Tamura, R., Uwano, T., Nishijo, H., Kaga, K., & Ono, T. (2001). Retrospective and prospective coding for predicted reward in the sensory thalamus. *Nature, 412*, 546-549.

Kononowicz, T. W., & Van Rijn, H. (2014). Decoupling interval timing and climbing neural activity: A dissociation between CNV and N1P2 amplitudes. *J. Neurosci., 34*, 2931-2939.

Lake, J. I., & Meck, W. H. (2013). Differential effects of amphetamine and haloperidol on temporal reproduction: Dopaminergic regulation of attention and clock speed. *Neuropsychologia, 51*, 284-292.

Lejeune, H. (1998). Switching or gating? The attentional challenge in cognitive models of psychological time. *Behav. Process., 44*, 127-145.

Lejeune, H. (2000). Prospective timing, attention and the switch: A response to 'Gating or switching? Gating is a better model of prospective timing' by Zakay. *Behav. Process., 52*, 71-76.

Lejeune, H., Ferrara, A., Simons, F., & Wearden, J. H. (1997). Adjusting to changes in the time of reinforcement: Peak interval transitions in rats. *J. Exp. Psychol. Anim. Behav. Process., 23*, 211-231.

Leon, M., & Shadlen, M. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron, 38*, 317-327.

Lépine, R., Bernardin, S., & Barrouillet, P. (2005). Attention switching and working memory spans. *Eur. J. Cogn. Psychol., 17*, 329-346.

Lewis, P. A., & Miall, R. C. (2009). The precision of temporal judgement: Milliseconds, many minutes, and beyond. *Philos. Trans. R. Soc. B., 364*, 1897-1905.

Lewis, P. A., & Miall, R. C. (2006). Remembering the time: A continuous clock. *Trends Cogn. Sci., 10*, 401-406.

Lustig, C., Matell, M. S., & Meck, W. H. (2005). Not "just" a coincidence: Frontal-striatal synchronization in working memory and interval timing. *Memory, 13*, 441-448.

Lustig, C., & Meck, W. H. (2005). Chronic treatment with haloperidol induces working memory deficits in feedback effects of interval timing. *Brain Cogn., 58*, 9-16.

MacDonald, C. J., Fortin, N. J., Sakata, S., & Meck, W. H. (2014). Retrospective and prospective views on the role of the hippocampus in interval timing and memory for elapsed time. *Timing Time Percept., 2*, 51-61.

Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval time behavior. *Bioessays, 22*, 94-103.

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence-detection of oscillatory processes. *Cogn. Brain Res., 21*, 139-170.

McCormack, T., & Russell, J. (1997). The development of recency and frequency memory: Is there a developmental shift from reliance on trace-strength to episodic recall? *J. Exp. Child Psychol., 66*, 376-392.

McLaren, I. P. L. (1994). Estimating recency and familiarity using trace strength. *Brit. J. Math. Stat. Psychol., 47*, 227-234.

Meck, W. H. (1984). Attentional bias between modalities: Effect on the internal clock, memory, and decision stages used in animal time discrimination. *Ann. N.Y. Acad. Sci., 423*, 528-541.

Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cogn. Brain Res., 3*, 227-242.

Meck, W. H. (2003). *Functional and neural mechanisms of interval timing*. Boca Raton, FL: CRC Press.

Meck, W. H. (2005). Neuropsychology of timing and time perception. *Brain Cogn., 58*, 1-8.

Meck, W. H., Cheng, R. K., MacDonald, C. J., Gainetdinov, R. R., Caron, M. G., & Çevik, M. Ö. (2012a). Gene-dose dependent effects of methamphetamine on interval timing in dopamine-transporter knockout mice. *Neuropharmacology, 62*, 1221-1229.

Meck, W. H., Doyère, V., & Gruart, A. (2012b). Interval timing and time-based decision making. *Front. Integr. Neurosci., 6*:13.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annu. Rev. Neurosci., 36*, 313-336.

Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Comput., 1*, 359-371.

Oprisan, S. A., & Buhusi, C. V. (2011). Modeling pharmacological clock and memory patterns of interval timing in a striatal beat-frequency model with realistic, noisy neurons. *Front. Integr. Neurosci., 5*:52.

Paninski, L., Pillow, J. W., & Simoncelli, E. P. (2004). Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model. *Neural Comput., 16*, 2533-2561.

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsaki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science, 321*, 1322-1327.

Penney, T. B., Gibbon, J., & Meck, W. H. (2008). Categorical scaling of duration bisection in pigeons (*Columba livia*), mice (*Mus musculus*), and humans (*Homo sapiens*). *Psychol. Sci., 19*, 1103-1109.

Penney, T. B., Holder, M. D., & Meck, W. H. (1996). Clonidine-induced antagonism of norepinephrine modulates the attentional processes involved in peak-interval timing. *Exp. Clin. Psychopharm., 4*, 82-92

Pouget, A., Dayan, P., & Zemel, R. S. (2003). Inference and computation with population codes. *Annu. Rev. Neurosci., 26*, 381-410.

Poynter, W. D. (1983). Duration judgment and the segmentation of experience. *Mem. Cognit., 11*, 77–82.

Rakitin, B. C., Gibbon, J., Penney, T. B., Malapani, C., Hinton, S. C., & Meck, W. H. (1998). Scalar expectancy theory and peak-interval timing in humans. *J. Exp. Psychol. Anim. Behav. Process., 24*, 15-33.

Reutimann, J., Yakovlev, V., Fusi, S., & Senn, W. (2004). Climbing neuronal activity as an event-based cortical representation of time. *J. Neurosci., 24*, 3295-3303.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nat. Rev. Neurosci., 8*, 657-661.

Shankar, K. H., & Howard, M. W. (2010). Timing using temporal context. *Brain Res., 1365*, 3-17.

Shaw, C., & Aggleton, J. P. (1994). The ability of amnesic subjects to estimate time intervals. *Neuropsychologia, 32,* 857-873.

Shi, Z., Church, R. M., & Meck, W. H. (2013). Bayesian optimization of time perception. *Trends Cogn. Sci., 17,* 556-564.

Simen, P., Balci, F., deSouza, L., Cohen, J. D., & Holmes, P. (2011). A model of interval timing by neural integration. *J. Neurosci., 31,* 9238-9253.

Staddon, J. (2005). Interval timing: Memory, not a clock. *Trends Cogn. Sci., 9,* 312-314.

Staddon, J. E. R., & Higa, J. J. (1999). Multiple time scales in simple habituation. *Psychol. Rev., 103,* 720-733.

Stuyven, E., Van der Goten, K., Vandierendonck, A., Claeys, K., & Crevits, L. (2000). The effect of cognitive load on saccadic eye movements. *Acta Psychol. (Amst.), 104,* 69-85.

Taatgen, N. A., van Rijn, H., & Anderson, J.R. (2007). An integrated theory of prospective time interval estimation: The role of cognition, attention, and learning. *Psychol. Rev., 114,* 577-598.

Teki, S., Grube, M., & Griffiths, T. D. (2012). A unified model of time perception accounts for duration- based and beat-based timing mechanisms. *Front. Integr. Neurosci. 5:*90.

van Rijn, H., Kononowicz, T. W., Meck, W. H., Ng, K. K., & Penney, T. B. (2011). Contingent negative variation and its relation to time estimation: A theoretical evaluation. *Front. Integr. Neurosci., 5:*91.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cogn. Affect. Behav. Neurosci., 3,* 255-274.

Wearden, J. H. (1991). Do humans possess an internal clock with scalar timing properties? *Learn. Motiv., 22,* 59-83.

Wearden, J. H. (2001). Internal clocks and the representation of time. In C. Hoerl & T. McCormack (Eds.), *Time and memory: Issues in philosophy and psychology* (pp. 37-58). Oxford, UK: Clarendon Press.

Zakay, D., & Block, R. A. (1995). An attentional-gate model of prospective time estimation. In M. Richelle, V. De Keyser, G. d'Ydewalle, & A. Vandierendonck (Eds.), *Time and the dynamic control of behaviour* (pp. 167-178). Liège, Belgium: Université de Liège.

Zakay, D., & Block, R. A. (2004). Prospective and retrospective duration judgments: An executive-control perspective. *Acta Neurobiol. Exp. (Wars.), 64,* 319-328.