What does scalar timing tell us about neural dynamics?

Harel Z. Shouval[1]*, Marshall G. Hussain Shuler[2], Animesh Agarwal[1,3] and Jeffrey P. Gavornik[4]

[1] Deptartment of Neurobiology and Anatomy, University of Texas Medical School at Houston, Houston, TX, USA
[2] Department of Neuroscience, Johns Hopkins University, Baltimore, MD, USA
[3] Department of Biomedical Engineering, The University of Texas at Austin, Austin, TX, USA
[4] Department of Brain and Cognitive Sciences, The Picower Institute of Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA, USA

The "Scalar Timing Law," which is a temporal domain generalization of the well known Weber Law, states that the errors estimating temporal intervals scale linearly with the durations of the intervals. Linear scaling has been studied extensively in human and animal models and holds over several orders of magnitude, though to date there is no agreed upon explanation for its physiological basis. Starting from the assumption that behavioral variability stems from neural variability, this work shows how to derive firing rate functions that are consistent with scalar timing. We show that firing rate functions with a *log-power* form, and a set of parameters that depend on spike count statistics, can account for scalar timing. Our derivation depends on a linear approximation, but we use simulations to validate the theory and show that *log-power* firing rate functions result in scalar timing over a large range of times and parameters. Simulation results match the predictions of our model, though our initial formulation results in a slight bias toward overestimation that can be corrected using a simple iterative approach to learn a decision threshold.

Keywords: scalar timing, Weber's law, temporal intervals, temporal coding, neural dynamics

1. INTRODUCTION

Errors estimating the intensity of a stimulus commonly scale linearly with the magnitude of the stimulus. This relationship, called *Weber's Law*, has proven to be a surprisingly general property of the brain that accurately describes perception across sensory modalities (Weber, 1843; Coren et al., 1984). We have previously used basic principles to argue that this scaling naturally emerges if neural processes representing stimulus magnitudes have tuning curves with a specific mathematical form and that the generality of the law implies that this is a fundamental organizing principal of neural computation (Shouval et al., 2013).

An analog of Weber's law in the temporal domain, called *linear scaling* or *scalar timing*, states that errors estimating temporal intervals scale linearly with the duration of the intervals (Gibbon, 1977; Church, 2003). Temporal perception has been extensively studied by psychologists and neuroscientists for over 150 years, starting in the 1860s with Fechner (Fechner, 1966), leading to considerable knowledge about the behavioral aspects of temporal perception. Much less is known, however, about the underlying neural substrate responsible for engendering observed timing behavior.

Over the years many theories have been proposed to account for scalar timing. The scalar expectancy theory (Gibbon, 1977; Church, 2003) is based on a counter and accumulator model, conceptually similar to counting the ticks of a mechanical clock, and variability arises from comparison errors with remembered reference values. Another class of models assumes an ensemble of neurons oscillating at different frequencies, and timing is produced by decision neurons which become active only when a precise set of the oscillating neurons are coactive (Matell and

Meck, 2000). These models are akin to a Fourier transform of the desired temporal response profile. Variability stems from the addition of stochastic noise to the ensemble dynamics, and it has recently been shown analytically that general addition of noise at various levels in the model can result in scalar timing (Oprisan and Buhusi, 2011, 2014). Note that these models are currently derived based on continuous dynamical systems, not spiking neural models. Drift diffusion models have also been proposed to provide a mechanistic basis of interval timing, though with spike-statistics that are inconsistent with scalar timing. Recent derivations of this model, with drift that is driven by opponent inhibitory and excitatory processes, can account for scalar timing (Rivest and Bengio, 2011; Simen et al., 2011). A final class of models, including this work, assume that timing is derived from the state of dynamic neural responses. For example, time can be estimated from the threshold crossing of decaying neural response (Staddon et al., 1999), or from a precisely designed set of leaky integrators (Shankar and Howard, 2012).

Here, by extending our earlier analysis (Shouval et al., 2013) to the temporal domain, we explore the relationship between neural dynamics and temporal perception and propose a theory of scalar timing based on experimentally verifiable physiological processes. Our approach is based on the assumption that estimates of a temporal interval vary on a trial-by-trial basis due to spike count variability (Dean, 1981; Tolhurst et al., 1983; Churchland et al., 2010). Although our analysis is mathematically homologous to the intensity variable case, the physiological substrates of time and intensity estimation are quite different. Here we show that neural processes with with activity levels that dynamically progress with

a log-power temporal profiles can account for scalar timing, in much the same way that our previous work showed that Weber's law results from log-power tuning curves. Our analysis is based on a linear approximation, with results that are less precise than we found when analyzing intensity variables. Though the log-power model does produces scalar timing, our initial formulation of the model with the linear approximation results in a small bias toward underestimation and slightly less variability than found in simulations. In this paper we also derive a discrete approximation, which is applicable only in the temporal domain, that precisely estimates simulated neural variability. We also demonstrate that the bias is a consequence of our initial threshold selection criteria that can be easily eliminated with a simple algorithm that learns the correct threshold value to accurately decode desired intervals.

2. METHODS AND RESULTS

We start with the assumption that the brain uses the temporal evolution of neural activity, which progresses with predictable stochastic dynamics, to estimate intervals. Specifically, we assume that some stimulus at time $t = 0$ initiates a neural process (describing either a single neuron or, more likely, a neural ensemble) that is characterized by a spike rate function $r(t)$ (see **Figures 1B,C**), which either increases (Roitman and Shadlen, 2002) or decreases (Shuler and Bear, 2006) monotonically. The average spike count within a window τ is:

$$R(t) = \int_{t-\tau/2}^{t+\tau/2} r(t+t')dt' \qquad (1)$$

which can be approximated as $R(t) \approx \tau r(t)$ for small τ values. As illustrated in **Figure 2**, the temporal interval described by this process is defined as the time required for $R(t)$ to reach some threshold R_0, which can be set as $R_0 = r(t_{tar})$ for a target time t_{tar}. Noise driven fluctuations of $r(t)$ result in variability of trial-by-trial estimates, t_{est}, of the encoded target time.

In this framework there is a direct relationship between the magnitude of temporal estimate errors (which can be easily recorded using standard psychophysical methods) and spike count statistics that can be used to infer a mathematical form of $r(t)$, and thus the underlying physiology, subject to the linear constraint on estimate errors as a function of t_{tar} specified by the scalar timing law. A simple linear approximation (illustrated in **Figure 1A**) of this relationship between the interval estimate and spike count has the form:

$$\sigma_t(t_{est}) \approx \frac{\sigma_R(t_{tar})}{|R'(t_{tar})|} = \frac{\sigma_R(t_{tar})}{|\tau r'(t_{tar})|} \qquad (2)$$

where $\sigma_R(t)$ is the standard deviation of the average spike count at time t, and $R'(t_{tar})$ is derivative of the spike count curve with respect to the time, estimated at the target time. Note that $\sigma_t(t_{est})$ is the standard deviation of the estimation of the time t over many trials.

The scalar law states that errors estimating t scale linearly with t. Using standard deviation as the error measure:

$$\sigma_t(t) = \alpha \cdot t. \qquad (3)$$

where α specifies the slope of linear scaling, equivalent to the "Weber fraction."

Combining Equations 2 and 3:

$$\tau \frac{dr(t)}{dt} = \frac{dR(t)}{dt} = \pm \frac{\sigma_R(t)}{\alpha t} \qquad (4)$$

where the $+$ sign is valid when the slope of $r(t)$ is positive and the $-$ sign when it is negative.

We assume that spike count variability can be characterized using a power-law model with the form:

$$\sigma_R(t) = \beta \left(\tau r(t)\right)^\rho \qquad (5)$$

where the parameters β and ρ specify the specific noise model. This power-law model can account for many forms of spike-count variability. For example, $\rho = 1/2$ and $\beta = 1$ result in Poisson noise, and the $\rho = 0$ case is the constant noise case, which means spike count variability does not depend on the spike count. Experimentally spike count variability is found to be close to Poisson and often with somewhat larger variability than Poisson ($\rho \approx> 1/2$). Although the power-law noise is a relatively general model, it obviously cannot account for all forms of noise.

Applying this form to Equation 4, we obtain a differential equation relating the neural firing rate to specific noise and estimate error models:

$$\frac{dr(t)}{dt} = \pm \left(\frac{\beta \tau^{\rho-1}}{\alpha}\right) \frac{r(t)^\rho}{t} \qquad (6)$$

The solution of Equation 6 has a *log-power* form:

$$r(t) = K \cdot \left(\pm log(t/t_0)\right)^n \qquad (7)$$

where $K = \frac{1}{\tau}(\beta(1-\rho)/\alpha)^n$, and $n = \frac{1}{1-\rho}$. This relationship holds whether $r(t)$ rises ($t \geq t_0$, "+"case) or falls ($t < t_0$,"−" case) monotonically. The integration constant t_0 has a simple interpretation: it is the minimal (or maximal) time that can be estimated using this specific monotonically increasing (or decreasing) firing rate function (as shown in **Figures 1B,C**) for different values of t_0. Note that all the parameters of the *log-power* function are determined by measurable spike statistics and behavioral performance; none of them are free parameters.

The specific shape of the general *log-power* form depends primarily on the spike count statistics. In the constant noise case ($\rho = 0$) this equation reduces to Fechner's law (Fechner, 1966). Hence, Fechner's law can be seen as making an implicit constant noise assumption. In the special and unrealistic case of proportional noise ($\rho = 1$) a power-law solution is obtained (Stevens, 1961).

Experimental recordings are often characterized by a nearly-linear relationship between mean spike count and variance(Dean, 1981; Tolhurst et al., 1983; Churchland et al., 2010). In the near-Poisson case, ($\rho = 1/2$), the *log-power* form has an exponent of $n = 2$ (note, the examples in **Figures 1B,C** assume Poisson statistics). We have previously shown for the case of magnitude
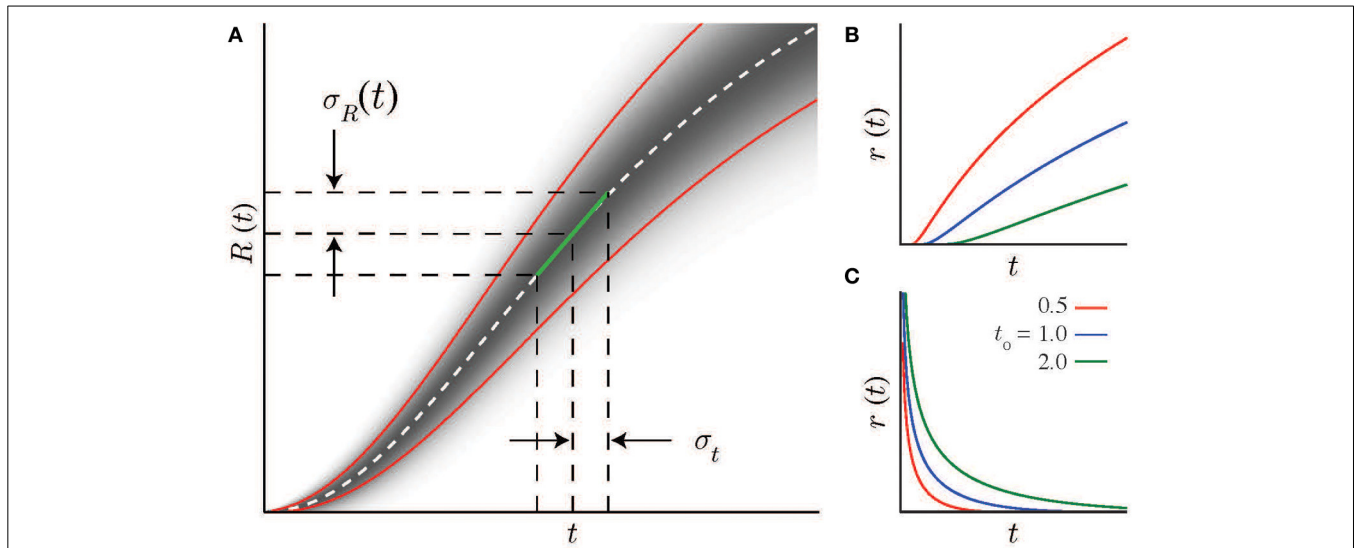
FIGURE 1 | Scalar timing and neural statistics. (A) A local linear approximation (green line, Equation 2) of the the average firing rate $R(t)$ (real distribution shown schematically by the gradient as a function of t, mean and standard deviations indicated by dashed-white and solid red lines) together with the scalar timing law leads to Equation 4, the solution of which (Equation 7 for the case of Poisson noise) is the firing rate curve $r(t)$. Note, R' is the slope of the linear approximation to $R(t)$. **(B,C)** Example firing rate curves with Poisson spike statistics for different values of the integration constant t_0. **(B)** Increasing solutions are defined above minimal values at t_0. **(C)** Decreasing solutions are defined below maximal values at t_0.
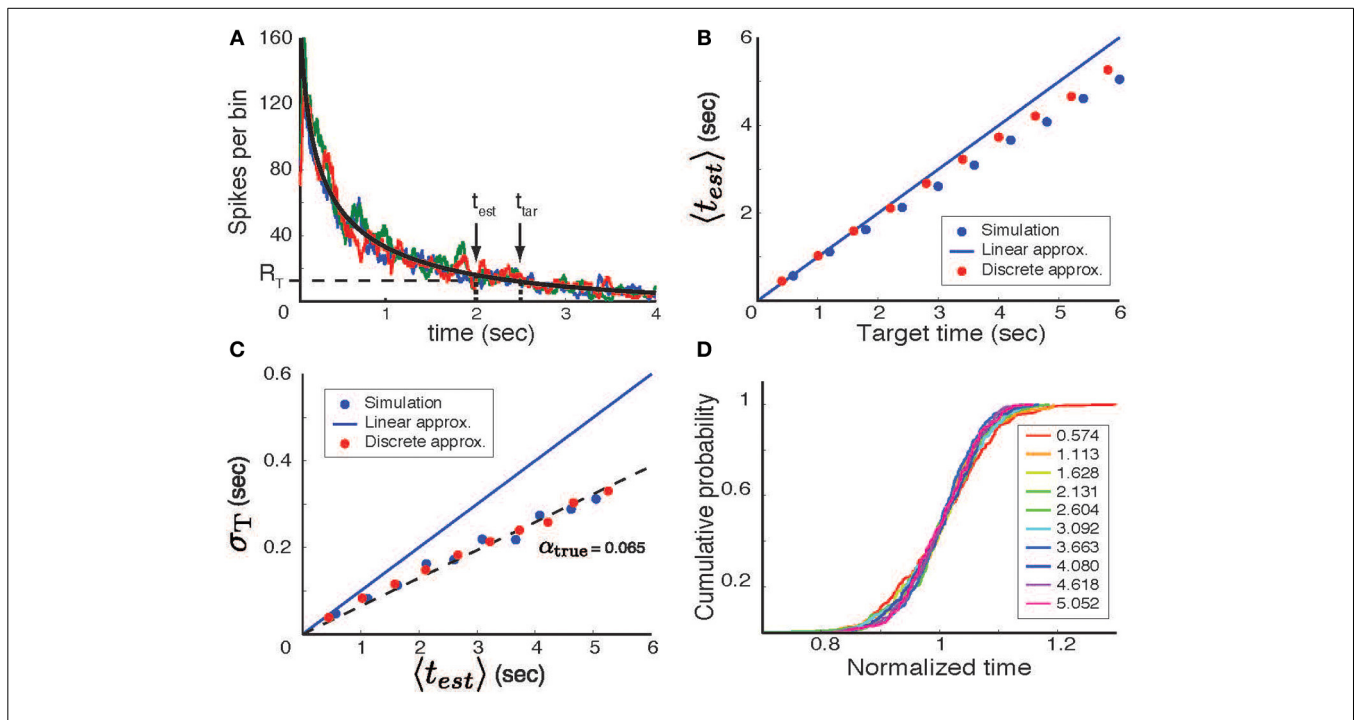

FIGURE 2 | Temporal interval estimation. (A) A stimulus (at time $t = 0$) initiates a neural process with a mean firing rate (black line, determined by linear approximation theory) that decreases with time. In each trial the actual number of spikes varies stochastically; three trial-by-trial examples of the spike count variable are shown by the colored lines. The time estimate in each trial is determined by the first threshold crossing (R_T - horizontal dashed line) of the spike count variable. The actual estimated time for one trial (t_{est}) is shown in comparison to the target time (t_{tar}). **(B)** The mean time predicted by the model ($\langle t_{est} \rangle$, averaged over 200 trials) as a function of the target time. Blue circles based on simulations, red circles using discrete approximation. **(C)** The standard deviation of the time estimate (σ_T) as a function of the mean predicted interval. **(D)** When rescaled by the mean estimated time (values specified by the color-code shown in the legend), the cumulative distributions of the actual response times overlap and are statistically indistinguishable (KS-test). These distributions were generated using Poisson statistics, a decreasing *log-power* function, $t_0 = 10$ and $\tau = 0.1$ sec.

estimation that Weber's law can be based on the tuning curves of either single neurons or neural ensembles (Shouval et al., 2013). Similarly here, while scalar timing can result if a single neuron's time-varying activity follows a *log-power* function, it is more likely to arise from the combined activity of a heterogenous population of neurons whose collective activity has the appropriate form.

To test the validity of our theory, we simulated a stochastic neural process with a monotonically falling spike rate in time (the increasing case, not discussed, is similar). Specifically, as per the derivations above, simulations were performed by generating spikes using a non-homogeneous Poisson process with a firing-rate parameter that decreased as a *log-power* function of time. Firing rates were determined by convolving the resultant spikes trains with a square window of width $\tau = 100$ ms. The estimate of the temporal interval (t_{est}) was defined, on a per-trial basis, as the time at which the firing rate first reached threshold (R_0 set to $r(t_{tar})$). The result of these simulations are shown in **Figures 2B,C**. The mean value of t_{est} is close to, but a bit shorter than that predicted by the linear approximation theory (**Figure 2B**, blue circle) and the standard deviation is a linear function of the mean estimated time (**Figure 2C**–blue circle), although with a slope lower than that predicted by the linear approximation. Nevertheless, the rescaled distributions are almost completely overlapping (**Figure 2D**) and paired Kolmogorov-Smirnov tests find that the differences between these distributions are not statistically significant (although small differences may emerge with longer time spans, larger α values, or more trials). This shows that the *log-power* firing rate function indeed produces scalar timing, but that the theory described above results in an overestimate of error and a small bias of the mean.

It is possible to obtain a discrete approximation that better captures the simulation results. This approximation is obtained by dividing time into non-overlapping time bins of length τ, such the average spike count within a time bin (designated by the integer i) is $R_i \approx \tau \cdot r((i - 0.5) \cdot \tau)$. Under the assumptions that the bins are non-overlapping and have no significant temporal correlations, the spike counts in each bin are conditionally independent. Then, for a given spike generation model, the probability of a threshold crossing within a time bin as time unfolds is:

$$P_c(i|R_0) = \sum_{n \le R_0} P_s(n|R_i), \qquad (8)$$

where $P_s(n|R)$ is the probability of emitting n spikes given the mean spike count R (note that this formulation assumes a decreasing function $r(t)$, for the increasing function the sum is over $n \ge R_0$). The probability that the first threshold crossing occurs in time bin j is

$$P_{FC}(j|R_0) = P_c(j|R_0)\Pi_{i=1}^{j-1}(1 - P_c(i|R_0)). \qquad (9)$$

This distribution can be used to calculate the mean, $\langle t_{est} \rangle$, and standard deviation, σ_T, of elapsed time estimates. Results of these calculations, (**Figures 2B,C**, red circles) agree closely with the numerical simulation results. The small discrepancy between this calculation and the actual simulations arises from partitioning

time into non-overlapping time bins. Coarse grained simulations in which zero crossings are allowed only at these discrete points agree perfectly with the results of this discrete approximation.

Despite the close agreement between theory and simulations, the model as described consistently underestimates the mean target time. This bias, which results from the somewhat arbitrary decision to select $R_0 = r(t_{tar})$, can be corrected if the thresholds are learned rather than chosen directly from the spike count function. To learn the threshold (R_0) we used a simple iterative learning rule:

$$\frac{dR_0}{dt} = \pm\eta(t_{tar} = t_{est}) \qquad (10)$$

where the $+$ sign corresponds to the monotonically falling firing rate case, and the $-$ sign is used for the monotonically increasing cases, and $\eta << 1$ is the learning rate. This procedure quickly converges to provide an unbiased estimate of the target times (**Figure 3A**); error still scales linearly with time (**Figure 3B**) and the discrete approximation accounts well for the slope.

3. DISCUSSION

Relating behavior to its underlying physiological mechanism is a fundamental aim of neuroscience. Here we have shown how, in a class of models based on the idea that time is estimated based on the dynamic state of neural processes, to relate scalar timing to the time varying firing rate of neurons. We show that, given firing rate statistics characterized by a power-law, scalar timing arises from a log-power firing rate function with parameters that depend on the spike statistics. Our derivation relies on a linear approximation, but we have also shown that a log-power function results in scalar timing irrespective of this approximation. The initial method for setting the detection threshold using the mean of the firing rate function causes a small estimate bias, but this can be corrected using an iterative procedure to find an appropriate detection threshold. Further, we have shown how to use a discrete approximation to calculate better estimates of encoded time and variability given a log-power firing rate function. These result depend on a mathematical analysis which is similar to the one used in our previous analysis of Weber's law for intensity variables (Shouval et al., 2013). Though it produces scalar timing, the linear approximation is less precise in the temporal domain than it was when we used it to analyze the coding of stimulus intensity. Accordingly, it was necessary to introduced a discrete approximation in order to accurately calculate simulated variability. We also showed how the decision-selection threshold can bias the encoded interval. Most models that account for Weber's law in the intensity domain are completely distinct from models that account for scalar timing. Our results show that a single mathematical approach can provide a unified explanation for these two distinct observations.

Although we derived firing-rate functions for the case of perfectly-linear scalar timing, the same procedure used to generate Equation 4 can also be used when if the relationship between estimation error and time is non-linear (Grondin, 2012). Similar derivations are possible for other functional forms of Equation 3. Indeed there are experiments showing that scalar-timing is pre-
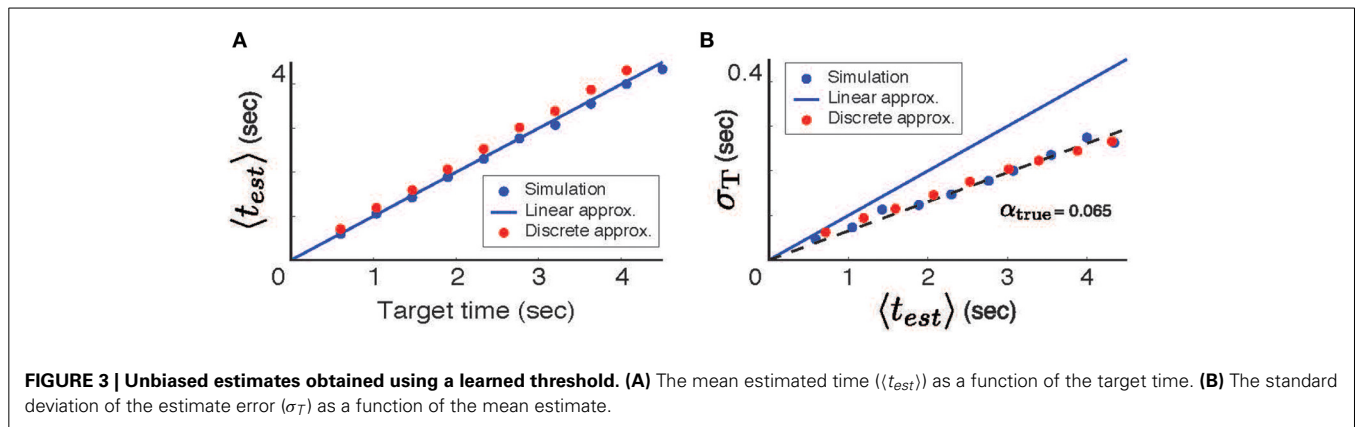
FIGURE 3 | Unbiased estimates obtained using a learned threshold. (A) The mean estimated time ($\langle t_{est} \rangle$) as a function of the target time. **(B)** The standard deviation of the estimate error (σ_T) as a function of the mean estimate.

cisely linear only in a limited range (Getty, 1975; Bizo et al., 2006; Grondin, 2012), and the exact forms of scaling observed in these experiments could be used to replace the linear scaling assumed here. There is no guarantee that an analytical result can be derived in such cases, but numerical solutions are always possible. Similarly, the same type of approach can be used to obtain a firing rate function, either analytically or numerically, assuming non-power-law forms for neural spike statistics.

Our analytical derivation produces a monotonically falling functions that are valid only below and upper threshold (t_0) and monotonically increasing functions valid only above a lower threshold (t_0). There is a lower threshold below which we can not evaluate time intervals, which would indicate that the monotonically increasing results are possibly more realistic. However, experimental results showing different Weber fractions at different temporal intervals could also be interpreted as indications that different processes are used for different time scales (Getty, 1975). One possibility is that falling functions could be used for very short temporal durations, on the order of a second or less, and increasing functions for longer durations.

As we outlined above, various models of interval timing have been proposed over the years to account for scaler timing (Gibbon, 1977; Matell and Meck, 2000; Church, 2003; Durstewitz, 2003; Oprisan and Buhusi, 2011, 2014; Rivest and Bengio, 2011; Simen et al., 2011; Shankar and Howard, 2012) and some share key properties of the model proposed here (Staddon et al., 1999; Durstewitz, 2003). An entirely different class of models is based on the idea that time can be read from the dynamic state of circuits in the cortical network (Buonomano and Mauk, 1994; Karmarkar and Buonomano, 2007), though the conditions for scalar timing in these models have not been analyzed. Some of the previously developed models of scalar timing are based on abstract entities such as counters and accumulators (Gibbon, 1977; Church, 2003), and some are dependent on continuous variables (Matell and Meck, 2000; Karmarkar and Buonomano, 2007; Oprisan and Buhusi, 2011), while others can be interpreted in terms of spiking inhibitory and excitatory neurons (Rivest and Bengio, 2011; Simen et al., 2011) and require nearly-perfect integration for the decision process. The model presented here is formulated directly in terms populations of spiking neurons with experimentally measurable variables. There are no free

parameters in our model, since all depend directly on neural and behavioral statistics. Therefore, our theory has the advantage that it can be tested experimentally at the physiological level.

Our analysis indicates a very precise log-power form for the firing rate function. One might wonder, rightly, if it realistic to expect a neural processes to have such a precise formulation. It is important to realize that our analysis does not require or claim that any single neuron should display precise log-power dynamics, though to get true linear scaling the relevant population of neurons must possess this form. The population can be composed of individual neurons with diverse response dynamics, as we demonstrated in the intensity domain (Shouval et al., 2013). A question not answered here is how single neurons or a population of neurons can develop firing rate functions with a desired form. Possible answers are provided by previous work showing how single neurons with active conductances (Durstewitz, 2004; Shouval and Gavornik, 2011) or networks of interacting neurons (Gavornik et al., 2009; Gavornik and Shouval, 2011) can be tuned to, or even learn de-novo, specific temporal dynamics. An additional possibility is that decision neurons can select (in the Hebbian sense) a sub-population of existing neurons with a combined spike rate that has a log-power form without requiring that the dynamics of any of the individual neurons change at all, though we can not here propose a biologically realistic mechanism for making this choice.

Recent experiments (Leon and Shadlen, 2003; Shuler and Bear, 2006; Chubykin et al., 2013) have made physiological recordings from cortical cells in animals as they learn temporal discrimination tasks. These results show that the firing rate function of cells change when animals learn different temporal intervals and theoretical models have been devised to account for them (Reutimann et al., 2004; Gavornik et al., 2009). Analyzing these cases, and determining a single framework that leads to scalar timing, is quite different in many respects from the analysis carried out here. We are currently studying this issue. Experimentally it requires many trials to change the firing rate function. One possibility, as mentioned above, is that the model presented here describes mechanisms used to discriminate times in a manner that requires little or no learning, whereas other models would be required to describe how representations of specific temporal intervals are encoded over many trials. Regardless, the work here

makes a strong prediction that any neural process used to encode temporal intervals that display scalar timing, with our minimal assumptions, will have firing rates that evolve as a log-power of time.

REFERENCES

Bizo, L. A., Chu, J. Y., Sanabria, F., and Killeen, P. R. (2006). The failure of Weber's law in time perception and production. *Behav. Process.* 71, 201–210. doi: 10.1016/j.beproc.2005.11.006

Buonomano, D. V., and Mauk, M. D. (1994). Neural network model of the cerebellum: temporal discrimination and the timing of motor responses. *Neural Comput.* 6, 38–55. doi: 10.1162/neco.1994.6.1.38

Chubykin, A. A., Roach, E. B., Bear, M. F., and Shuler, M. G. H. (2013). A cholinergic mechanism for reward timing within primary visual cortex. *Neuron* 77, 723–735. doi: 10.1016/j.neuron.2012.12.039

Church, R. (2003). "A coincise introduction to scalar timing theory," in *Functional and Neural Mechanisms of Interval Timing, Chapter 1*, ed W. Meck (Boca Raton, FL: CRC Press), 3–22.

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* 13, 369–378. doi: 10.1038/nn.2501

Coren, S., Porac, C., and Ward, L. M. (1984). *Sensation and Perception. 2nd Edn.* San Diego, CA: Harcourt Brace Janovich.

Dean, A. F. (1981). The variability of discharge of simple cells in the cat striate cortex. *Exp. Brain Res.* 44, 437–440. doi: 10.1007/BF00238837

Durstewitz, D. (2003). Self-organizing neural integrator predicts interval times through climbing activity. *J. Neurosci.* 23, 5342–5353.

Durstewitz, D. (2004). Neural representation of interval time. *Neuroreport* 15, 745–749. doi: 10.1097/00001756-200404000-00001

Fechner, G. (1966). *Elements of Psychophysics* (H. E. Adler, Trans). New York, NY: Holt, Reinhart and Winston (Original work published 1860).

Gavornik, J. P., and Shouval, H. Z. (2011). A network of spiking neurons that can represent interval timing: mean field analysis. *J. Comput. Neurosci.* 30, 501–513. doi: 10.1007/s10827-010-0275-y

Gavornik, J. P., Shuler, M. G. H., Loewenstein, Y., Bear, M. F., and Shouval, H. Z. (2009). Learning reward timing in cortex through reward dependent expression of synaptic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6826–6831. doi: 10.1073/pnas.0901835106

Getty, D. J. (1975). Discrimination of short temporal intervals: a comparison of two models. *Percept. Psychophys.* 18, 1–8. doi: 10.3758/BF03199358

Gibbon, J. (1977). Scalr expectency theory and weber's law in animal timing. *Psychol. Rev.* 84, 279–325. doi: 10.1037/0033-295X.84.3.279

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. doi: 10.1037/a0027188

Karmarkar, U. R., and Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron* 53, 427–438. doi: 10.1016/j.neuron.2007.01.006

Leon, M. I., and Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* 38, 317–327. doi: 10.1016/S0896-6273(03)00185-5

Matell, M. S., and Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays* 22, 94–103. doi: 10.1002/(SICI)1521-1878(200001)22:1<94::AID-BIES14>3.0.CO;2-E

Oprisan, S. A., and 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. doi: 10.3389/fnint.2011.00052

Oprisan, S. A., and Buhusi, C. V. (2014). What is all the noise about in interval timing? *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20120459. doi: 10.1098/rstb.2012.0459

Reutimann, J., Yakovlev, V., Fusi, S., and Senn, W. (2004). Climbing neuronal activity as an event-based cortical representation of time. *J. Neurosci.* 24, 3295–3303. doi: 10.1523/JNEUROSCI.4098-03.2004

Rivest, F., and Bengio, Y. (2011). Adaptive drift-diffusion process to learn time intervals. *arXiv*:1103.2382v1. Available online at: http://arxiv.org/abs/1103.2382

Roitman, J. D., and Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci. Off. J. Soc. Neurosci.* 22, 9475–9489.

Shankar, K. H., and Howard, M. W. (2012). A scale-invariant internal representation of time. *Neural Comput.* 24, 134–193. doi: 10.1162/NECO_a_00212

Shouval, H. Z., Agarwal, A., and Gavornik, J. P. (2013). Scaling of perceptual errors can predict the shape of neural tuning curves. *Phys. Rev. Lett.* 110, 168102. doi: 10.1103/PhysRevLett.110.168102

Shouval, H. Z., and Gavornik, J. P. (2011). A single spiking neuron that can represent interval timing: analysis, plasticity and multi-stability. *J. Comput. Neurosci.* 30, 489–499. doi: 10.1007/s10827-010-0273-0

Shuler, M. G., and Bear, M. F. (2006). Reward timing in the primary visual cortex. *Science* 311, 1606–1609. doi: 10.1126/science.1123513

Simen, P., Balci, F., de Souza, L., Cohen, J. D., and Holmes, P. (2011). A model of interval timing by neural integration. *J. Neurosci.* 31, 9238–9253. doi: 10.1523/JNEUROSCI.3121-10.2011

Staddon, J., Higa, J., and Chelaru, I. (1999). Time, trace, memory. *J. Exp. Anal. Behav.* 71, 293–301. doi: 10.1901/jeab.1999.71-293

Stevens, S. S. (1961). To honor fechner and repeal his law: a power function, not a log function, describes the operating characteristic of a sensory system. *Science* 133, 80–86. doi: 10.1126/science.133.3446.80

Tolhurst, D. J., Movshon, J. A., and Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* 23, 775–785. doi: 10.1016/0042-6989(83)90200-6

Weber, E. (1843). *De Pulsu Resortione, Auditu et Tactu: Annotationes Anatomicae et Physiologicae*. Leipzig: Koehler.