A firing-rate model of Parkinsonian deficits in interval timing

Eric Shea-Brown[1,2], John Rinzel[1,2], Brian C. Rakitin[3], and Chara Malapani[4,5]

[1] Courant Institute of the Mathematical Sciences,
[2] Center for Neural Science,
New York University, New York, NY 10012, U.S.A.
[3] Cognitive Neuroscience Division, Taub Institute,
[4] Department of Psychiatry, Columbia University,
[5] Division of Biopsyhology, New York State Psychiatric Institute,
New York, NY 10032, U.S.A.

September 16, 2005

15 pages, 7 figures

Corresponding author:
E. Shea-Brown, Courant Institute of the Mathematical Sciences, 251 Mercer
St., New York, NY 10012
ebrown@math.nyu.edu, PH: 212-998-3132, FAX: 212-995-4121, www.math.nyu.edu/~ebrown

1

Abstract

To account for deficits in interval timing observed in Parkinson's Disease (PD) patients, we develop a model based on the accumulating firing rate of a neural population with recurrent excitation. This model naturally produces the curvilinear accumulation of neural activity introduced to timing psychophysics by Miall (*Models of Neural Timing*, Elsevier Science, 1996), and implicated in Parkinsonian timing by Malapani and Rakitin (*Functional and Neural Mechanisms of Interval Timing*, CRC Press, 2003). The parameters essential for our model are the strength of the net neural feedback and the mean rate of inputs to the population from external brain areas. Systematic variations in these parameters reproduce the PD migration effect, in which estimates of long and short intervals drift towards each other, as well as uniform slowing of time estimates observed under other experimental conditions. For example, our model suggests that dopamine depletion in PD patients increases the neural feedback parameter and decreases the effective input parameter for populations involved in the production of time estimates. The model also explains why the migration effect will be associated with a violation of the *scalar property*, the linear increase in the standard deviation of time estimates with the duration of the target interval that is ubiquitous in healthy participants. We also show that the effect of systematically decreasing the input rate parameter in our model is equivalent to increasing thresholds, so that either of these changes may be associated with the Parkinsonian state.

Theme: NEURAL BASIS OF BEHAVIOR
Topic: Cognition
Key Words: Interval timing, Parkinson's Disease, scalar property, migration, curvilinear accumulator, firing rate model

1. INTRODUCTION

This paper is motivated by experiments that revealed specific deficits in the interval timing behavior of Parkinson's Disease (PD) patients [37, 41, 40, 42]. In the most developed of these experiments, the timing abilities of PD patients were assessed using the Peak Interval (PI) timing task [54] immediately after learning a pair of target time intervals with the assistance of behavioral feedback and 24 hours later without the benefit of such feedback [37]. The results indicated two separable, dopamine-dependent error patterns associated with storage of time intervals (i.e., *encoding*) and retrieval from temporal memory (i.e., *decoding*) [37]. These patterns cannot be immediately reconciled with established models of the timing process, especially Scalar Expectancy Theory, or SET [17, 8, 16, 18]. Our goal is to present a modification of SET based on simple models of the behavior of neural populations that can explain the Parkinsonian timing phenomena.

The PI task requires participants to produce behavioral responses (i.e., a burst of button presses) at one of two memorized time intervals. Originally developed for animal research [7, 9, 57], it was adapted for human use in healthy subjects (e.g, [54, 24, 25, 35, 53]) and patients with neurological disease (e.g., [37, 41, 36, 38]). Given the high ratio of variance associated with timing versus motor components [52] (as motor requirements are minimal and time ranges are of seconds in duration), the PI task is well suited for the assessment of timing functions in patient populations encountering severe motor deficits, as in PD. Indeed, PD patients experience specific problems when producing movements, such as increased reaction time [6, 11, 39], movement time [5, 58] and speech production time [34, 63], as well as deficits in programming and synchronizing motor responses [48, 67, 51, 50]. These motor deficits may add a constant ("motor") variance in timing performance [68], particularly in tasks with substantial motor requirements such as repetitive tapping [29, 51, 50, 67].

Demonstrating separable properties of interval storage and retrieval in PD patients requires two distinct experimental sessions, the first performed in the presence of behavioral feedback and immediately following demonstration of a standard interval, and the second without feedback or demonstration. This was originally reported as the "encode-decode" task design [37]. In this design, during the first session the standard interval that subjects are to reproduce is first demonstrated. Then, "production" trials begin, in which subjects are asked to reproduce the standard intervals via timed button presses following a cue; on a fraction of these trials, the standard interval is again demonstrated to the subject. Additional behavioral feedback is delivered after other production trials in the form of a histogram indicating whether the response was too short or too long. Two separate blocks of such trials are performed, one with each of a longer (17 sec) and a shorter (6 sec) standard interval. In the second session,

performed on the following day, subjects produce both intervals (in separate blocks) without further demonstration of the standard intervals and with no behavioral feedback. Therefore, participants learn and store the target time intervals only during the first session (henceforth referred to as the *training* session), but retrieve these intervals from temporal memory during both the training session and the second day's testing session.

PD patients' drug state (ON or OFF medication) on the two successive days varied according to their assignment to one of four experimental groups. The ON-ON group was provided with L-Dopa during both training and testing sessions, and the OFF-OFF group was tested without L-Dopa for both sessions. The ON-OFF group was provided with L-Dopa during the training but not the testing session, and vice-versa for the OFF-ON group. Clinical measures for all four groups are given in Table 1. The motivation was the hope that by crossing patients' drug state with the availability of information about the accuracy of reproductions of the standard intervals, we could determine whether dopamine (DA) deficiency (associated with being in an OFF state) selectively affected memory storage (*encoding*), retrieval (*decoding*), or both.

Figure 1 summarizes the resulting behavioral performance during the testing session. Correct estimates are obtained when both storage and retrieval occur ON medication (see Panel (A), left). However, (B, left) shows that retrieving the trace of two different time intervals while OFF medication results in "migration," a pattern of bi-directional errors such that reproductions of each interval drift in the direction of the midpoint of the intervals. This migration effect is seen for the OFF-OFF group (see panel (D, left)). However, when time intervals are stored OFF medication, but retrieved ON, Panel (C, left) shows that both intervals are overestimated. In addition to migration and overestimation of mean time estimates, retrieval of time estimates OFF medication results in a violation of the *scalar property of timing variability* (Figure 1 (B,D, right panels)). (The scalar property implies a linear proportionality between standard deviation and mean of time estimates, see below.) By contrast, for retrieval ON medication, the scalar property holds, regardless of whether intervals were encoded ON or OFF medication (A,C, right panels). Note that migration is always accompanied by a violation of the scalar property (specifically, by estimates for the shorter interval having a relatively broad distribution), a fact we will return to in the modelling below. While we show here only data from the testing session, similar migration and scalar timing effects occur for patients OFF L-dopa during the training sessions. This indicates that the availability of behavioral feedback is not a critical factor in producing the statistical trends just discussed.

These results forced a rethinking of some underlying assumptions of Scalar Expectancy Theory (SET), a framework for modeling timing originally developed by Gibbon in 1977 [15] and still strongly influential in the field [17, 8, 16,

2, 14, 65]. SET is an information-processing model that describes the sources of timing errors that lead to the scalar property, the behavioral phenomenon in which distributions of time estimates for different length target intervals appear as scaled versions of a single fundamental distribution and which is robustly observed in both animals [8, 14] and humans [54, 19, 28, 64]. According to this model, pulses from an internal pacemaker are integrated by an accumulator when attention is directed to time [32, 46, 69, 52]. The value of the accumulator increases linearly with time, and is stored in memory upon the occurrence of reinforcement or feedback [17, 30]. A ratio-based decision process then compares values of the accumulator to remembered values in order to determine responses on future trials [4, 3]. The SET model parameters can fluctuate from trial to trial, but their overall statistics are independent of the target time interval. This produces the scalar property [17, 8], and explains the fact that most experimental manipulations produce monotonic changes across a range of target intervals [40]. Indeed, SET could accommodate the unidirectional, proportional rightward shift seen in the OFF-ON groups by relative slowing of the pacemaker rate, which sets the temporal slope of the accumulator (see [18] for a more detailed discussion). The migration effect, however, cannot be reconciled with SET or other similar timing theories (e.g., [69]). The same is true of the violation of the scalar property for times produced OFF L-dopa. As a result, new modeling approaches are necessary.

Malapani and Rakitin [40] produced a theoretical account of both the migration effect and uniform overestimation trends in PD interval timing data by modifying the accumulator so that its value increased with a curvilinear dependence on time. The critical notion was that if L-dopa altered the curvature of the accumulators, then functions associated with timing ON and OFF L-dopa could be made to cross, and duration-dependent errors like migration could emerge. The underlying computational model was a stochastic, neural network-type architecture comprised of binary-valued neural units proposed by Miall [47], in which the mean value of network activity exponentially approaches an equilibrium value from its initial state. In this paper, we establish a different, idealized model that also produces curvilinear accumulation of network activity, but which is immediately tractable analytically and which may be more directly related to the firing rate dynamics of neural populations.

Specifically, we consider a related class of 'leaky integrator' models derived from idealized equations for the firing rate of a recurrently connected neural population receiving input from another, separate neural population. Models of this type have a long history in neural modelling (cf. [59, 61, 66, 62, 26, 1, 45, 23]). The input to the recurrent population and its firing rate, respectively, correspond to the pacemaker and accumulator components of SET. In the variant that we use here, the time-dependent neural activation (firing rate) can display three qualitatively different trajectories, either approaching a steady

state (as for the model of [47]), increasing linearly in time, or increasing at an accelerating rate. Therefore, this model can be used to further study the effects of curvilinear accumulation in interval timing.

In particular, this variety of temporal dynamics allows us to address an additional feature of the behavioral data not treated in [40] - that of the scalar property of variability in time estimates. In particular, we show that model parameter values corresponding to the ON-drug condition produce timing distributions with a fixed ratio of mean to standard deviation (the scalar property). This results from a linear accumulation of firing rates in time, as in classical models [16]. In contrast, parameter values modelling the OFF-drug condition result in violation of the scalar property, with proportionally broader distributions at shorter intervals (in agreement with experimental data).

The present paper extends the work of [40] in three additional ways. First, we use explicit solutions to our simple firing rate-based model to show the equivalence of timing models that use two distinct mechanisms to time different intervals. We also group parameters into sets that have identical effects, identifying only two combined parameters that determine the model's dynamics. Finally, we note that our model, while substantially removed from the underlying physiology, is parameterized by values that may nevertheless be related to averaged or 'mean field' descriptions of asynchronously firing neural populations (e.g., [66, 12, 49]). Therefore, we are able to draw preliminary conclusions about the changes in effective parameters in neural timing circuits that may result from transitions from ON- to OFF-drug states.

The balance of the paper proceeds as follows. In the methods section, we introduce the firing rate model, discuss the form of its solutions, and identify simplifying parameter combinations. Next, we give the results of the model for both trends in mean values of time estimates across the various experimental conditions, and adherence to or violation of the scalar property. Here, we emphasize how parameters must vary across experimental conditions in order to reproduce the trends observed in the data of [40], and develop several predictions that follow from these necessary variations in parameters. We then show that these conclusions hold regardless of whether thresholds or accumulator slopes are changed to time different intervals, and identify an equivalence between increasing external input to the recurrent network and decreasing thresholds. Finally, we discuss our findings in the context of physiological mechanisms, covering both the model's shortcomings and its predictions for future experiments.

2. METHODS

A self-excitatory firing rate model of interval timing

An idealized model for the firing rate r of a self-excitatory neural population is

$$\tau \frac{dr}{dt} = -r + f(\beta r + c) \tag{1}$$

where τ is the effective recruitment timescale of the population and $f(\cdot)$ is the population input-output function (i.e., the steady state firing rate vs. current relationship). Additionally, β is the strength of excitatory recurrent connections, and c is the strength of inputs from areas external to the population itself. See Fig. 2. We do not specify the location of the population, but view it as a simplified aggregate model of the timing circuit. Therefore, the variable r represents in a highly abstracted way the activity of dopamine modulated pathways of the basal ganglia, as well as their input and output structures (contrasting the region-specific modelling of, e.g. [44, 19, 10] as well as architecturally-based models that are not directly related to timing [22, 13]).

We further assume that the net input $\beta r + c$ remains in a range where $f(\cdot)$ may be approximated by the linearization $f(x) = gx$, where g is the 'gain' of the population, so that Eqn. 1 becomes

$$\tau \frac{dr}{dt} = \lambda r + I \ , \tag{2}$$

where we have combined terms to define the overall neural feedback parameter $\lambda = \beta g - 1$ and the effective input $I = gc$. Below, we will tacitly assume that time is measured in units of τ, eliminating this parameter in the above. This choice is for simplicity: we note that effects of changing τ can be captured by simultaneously rescaling λ and I.

We take initial conditions $r(0) = 0$ for Eqn. (2), giving the solution

$$r(t) = \frac{I}{\lambda} \left(\exp \lambda t - 1 \right) \tag{3}$$

Fig. 3 displays the trajectories that this solution takes for different values of the parameters λ and I, demonstrating that each plays a distinguished role: λ sets the curvature of the accumulating firing rate while I determines its initial slope. Note that in the special case $\lambda = 0$, $r(t)$ is linear: $r(t) = I t$.

The firing rate $r(t)$ produces time estimates as follows: the first time at which $r(t)$ reaches a preset threshold θ is the time estimate on that trial. To time different intervals, there are two possibilities: (i) the parameters λ, I could differ or (ii) the thresholds θ could differ between the intervals. We call these options, respectively, the *one-threshold* and *two-threshold* models, and show below that they are essentially equivalent.

3. RESULTS

3.1 The two-threshold model and predicted interval timing behavior

Encoding and decoding time intervals

We assume that, during the training or *encode* phase of the experiments, parameters in the timing model are adjusted so that (the subject's) estimates match the target time intervals assigned by the experimenter. We first consider a framework in which the threshold parameter $\bar{\theta}$ is the adjustable parameter, so that for two-interval tasks there are two thresholds $\bar{\theta}_1$, $\bar{\theta}_2$. Later, however, we will see that tuning of other parameters to the various intervals is in fact equivalent.

Following [40], we allow the parameters describing accumulating firing rates to vary depending on medication state (ON vs. OFF L-dopa) and on whether the accumulator results in storage (encoding) or production (decoding) of time estimates. The crucial notion is that one curve is used to set thresholds (denoted by $r_{encode}(t)$), while another, possibly different, curve $r_{decode}(t)$ is used to translate these thresholds into time estimates (henceforth called time productions). Disparities between target times and time productions arise when the curves $r_{encode}(t)$ and $r_{decode}(t)$ differ.

Specifically, we will show that the parameters of the accumulator (2) must change as follows among experimental conditions to reproduce experimentally observed effects on both mean time estimates and the scalar property:

$$0 = \lambda_{ON} \quad \leq \quad \lambda_{OFF,\,encode} < \lambda_{OFF,\,decode} \tag{4}$$

$$I_{OFF,\,decode} \quad < \quad I_{ON} \leq I_{OFF,\,encode} \quad, \tag{5}$$

where at least one of the $\leq$ relations must be a strict inequality. We specify a single set of parameters $(\lambda_{ON},\ I_{ON})$ which applies to both encoding and decoding in the ON-drug condition. Note that, for accumulators operating in ON-drug conditions, recurrent feedback balances leak (giving $\lambda = 0$), but for OFF-drug states, the neural feedback parameter λ must incrementally increase. Additionally, for the decode OFF condition, the net input I must also decrease. We now explain why these parameter orderings reproduce the experimental data. Below, we will also show why the parameter orderings of (4)-(5), as well as the value $\lambda_{ON} = 0$, are the only choices that could have been made.

Trends in time estimates for the various experimental conditions

In our model the initial slope of the firing rate $r(t)$ is I. Therefore, our parameter choices (4)-(5) guarantee that for all sufficiently early times, the firing

rates for the different experimental contingencies are ordered as follows:

$$r_{OFF,decode}(t) < r_{ON}(t) < r_{OFF,encode}(t) \ . \tag{6}$$

However, for sufficiently later times, the exponential growth of $r(t)$ dominates and we have

$$r_{ON}(t) < r_{OFF,encode}(t) < r_{OFF,decode}(t) \ . \tag{7}$$

Following [16], we assume that mean threshold values for two standard intervals T_1 and T_2 are determined by a memory system which stores the value of the appropriate rate $r(t)$ at these target times. Figure 4 shows the resulting thresholds when this threshold encoding is done ON and OFF medication as solid and chain-dotted lines respectively. Parameters have been chosen so that the inequality (6) holds for the short interval $T_1 = 1$, while (7) is valid for the long interval $T_2 = 3$.

Figure 4 also demonstrates the trends in time production that result from the inequalities (6)-(7). Since the same accumulator is used to encode and decode thresholds in the $ON - ON$ experimental condition (upper left), time productions will have the correct means. However, when thresholds that were established using the accumulator $r_{ON}(t)$ are later decoded using $r_{OFF,decode}(t)$ (the ON-OFF condition), migration results: the short interval is overestimated, while the longer is underestimated. To model the $OFF - ON$ and $OFF - OFF$ conditions, thresholds are set using a third process $r_{OFF,encode}(t)$ and then decoded into time productions via $r_{ON}(t)$ and $r_{OFF,decode}(t)$, respectively. This results in consistent overestimation of intervals in the $OFF - ON$ case, and migration in the $OFF - OFF$ condition.

We emphasize that these general trends occur for any settings of the accumulator parameters that satisfy the inequalities above. For Figure 4 and all subsequent figures, we chose $I_{ON} = 1$, $I_{OFF,encode} = 1.25$, $I_{OFF,decode} = 0.35$ (as per Eqn. (4)), and $\lambda_{ON} = 0$, $\lambda_{OFF,encode} = 0$, $\lambda_{OFF,decode} = 1$ (all in accordance with Eqn. (5)). Additionally, we took target intervals to be $T_1 = 1$ and $T_2 = 3$, where time is in units of τ (for direct comparison with the experimental results of [37], where $T_1 = 6$ sec. and $T_2 = 17$ sec., take $\tau \approx 6$ sec.).

The scalar property and its violation

Throughout this paper, we adopt Gibbon's 'ratio rule' for determining thresholds [16]. This results in Gaussian distributions of thresholds, with the standard deviations of these distributions assumed to be fixed proportions of their means. As is well known, such distributions of thresholds, in conjunction with linear accumulator processes, give rise to the scalar property (in particular, distributions of time estimates also have standard deviations proportional to their means). However, for curvilinear accumulators, as result from setting $\lambda \neq 0$, the scalar property is *violated* in general.

9

To see this, let $p_\theta(\theta)$ be the probability density of thresholds. We say that a time estimate of duration t is produced if the accumulating rate reaches the threshold for a given trial at time t; that is, if $r(t) = \theta$. Therefore, the density of interval estimates $p(t)$ is described by $p(t)dt = dr(t)\, p_\theta\left(r(t)\right)$, or

$$p(t) = \frac{dr(t)}{dt} p_\theta\left(r(t)\right) \quad . \tag{8}$$

For linear accumulators, $r'(t) \equiv I$ (where $' = \frac{d}{dt}$), so that $p(t)$ is simply a rescaled version of the threshold density $p(\theta)$. Thus the scalar property of the threshold densities is directly inherited in densities of interval estimates.

However, for curvilinear accumulators, $r'(t)$ is no longer constant, so these arguments no longer apply. In particular, when $\lambda > 0$, $r'(t)$ is an increasing function, so we expect densities $p(t)$ to be 'tighter' for longer intervals than for shorter intervals, relative to the scalar distributions of thresholds $p(\theta)$ (see Figure 5). This is the case whenever time estimates are produced OFF drug, and, using the parameters described in the previous section, gives rise to the densities of time productions in the various task conditions shown in Figure 6. This figure shows that two key trends in variability of behavioral time productions are captured by the model: i) distributions possess the scalar property whenever they are produced for parameters representing the ON L-dopa condition, and ii) whenever times are produced OFF medication, distributions for the shorter interval are relatively more variable than would be predicted by the scalar property [37].

Taken as a whole, however, distributions of time estimates produced OFF L-dopa are relatively less variable, in comparison with time estimates ON L-dopa, for the model than for the data. We have verified that this discrepancy can easily be remedied by assuming a broader distribution of thresholds in the OFF vs. ON L-dopa condition (specifically, with twice the standard deviation).

Necessity of parameter settings

We now show that the parameter values and orderings of (4)-(5) are in fact necessary. That is, there are no other choices for which our model would reproduce the following primary features of the experimental data: i) time estimates with correct means and satisfying the scalar property in the ON-ON case, ii) migration and violation of the scalar property in the ON-OFF and OFF-OFF conditions, and iii) overestimation of both intervals, but with the scalar property, in the OFF-ON condition.

Experimental fact i) implies that there is a single accumulator $r_{ON}(t)$ for both the encode and decode processes, and that $\lambda_{ON} = 0$. Next, from fact ii), for migration in the ON-OFF case, we require $\lambda_{ON} < \lambda_{OFF,\,decode}$ and $I_{OFF,\,decode} < I_{ON}$ (so that $r_{OFF,\,decode}(t)$ crosses $r_{ON}(t)$ from below); similarly,

for the OFF-OFF case we require $\lambda_{OFF,\,encode} < \lambda_{OFF,\,decode}$ and $I_{OFF,\,decode} < I_{OFF,\,encode}$. Finally, fact iii) implies $\lambda_{OFF,\,encode} \leq \lambda_{ON}$ and $I_{OFF,\,encode} \leq I_{ON}$, where at least one of these relations must be strict inequality, (so that $r_{OFF,\,encode}$ lies everywhere above r_{ON}). Putting these conclusions together exactly yields Equations (4)-(5).

We remark additionally that the "less than *or* equals" ambiguity in the relations $\lambda_{OFF,\,encode} \leq \lambda_{ON}$ and $I_{OFF,\,encode} \leq I_{ON}$ may be resolved by the additional observation that the extent to which each interval is overestimated in the OFF-ON experimental data is a fixed proportion of the target time for each interval. This requires $\lambda_{OFF,\,encode} = 0 = \lambda_{ON}$ and hence $I_{OFF,\,encode} > I_{ON}$, which are the choices we have made in producing the figures in this paper. However, this latter level of parameter specificity is not required for the additional predictions of our model that we discuss next.

Predictions of the curvilinear accumulator model

Above, we showed that, if our accumulator model is to reproduce the basic features i), ii), and iii) of the behavioral data listed there, then model parameters *must* be constrained as per Eqns. (4)-(5). This constraint allows us to make separate, additional predictions for statistical patterns that should be present in the existing behavioral data.

The first set of these predictions concerns the relative extent of overestimation and underestimation of time intervals in the various experimental conditions. For all parameter choices satisfying (4)-(5), underestimation of the longer interval will be accentuated for the ON-OFF relative to the OFF-OFF condition, because the curves $r_{ON}(t)$ and $r_{OFF,\,decode}(t)$ are necessarily further apart around the longer time than the curves $r_{OFF,\,encode}(t)$ and $r_{OFF,\,decode}(t)$. Similar reasoning shows that the overestimation of the shorter interval will be accentuated for test session in the OFF-OFF relative to the OFF-ON condition, and likewise for the OFF-OFF relative to the ON-OFF group. All three of these model predictions are consistent with the experimental data of [37], as statistical reanalysis of this data confirms.

In particular, overestimation of the short interval is significantly higher in the OFF-OFF relative to the OFF-ON group ($p < .005$; F=9.6) and likewise for the OFF-OFF relative to the ON-OFF condition ($p < .05$; F=4.2). The predicted trend is present in the behavioral data for the accentuated underestimation of the longer interval in the ON-OFF relative to the OFF-OFF group, but the difference did not reach significance. This may be due to the low number of subjects included in the OFF-OFF group, which is a question to be addressed in future experiments.

Our model, coupled with the necessary parameter values derived above, also predicts that distributions of time estimates will be more skewed towards

11

shorter times for all intervals produced in the OFF relative to the ON condition (in other words, the OFF distributions will have a greater "leftward" skew or, equivalently, a lesser "rightward" skew). This follows from the fact that $\lambda_{OFF,\,decode} > 0$, which leads to rates $r_{OFF,\,decode}(t)$ which accelerate in time and hence, via Eqn. (8), to a compression of the tail of the distribution corresponding to longer estimates. However, statistical analyses have not found this trend to be reliably present in the existing experimental data.

3.2 Equivalent models and parameters

As already mentioned, Gibbon's rule for thresholds states that they are normally distributed with a standard deviation proportional to their mean $\bar{\theta}$. Explicitly:

$$p_\theta(\theta) = \frac{1}{\sqrt{2\pi k^2 \bar{\theta}^2}} \exp\left(-\frac{(\theta - \bar{\theta})^2}{2k^2 \bar{\theta}^2}\right) \ ,$$

where k is the proportionality constant. Rewriting, this is

$$p_\theta(\theta) = \frac{1}{\bar{\theta}}\frac{1}{\sqrt{2\pi k^2}} \exp\left(-\frac{(\theta/\bar{\theta} - 1)^2}{2k^2}\right) \equiv \frac{1}{\bar{\theta}}q(\frac{\theta}{\bar{\theta}}) \ . \tag{9}$$

Inserting this latter expression into (8), we have

$$p(t) = r'(t)\frac{1}{\bar{\theta}}q(\frac{r(t)}{\bar{\theta}}) \ . \tag{10}$$

Then, substituting in the solution (3) gives

$$p(t) = \frac{I}{\bar{\theta}}\exp(\lambda t)q\left(\frac{I}{\bar{\theta}}\frac{(\exp(\lambda t) - 1)}{\lambda}\right) \ . \tag{11}$$

This expression reveals that densities of time estimates will be Gaussian only when $\lambda = 0$, and will otherwise be skewed. Additionally, Eqn. (11) also shows that the densities of response times predicted by our model depend on only two quantities: λ and the *ratio* $I/\bar{\theta}$. This fact reduces the number of free parameters in the model. As we now explain, it also implies that another modeling paradigm is actually exactly equivalent to that discussed above, and that different parameter variations can have the same effects on predicted distributions of time productions.

The equivalent one-threshold model

An alternative mechanism for timing different intervals is suggested by firing rate recordings from frontal brain areas during delayed match to sample tasks [33, 56]. These data suggest that there is a single threshold that is used to

time intervals of various duration, with the accumulation occurring at different speeds for the different target times. We represent this situation via a *one-threshold model*, in which there is a single threshold $\bar{\theta}$, but two different values of I (one for each of the two intervals being timed) which are separately adjusted during encoding so that $r(T) = \bar{\theta}$ at each target time T.

As an example, Fig. 7 illustrates this situation for the ON-OFF condition. For comparison with the results of the two-threshold model already discussed, we take $T_1 = 1$, $T_2 = 3$, $\lambda_{ON} = 0$, and $\lambda_{OFF,decode} = 1$, exactly as above. Additionally, we take $\bar{\theta} = 3$, and 'tune' inputs to the target times, so that $I_1 = 3$ and $I_2 = 1$. Figure 7 (right) also shows the resulting distribution of time estimates, which exactly match those for the corresponding two-threshold model (compare with Fig. 6). We now explain why the distributions under the one- and two-threshold models are identical.

The key observation is that, following encoding, either the mean threshold $\bar{\theta}$ (in the *two threshold* model) or the input I (in the *one threshold* model) has been adjusted so that, for the target time T, $r(T) = \bar{\theta}$. Substituting this into (3) gives

$$\frac{I}{\bar{\theta}} = \frac{\lambda}{e^{\lambda T} - 1} \ . \tag{12}$$

We emphasize that this relationship holds in *both* the one threshold and two threshold models: although the values of I and $\bar{\theta}$ may be different in each case, their ratio will be the same for fixed values of λ and T. This is a consequence of the simple form of (3), in which the input I enters as a multiplicative parameter; for a general accumulator, one would *not* expect that the different learning procedures in which thresholds or inputs are adjusted during encoding would both give the same accumulator dynamics.

We now consider how this encoded ratio of $\frac{I}{\theta}$ is decoded to produce time estimates in the one- and two- threshold models. As discussed above, I and λ may be different for accumulators assumed to encode vs. decode target times. We now show that, if these parameters are changed identically in the one- and two-threshold models, then both of these models produce identical distributions of time productions.

In particular, assume for both the one- and two-threshold models that the value of I during decoding of an arbitrary time interval is varied by a multiplicative constant k from its value during encoding. Then, because the ratios $\frac{I}{\theta}$ must be identical for both the one- and two-threshold models during encoding, they will continue to take the same value during decoding. If we additionally assume that the value of λ is the same for decoding in both the one-threshold and two-threshold models, then we see that both values $\frac{I}{\theta}$ and λ, which completely determine the distribution of time productions $p(t)$ via Equation (11), are identical for decoding in both the one-threshold and two-threshold models. In other words, these distributions are necessarily identical

for the one-threshold and two-threshold models, no matter what the 'fixed' input level I is taken to be in the two-threshold model or what the 'fixed' single threshold is taken to be in the one-threshold model.

Therefore, all of the observations made above about migration when decoding is performed in the OFF condition, accuracy or overestimation in the ON condition, and violation or preservation of the scalar property in these two cases carry over exactly to the one-threshold model. Therefore, so do results on the necessity of parameter settings given by Eqns. (4)-(5), as well as the predictions for behavioral data that follow from these settings.

An equivalent hypothesis on parameter variations between task conditions

Above, we allowed the parameters λ and I to differ for the accumulators $r(t)$ used in encoding vs. decoding time intervals ON and OFF L-dopa, but assumed that there is no variation in the mean threshold $\bar{\theta}$ between these experimental conditions. That is, we assumed that $\bar{\theta}$ is determined via Eqn. (12) during the encode process, regardless of medication state, and that it is exactly this same value of $\bar{\theta}$ which is later used for decoding, again regardless of medication state. However, another possibility is that $\bar{\theta}$ changes between experimental condition instead of I, while λ continues to vary as above. These two possibilities are equivalent if the corresponding parameters are varied inversely, because Eqn. (11) shows that only the ratio $\frac{I}{\theta}$ determines distributions of time productions.

To explore this possibility, we assume that there is a multiplicative term b, analogous to the 'criterion factor' of scalar expectancy theory [15], which relates accumulator rates $r(t)$ to threshold values. Specifically, during the encode process we assume that rates $r(T)$ at target times T are encoded as thresholds $\bar{\theta} = b\, r(T)$, and that during decoding time estimates are made when $r(t) = \frac{\theta}{b}$, where the value of b is specific to task condition (encode vs. decode, OFF vs. ON). In this case, the parameter ordering

$$b_{OFF,\,decode} < b_{ON} \leq b_{OFF,\,encode} \tag{13}$$

along with the ordering of the associated λ values given by Eqn. (4) will yield exactly the same results as those developed above for covariation of λ and I between conditions.

4. DISCUSSION

In this paper, we have shown how a threshold-based mechanism for interval timing, similar to that of scalar expectancy theory [17, 8] but extended to

include accumulating firing rates that have a curvilinear dependence on time, can reproduce both the migration and uniform overestimation trends observed in two-interval timing experiments with Parkinson's patients [37]. Our model also accounts for the scalar property that these time estimates display whenever they are produced ON L-dopa medication, and the violation of this property OFF L-dopa: compare Figs. 1 and 6. For the model to reproduce these trends, the dynamics of the accumulating firing rate must vary both between the different experimental conditions (ON vs. OFF drug therapy) and between the different task stages (encoding vs. decoding).

The model is based on an idealized model of a recurrent, excitatory neural network, and is characterized by two parameters: λ, the net neural feedback, I the external drive to the population. These parameters must vary among experimental conditions as in Eqns. (4)-(5) to reproduce the primary features of the experimental data. We have shown that this result continues to hold regardless of whether one assumes that thresholds or accumulator slopes are adjusted to time different intervals (i.e., the one-threshold vs. two-threshold models). Additionally, we demonstrated that systematic biases in relating thresholds to firing rates can play the role of variations of the neural drive I among task conditions.

In the methods section, we rescaled and grouped parameters to define the neural feedback $\lambda = \beta g - 1$ and the neural drive $I = gc$, which completely determine the dynamics of our model. Here, β is the strength of the excitatory recurrent connections, c measures inputs from areas external to the population, and g is the input-output gain of the population. Informed by Eqns. (4)-(5), we can assess how these values may vary in order to affect the required changes in λ and I among task conditions. For example, the transition from ON-drug encoding to OFF-drug decoding (in which λ increases and I decreases) could be caused by an increase in the weight β with a (smaller) decrease in gain g, by an increase in gain g with a (greater) decrease in external inputs c, or by other covariations in parameters. Any of these choices results in the key dynamical effect, that the putative neural accumulators which produce time estimates OFF medication start increasing at relatively slow rates but accelerate over time, due to diminished drive but excessive positive feedback relative to the ON medication case. Despite these myriad options, if one views our recurrent neural population as the simplest imaginable model of a dopamine-modulated timing circuit, our general conclusions about parameter variations in the OFF vs. ON L-dopa state do provide constraints on the dynamics of accumulator-based interval timing in Parkinson's patients.

Three remarks are appropriate in considering how the present, extremely simple model might be implemented biologically. First, as emphasized in [59, 60], we note that the single firing rate r may be thought of as characterizing a distinguished subpopulation of cells whose activity increases along

15

the 'line attractor' of a more complex recurrent network. Second, we note that when $\lambda = 0$ (as in the ON-medication state) our model possesses a continuum of steady states in the absence of external drive I. Such systems are often referred to as neural integrators, and have been studied widely in the context of oculomotor control (e.g., [59, 61, 21]). However, as [59] and references therein have pointed out, this neural integrator property requires fine tuning of parameters (i.e., setting λ perfectly equal to 0), and even small deviations in this tuning result in errors that accumulate exponentially in time. This raises the legitimate question of the plausibility of our model of interval timing *ON* L-dopa, which does require $\lambda = 0$. Recent work addresses the fine tuning problem by introducing bistability and hysteresis in subunits of the recurrent network [31, 21]. While these papers do explore some cases in which these properties enhance the robustness of neural integration, is not clear whether bistability and hysteresis alone lead precisely to robust ramping dynamics, or whether additional assumptions on network architecture, beyond those already explored, will be required. Nevertheless, we note that *in vivo* data from animals performing timing tasks does provide evidence for linearly accumulating firing rates (e.g., [56, 33]) or linearly accumulating firing rates in concert with other firing patterns [43]. A direction of our current work is to determine whether physiologically motivated parameter changes can produce instabilities in otherwise robust integrator models which result in the type of curvilinear accumulation patterns that are attributed to the OFF-drug state above.

Thirdly, we comment on the mechanisms of trial-to-trial variability in time estimates. Above, we adopt Gibbon's assumption of distributed thresholds [17, 16]. However, we believe that our model would give similar results if, as proposed by [55], this variability followed instead from the rapid stochastic variations in $r(t)$ itself that may be expected from fluctuations inherent in finite size neural populations (or, similarly, from rapid variations in the separate mechanism that detects threshold crossings, not modelled here). This is because the mean value and temporal spread of time estimate distributions in this case would still depend on the time and rate with which the mean value of accumulating firing rates crosses through a range of values near threshold.

While our model explains the two-interval behavioral data of [37], without further assumptions it cannot account for the results of experiments which show that the migration effect vanishes when only a *single* interval must be timed OFF L-dopa. Specifically, in tasks following the encode ON - decode OFF protocol but in which only a single interval must be encoded and reproduced, each duration is typically overestimated [41]. Without modification, the present model would predict that when this experiment was performed with single, longer intervals, these longer intervals would in fact be underestimated, exactly as for the two-interval task studied above. To reconcile this

situation, one can assume that accumulator parameters depend on the range of time intervals being encoded, so that the crossing point of, for example, $r_{ON}(t)$ and $r_{OFF,decode}(t)$, itself varies from task to task (in particular, that this point always lies to the 'right' on the time axis from the single intervals that were tested in [41]), but this idea needs further investigation.

We also note that our model addresses only averages of time estimates produced over an entire block of trials, and that trial-to-trial variations in behavior are beyond its scope. Future investigations into the trial-to-trial adjustment of behavior and inferred trial-to-trial variations in model parameters could address questions including that just raised, as they could enable studies of how the 'crossing point' referred to above moves as additional intervals are added to a subject's timing repertoire. We further note that our model was developed to account for timing of intervals in the seconds range using the peak interval procedure, and it is an open question whether similar mechanisms apply to other tasks and interval ranges [27, 20].

The present model makes a number of testable predictions. First, the magnitudes of timing errors should differ in a predictable way across experimental protocols, as explained and then verified via a reanalysis of experimental data in the section "predictions of the curvilinear accumulator model." Second, as also discussed in that section, distributions of time estimates produced OFF L-dopa should display relatively greater leftward skew (or less rightward skew) relative to comparable distributions produced ON L-dopa. Statistical analyses did not find this trend to be reliably present in the existing experimental data.

Other predictions of our model will require future experiments. The most direct of these is that recordings from the timing circuits of animals that have been, for example, pharmacologically induced to display migration behavior in timing tasks should reveal firing rates with a curvilinear dependence on time. Another, fourth prediction of the model is that for any task design there will be a single critical interval duration toward which migration will occur (as noted in [40]). In other words, regardless of the number of different intervals encoded, intervals below this critical duration will be overestimated while intervals above this duration will be underestimated. Fifth, we predict that violation of the scalar property, with estimates of the shortest interval(s) displaying excessive variability, will occur whenever migration in mean values of time estimates are observed.

References

[1] L. Abbott. Firing-rate models for neural populations. In O. Benhar, C. Bosio, P. Del Giudice, and E. Tabat, editors, *Neural Networks: from Biology to High-Energy Physics*, pages 179–196. ETS Editrice, Pisa, 1991.

[2] L.G. Allan. The influence of the scalar timing model on human timing research. *Behavioural Processes*, 44:101–117, 1998.

[3] L.G. Allan and K. Gerhardt. Temporal bisection with trial referents. *Perception and Psychophysics*, 63:Perception and Psychophysics, 2001.

[4] L.G. Allan and J. Gibbon. Human bisection and the geometric mean. *Learning and Motivation*, 22:39–58, 1991.

[5] R. Benecke, J. C. Rothwell, J. P. R. Dick, B. L. Day, and C. D. Marsden. Performance of simultaneous movements in patients with parkinsons disease. *Brain*, 109:739757, 1986.

[6] C.A. Bloxham, D.J. Dick, and M. Moore. Reaction times and attention in parkinsons disease. *Journal of Neurology, Neurosurgery and Psychiatry*, 50:11781183, 1987.

[7] A.C. Catania. Reinforcement schedules and psychophysical judgments: A study of some temporal properties of behavior. In W.N. Schoenfeld, editor, *The theory of reinforcement schedules*, pages 1–42. Appleton-Century-Crofts, New York, 1970.

[8] R.M. Church. A concise introduction to scalar expectancy theory. In W.H. Meck, editor, *Functional and neural mechanisms of interval timing*. CRC Press, Boca Raton, FL, 2003.

[9] R.M. Church, W.H. Meck, and J. Gibbon. Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes*, 20:135–155, 1994.

[10] J.L. Contreras-Vidal and W. Schultz. A predictive reinforcement model of dopamine neurons for learning approach behavior. *Journal of Computational Neuroscience*, 6:191–214, 1990.

[11] E.V. Evarts, H. Teravainen, and D.B. Calne. Reaction time in parkinsons disease. *Brain*, 104:167, 1981.

[12] N. Fourcaud and N. Brunel. Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural Comp.*, 14:2057–2110, 2002.

[13] M.J. Frank, B. Loughry, and R.C. O'Reilly. Interactions between the frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience*, 1:137–160, 2001.

[14] C.R. Gallistel and J. Gibbon. Time, rate, and conditioning. *Psychological Review*, 107:289–344, 2000.

[15] J. Gibbon. Scalar Expectancy Theory and Weber's Law in animal timing. *Psychological Review*, 84:279–335, 1977.

[16] J. Gibbon. Ubiquity of scalar timing with a Poisson clock. *J. Math. Psych.*, 35:283–293, 1992.

[17] J. Gibbon, R.M. Church, and W.H. Meck. Scalar timing in memory. *Ann. NY Acad. Sci.*, 423:52–77, 1984.

[18] J. Gibbon and C. Malapani. Neural basis of timing and time perception. In L. Nabel, editor, *Encyclopedia of Cognitive Science*, pages 305–311. Macmillan Reference, London, 2002.

[19] J. Gibbon, C. Malapani, C. Dale, and C. R. Gallistel. Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology*, 7:170–179, 1997.

[20] J. Gibbon, C. Malapani, C.L. Dale, and C.R. Callistel. Toward a neurobiology of temporal cognition: advances and challenges. *Curr. Opin. Neurobiol.*, 7:170–184, 1997.

[21] Levine J.H. Major G. Tank D.W. Goldman, M.S. and H.S. Seung. Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cerebral Cortex*, 1:1185–1195, 2003.

[22] A.M. Graybiel and M. Kimura. Adaptive neural networks in the basal ganglia. In J.C. Houk, J.L. Davis, and D.G. Beiser, editors, *Models of Information Processing in the Basal Ganglia*, pages 103–116. MIT Press, Cambridge, 1995.

[23] S. Grossberg. Nonlinear neural networks: principles, mechanisms, and architectures. *Neural Networks*, 1:17–61, 1988.

[24] S.C. Hinton and W.H. Meck. Frontal-striatal circuitry activated by human peak-interval timing in the supra-seconds range. *Cognitive Brain Research*, 21:171–182, 2004.

[25] S.C. Hinton and S.M. Rao. One-thousand one...one-thousand two...: Chronometric counting violates the scalar property in interval timing. *Psychonomic Bulletin and Review*, 11:24–30, 2004.

[26] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. USA*, 82:3088–3092, 1984.

[27] R. B. Ivry. The representation of temporal information in perception and motor control. *Curr. Opin. Neurobiol.*, 6:851–857, 1996.

[28] R. B. Ivry and E. Hazeltine. Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance*, 21:3–18, 1995.

[29] R. B. Ivry and S. W. Keele. Timing functions of the cerebellum. *Journal of Cognitive Neuroscience.*

[30] L.A. Jones and J.H. Wearden. More is not necessarily better: examining the nature of the temporal reference memory component in timing. *Quarterly Journal of Experimental Psychology*, 56:321–343, 2003.

[31] A. Koulakov, S. Raghavachari, A. Kepecs, and J. Lisman. Model for a robust neural integrator. *Nature Neurosci.*, 5:775 – 782, 2002.

[32] H. Lejeune. Switching or gating? the attentional challenge in cognitive models of psychological time. *Behavioral Processes*, 44:127–145, 1999.

[33] M. Leon and M. Shadlen. Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron*, 38:317–327, 2003.

[34] P. Lieberman, E. Kako, J. Friedman, G. Tajchman, L.S. Feldman, and E.B. Jiminez. Speech production, syntax comprehension, and cognitive deficits in parkinson's disease. *Brain Lang.*, 43:169–189, 1992.

[35] C. Lustig and W. Meck. Paying attention to time as one gets older. *Psychological Science*, 12:478–484, 2001.

[36] C. Lustig and W. H. Meck. Chronic treatment with haloperidol induces deficits in working memory and feedback effects of interval timing. *Brain and Cognition (in press)*, 2005.

[37] C. Malapani, B. Deweer, and J. Gibbon. Separating storage from retrieval dysfunction of temporal memory in Parkinsons disease. *J. Cognitive Neuroscience*, 14:1–12, 2002.

[38] C. Malapani, B. Dubois, G. Rancurel, and J. Gibbon. Cerebellar dysfunctions of temporal processing in the seconds range in humans. *Neuroreport*, 9:3907–3912, 1998.

[39] C. Malapani, B. Pillon, B. Dubois, and Y. Agid. Impaired simultaneous cognitive task performance in parkinsons disease; a dopamine related dysfunction. *Neurology*, 44:319326, 1994.

[40] C. Malapani and B. Rakitin. Interval timing in the dopamine-depleted basal ganglia: From empirical data to timing theory. In W.H. Meck, editor, *Functional and Neural Mechanisms of Interval Timing.* CRC Press, Boca Raton, FL, 2003.

[41] C. Malapani, B. Rakitin, R. Levy, W.H. Meck, B. Deweer, B. Dubois, and J. Gibbon. Coupled temporal memories in Parkinson's disease: a dopamine-related dysfunction. *J. Cogn. Neurosci.*, 10:316–331, 1998.

[42] C. Malapani, B.C. Rakitin, S. Fairhurst, and J. Gibbon. Neurobiology of timing. *Cognitive Processsing*, 3:3–20, 2002.

[43] M. Matell, W. Meck, and M. Nocolelis. Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behavioral neuroscience*, 117:760–773, 2003.

[44] M.S. Matell and W.H. Meck. Neurophysiological mechanisms of interval timing behavior. *BioEssays*, 22:94–103, 2000.

[45] J.L. McClelland. On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86:287–330, 1979.

[46] W. H. Meck and R.M. Church. A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 9:320–334, 1983.

[47] C. Miall. Models of neural timing. In M. Pastor and J. Artieda, editors, *Time, Internal Clocks, and Movement.* Elsevier Science, 1996.

[48] R. Nakamura, H. Nagasaki, and H. Narabayashi. Disturbances of rhythm formation in patients with parkinsons disease: Part 1. characteristics of tapping response to the period signals. *Brain Lang.*

[49] D. Nykamp and D. Tranchina. A population density approach that facilitates large-scale modeling of neural networks: analysis and application to orientation tuning. *J. Comp. Neurosci.*, 8:19–50, 2000.

[50] D. J. OBoyle. On the human neuropsychology of timing of simple repetitive movements. In C. M. Bradshaw and E. Szabadi, editors, *Time and behaviour: Psychological and neurobehavioural analyses*, page 459515. Amsterdam: North-Holland/Elsevier, 1997.

[51] M. A. Pastor, M. Jahanshahi, J. Artieda, and J. A. Obeso. Performance of repetitive wrist movements in parkinsons disease. *Brain.*
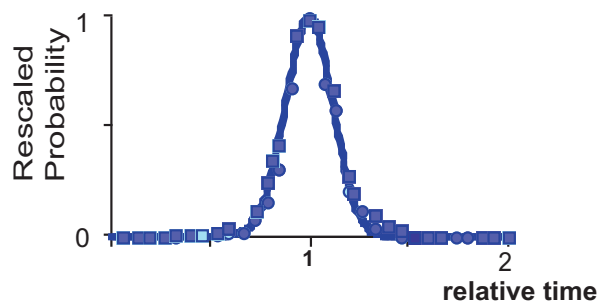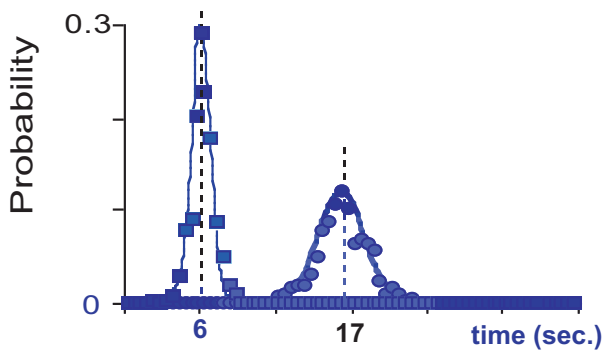
[52] B. Rakitin. Effects of spatial stimulus-response compatibility on choice time production accuracy and variability. *Journal of Experimental Psychology: Human Perception and Performance (in press)*, 2005.

[53] B. C. Rakitin, Y. Stern, and C. Malapani. The effects of aging on time production in delayed free-recall. *Brain and Cognition (in press)*, 2005.

[54] B.C. Rakitin, S.C. Hinton, T.B. Penney, C. Malapani, J. Gibbon, and W.H. Meck. Scalar expectancy theory and peak-interval timing in humans. *Experimental Psychology: Animal Behavior Processes*, 24:1–19, 1998.

[55] A. Renart and X.-J. Wang. A robust biophysical mechanism for scalar timing through derivative feedback. Abstract 768.18, Society for Neuroscience, San Diego, 2005.

[56] J. Reutimann, V. Yakovlev, S. Fusi, and W. Senn. Climbing neuronal activity as an event-based cortical representation of time. *J. Neurosci.*, 24:3295–3303, 2004.

[57] S. Roberts. Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, 7:242–268, 1981.

[58] E.A. Roy, J. Saint-Cyr, A. Taylor, and A. Lang. Movement sequencing disorders in parkinson's disease. *Int. J. Neurosci.*, 73:183–194, 1993.

[59] H.S. Seung. How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. USA*, 93:13339 – 13344, 1996.

[60] H.S. Seung. Amplification, attenuation, and integration. In M.A. Adbib, editor, *The Handbook of Brain Theory and Neural Networks, 2nd edition*, pages 183–215. Springer, New York, 2003.

[61] H.S. Seung, D.D. Lee, B.Y. Reis, and D.W. Tank. Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron*, 26:259–271, 2000.

[62] M. Usher and J.L. McClelland. On the time course of perceptual choice: The leaky competing accumulator model. *Psych. Rev.*, 108:550–592, 2001.

[63] J. Volkmann, H. Hefter, H. W. Lange, and H. J. Freund. Impairment of temporal organization of speech in basal ganglia diseases. *Brain Lang.*, 43:386399, 1992.

[64] J. H. Wearden and B. McShane. Interval production as an analogue of the peak procedure: evidence for similarity of human and animal timing processes. *The Quarterly Journal of Experimental Psychology*, 40B:363–375, 1988.

[65] J.H. Wearden. 'beyond the fields we know...': exploring and developing scalar timing theory. *Behavioural Processes*, 45:3–21, 1999.

[66] H. Wilson and J. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12:1–24, 1972.

[67] A. M. Wing, S. W. Keele, and D. I. Margolin. Motor disorder and the timing of repetitive movements. In J. Gibbon and L. Allan, editors, *Timing and time perception*, pages 183–192. New York: New York Academy of Sciences, 1984.

[68] A.M. Wing and A.B. Kristofferson. Response delays in the timing of discrete motor responses. *Perception and Psychophysics*.

[69] D. Zakay and R. A. Block. Temporal cognition. *Current Directions in Psychological Science*, 6:12–16, 1997.

Acknowledgements

A)

On - On

Probability

0.3

0

6 17

time (sec.)

Rescaled Probability

1

0

1 2

relative time

B)

On - Off

Probability

0.3

0

6 17

time (sec.)

Rescaled Probability

1

0

1 2

relative time

C)

Off - On

Probability

0.3

0

6 17

time (sec.)

Rescaled Probability

1

0

1 2

relative time

D)

Off - Off

Probability

0.3

0

6 17

time (sec.)

24

Rescaled Probability

1

0

1 2
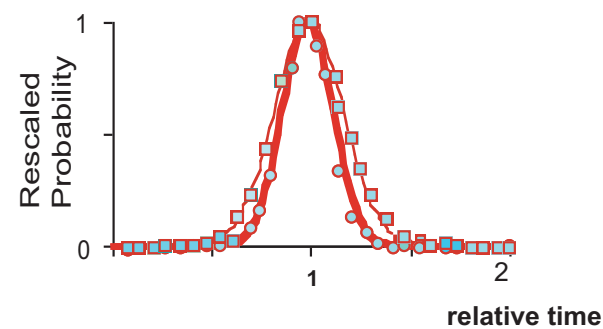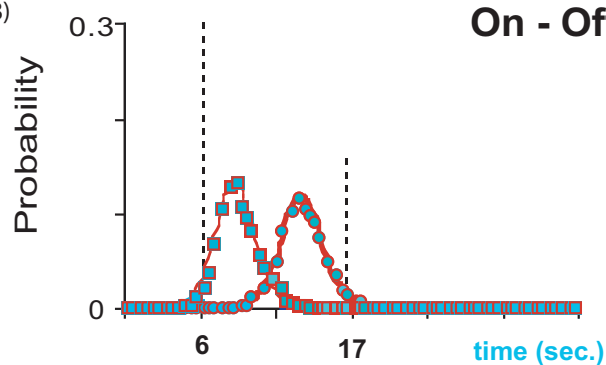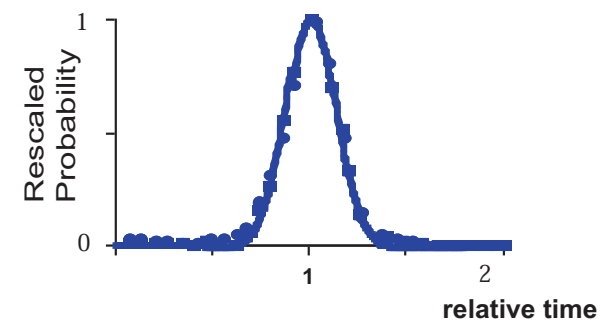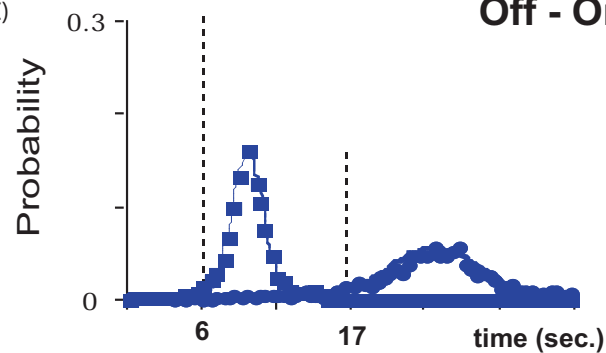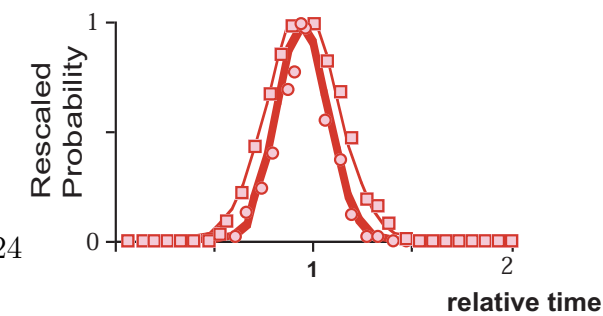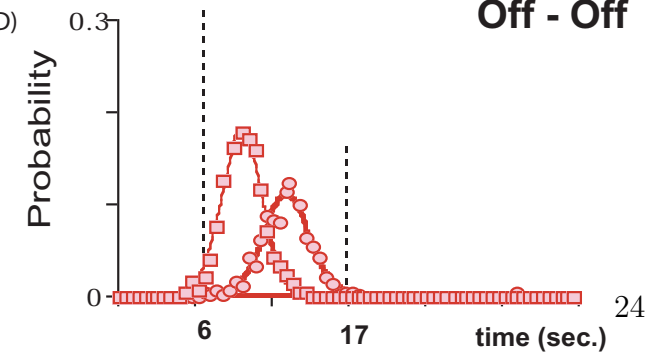
relative time

Figure 1: Distributions of time productions for each experimental group during the testing session, measured in actual laboratory time (left column) and relative time (right column), from [37]. The distributions in relative time are obtained by rescaling experimental time – specifically, the time axis for each subjects distribution is rescaled so that it has the same median as the group-averaged median, and then an average is taken across subjects. The smooth curves are gaussian fits to the data. The ON-ON group (panel A), which received L-dopa in both experimental sessions, exhibits veridical accuracy for both the shorter, 6 sec. interval (thin curve, squares) and longer, 17 sec. interval (heavy curve, circles). Distributions of time productions superpose when rescaled (right), indicating normative, or "scalar," timing when ON medication. Subjects trained ON medication but tested OFF medication without further behavioral feedback (panel B) show migration in the OFF testing condition (left) and violation of scalar timing (right). Panel C shows performance for subjects trained OFF medication but tested ON medication. The ON testing condition shows overestimation for both times (left), but the distributions for the two time values superpose in relative time (right). Panel D shows data for the OFF-OFF group, in which migration again occurs (left) and the scalar property is violated (right).
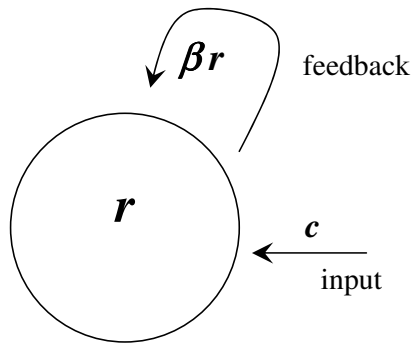

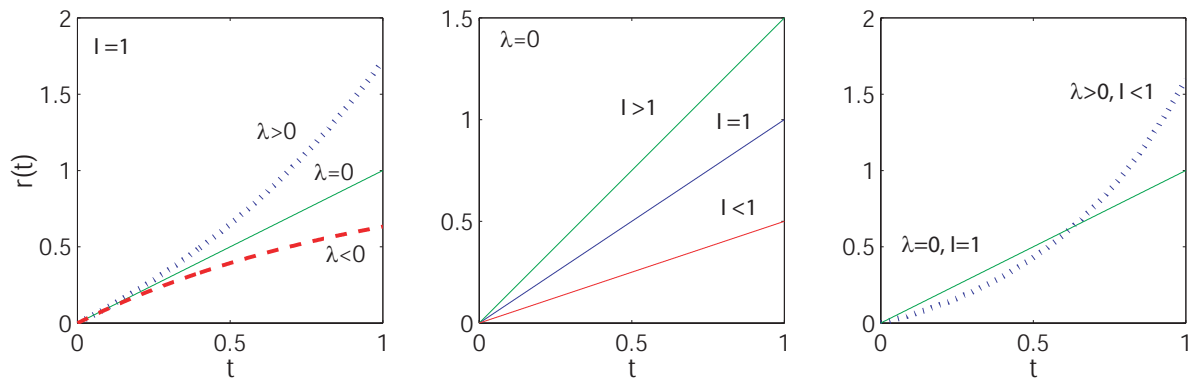Figure 2: Schematic of a recurrently excitatory neural population.

Figure 3: Time course of firing rates $r(t)$, from Eqn. (3). (left) concavity is determined by the sign of λ; (center), initial slopes determined by I ($\lambda = 0$ for all curves shown); (right), firing rates can cross if parameters λ and I are covaried.
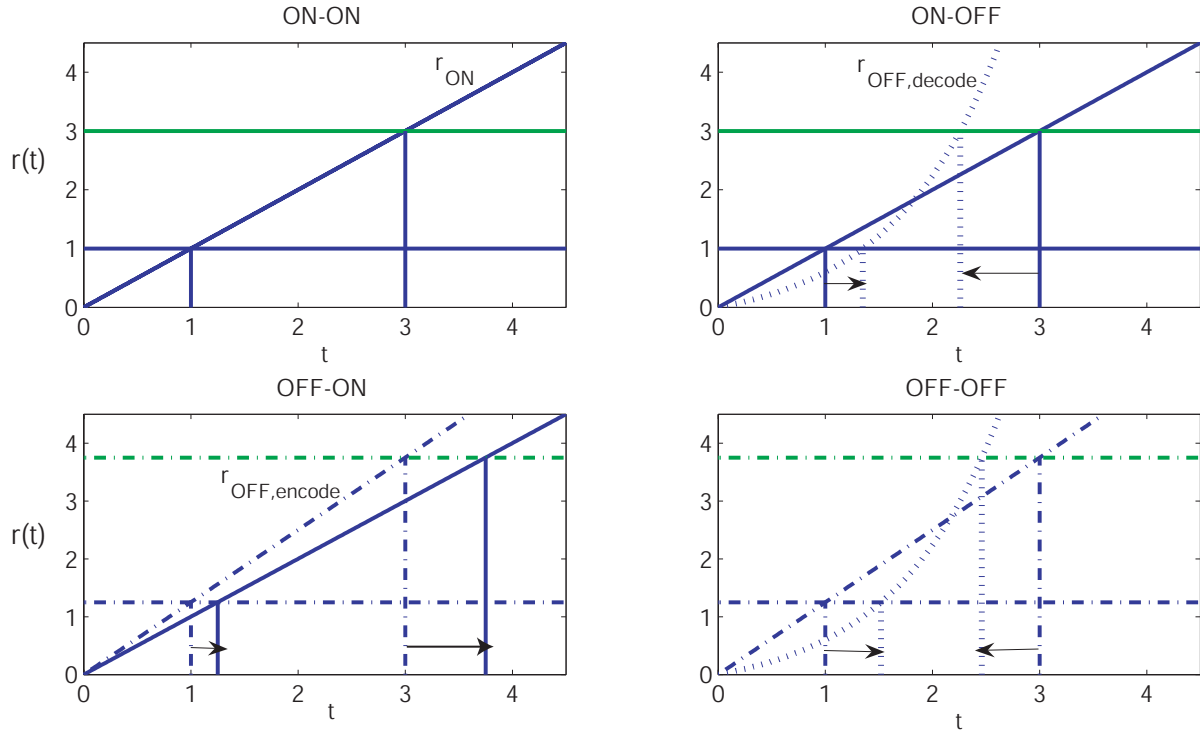
Figure 4: Accumulating firing rates $r(t)$ in the different task protocols. In the ON-ON and ON-OFF protocols, $r_{ON}(t)$ (solid) is used to set thresholds $\bar{\theta}_1$ and $\bar{\theta}_2$ for the two target intervals (1 and 3 time units). In the OFF-ON and OFF-OFF protocols, $r_{OFF,encode}(t)$ (dash-dotted lines) is used to set these thresholds. Following this threshold setting, these thresholds are used in the 'test' stage of the experiment in order to produce time estimates. Depending on the medication state, the times of intersection between either $r_{ON}(t)$ or $r_{OFF,decode}(t)$ (dotted) and the encoded thresholds determine the time estimates produced. Arrows indicate drift from correct values. The horizontal and vertical lines, with same line type as for $r(t)$ during encode, correspond to setting of thresholds for target times.
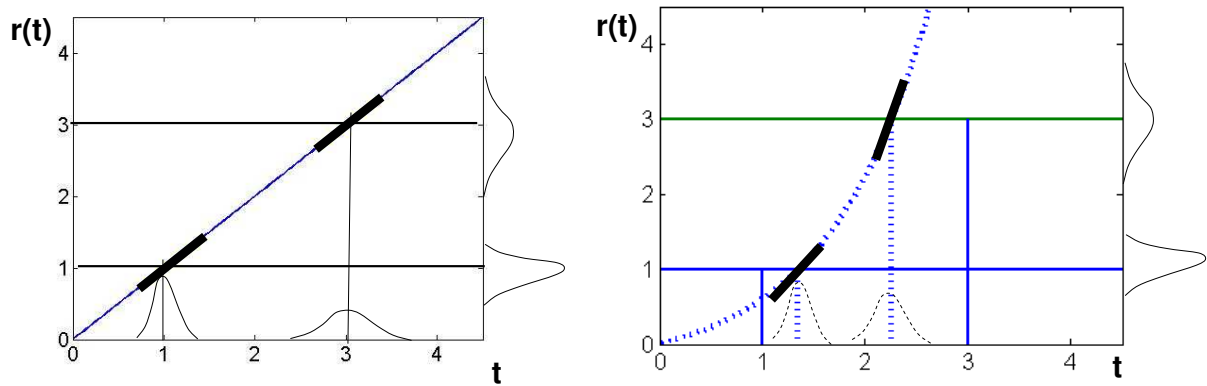
Figure 5: For linear accumulators (left), distributions of thresholds $p_\theta(\theta)$ with standard deviations proportional to their mean are mapped to time estimates $p(t)$ with the same 'scalar' property. However, for curvilinear accumulators (right), the different slope $r'(t)$ in different ranges (shown as thick line segments), produces time estimates that lack the scalar property – here, the distribution of time estimates produced for the longer interval are relatively compressed.
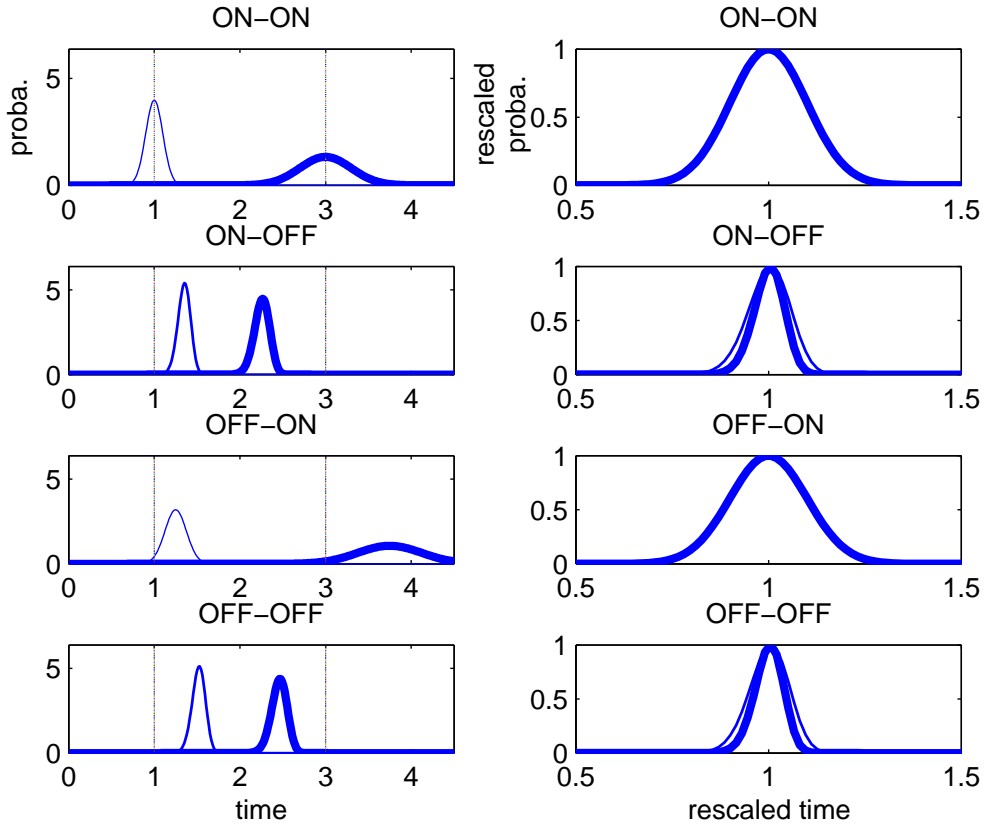
Figure 6: (left) Probability densities of time estimates during the decoding session for the various behavioral contingencies, from model: thin line, shorter target interval $T_1 = 1$, thick line, longer interval $T_2 = 3$. (right) As in the left panels, but measured in relative time; relative time is absolute time divided by the mean of the distribution. In the right panels, vertical axes are also rescaled to give each distribution a maximum value of 1. Whenever time estimates are produced in the OFF condition, the scalar property is violated. In particular, the distribution of estimates for the shorter time interval T_1 is relatively 'too broad' in these cases, as observed in the behavioral data: compare with Fig. 1.
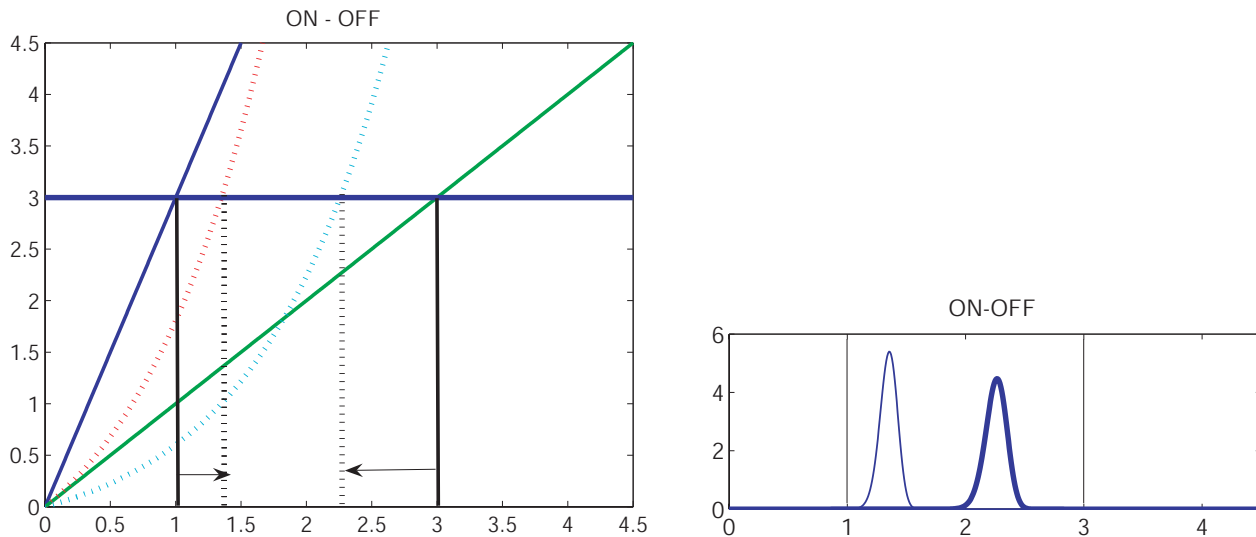
Figure 7: (left), accumulators in the one-threshold model for the ON-OFF condition. Two different accumulators (but a single threshold) are used to time the target times $T_1 = 1$ and $T_2 = 3$. Input values I_1 and I_2 are set so that the accumulators $r_{ON}^1(t)$, $r_{ON}^2(t)$ used for encoding reach the threshold $\bar{\theta}$ at the target times T_1, T_2. Decoding using the accumulator $r_{OFF,decode}(t)$ in the OFF condition results in drift in the median time estimates as indicated by arrows. Solid lines represent $r_{ON}(t)$; dotted, $r_{OFF,decode}(t)$. (right), densities of time productions, which exactly match those for the two-threshold model in the ON-OFF condition (and all other conditions); see text.

Table 1: Clinical Motor and Neuropsychological Profiles of the Four PD Groups (Malapani, Deweer & Gibbon, 2002)

	Group 1 (ON-ON)	Group 2 (ON_OFF)	Group 3 (OFF-ON)	Group 4 (OFF-OFF)
N	6	12	12	6
F/M	2/4	5/7	4/8	3/3
Age	52 (5.2)	56.5 (2.7)	54.7 (3.03)	56.8 (2.9)
Education	4.2 (0.2)	4.08 (0.58)	4.33 (0.4)	4.5 (0.56)
Evolution (years)	10 (2.3)	9.67 (0.89)	10.75 (1.5)	11.3 (2.4)
Hohen and Yahr	3.4 (0.5)	2.9 (0.4)	3.04 (0.2)	2.66 (0.3)
UPDRS-ON	17.4 (3.01)	15.2 (2.5)	13.54 (2.5)	15.8 (1.7)
Tremor	0.4 (0.6)	0.3 (0.7)	0.3 (2.7)	0.4 (1.5)
Akinesia	10.3 (5.6)	9.3 (4.5)	9.7 (2.5)	10.2 (1.5)
Rigidity	6.0 (3.1)	5.1(3.2)	4.9 (2.3)	5.5 (2.1)
UPDRS-OFF	51.2 (5.6)	43.4 (4.1)	42.7 (4.5)	43.5 (3.3)
Tremor	4.4 (3.2)	3.4 (2.3)	3.7 (3.3)	3.1 (2.3)
Akinesia	22.6 (7.3)	19.6 (6.7)	19.8 (6.7)	20.2 (6.2)
Rigidity	11.6 (4.1)	10.8 (3.8)	11.8 (2.8)	11.6 (2.9)
% Improvement	66.2 (3.5)	66.7 (3.6)	70.9 (4.6)	62.8 (4.4)
MMS	29.4 (0.4)	28.8 (0.4)	29.1 (0.2)	29.1 (0.4)
MATTISE				
Global	140.4 (1.2)	140.6 (1.05)	140.7 (0.8)	141.1 (0.9)
Attention	36.4 (0.2)	36.7 (0.1)	36.5 (0.2)	36.3 (0.3)
Initiation	34.2 (1.2)	35.6 (0.5)	35.8 90.5)	36.5 (0.3)
Constancy	5.8 (0.2)	5.9 (0.08)	6 (0)	6 (0)
Concept	38.2 (0.6)	38.3 (0.4)	37.8 (0.4)	38 (0.4)
Memory	24.8 (0.2)	24.3 (0.2)	24.6 (0.18)	24.3 (0.2)
Grober & Booschke				
Encoding	16 (0)	16 (0)	16 (0)	16 (0)
Recognition	16 (0)	15.2 (0.2)	14.7 (0.3)	15.6 (0.3)

[1]

[1] All patients included in the study had bilateral symptoms at the time of testing. Subjects with severely disabling pharmacological side effects, such as involuntary movements ("off" or "on" state dyskinesias) were excluded.