Research Report

A firing rate model of Parkinsonian deficits in interval timing

Eric Shea-Brown\textsuperscript{a,b,*}, John Rinzel\textsuperscript{a,b}, Brian C. Rakitin\textsuperscript{c}, Chara Malapani\textsuperscript{d,e}

\textsuperscript{a}Courant Institute of the Mathematical Sciences, New York University, New York, NY 10012, USA
\textsuperscript{b}Center for Neural Science, New York University, New York, NY 10012, USA
\textsuperscript{c}Cognitive Neuroscience Division, Taub Institute, New York, NY 10032, USA
\textsuperscript{d}Department of Psychiatry, Columbia University, New York, NY 10032, USA
\textsuperscript{e}Division of Biopsyhology, New York State Psychiatric Institute, New York, NY 10032, USA

ARTICLE INFO

Article history:
Accepted 1 October 2005
Available online 17 January 2006

Keywords:
Interval timing
Parkinson’s Disease
Scalar property
Migration
Curvilinear accumulator
Firing rate model

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.

© 2005 Elsevier B.V. All rights reserved.

1. Introduction

This paper is motivated by experiments that revealed specific deficits in the interval timing behavior of Parkinson’s Disease (PD) patients (Malapani et al., 1998b, 2002a,b; Malapani and Rakitin, 2003). In the most developed of these experiments, the timing abilities of PD patients were assessed using the Peak Interval (PI) timing task (Rakitin et al., 1998) immediately after learning a pair of target time intervals with the assistance of behavioral feedback and 24 h later without the benefit of such feedback (Malapani et al., 2002a). 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) (Malapani et al., 2002a). These patterns cannot be immediately reconciled with established models of the timing process, especially Scalar
Expectancy Theory, or SET (Gibbon et al., 1984; Church, 2003; Gibbon, 1992; Gibbon and Malapani, 2002). 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 (Catania, 1970; Church et al., 1994; Roberts, 1981), it was adapted for human use in healthy subjects (e.g., Rakitin et al., 1998, 2005; Hinton and Meck, 2004; Hinton and Rao, 2004; Lustig and Meck, 2001) and patients with neurological disease (e.g., Malapani et al., 1998a,b, 2002a; Lustig and Meck, 2005). Given the high ratio of variance associated with timing vs. motor components (Rakitin, 2005) (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 (Bloxham et al., 1987; Evarts et al., 1981; Malapani et al., 1994), movement time (Benecke et al., 1986; Roy et al., 1993), and speech production time (Lieberman et al., 1992; Volkman et al., 1992), as well as deficits in programming and synchronizing motor responses (Nakamura et al.; Wing et al., 1984; Pastor et al.; OBoyle, 1997). These motor deficits may add a constant (“motor”) variance in timing performance (Wing and Kristofferson), particularly in tasks with substantial motor requirements such as repetitive tapping (Ivy and Keele; Pastor et al.; OBoyle, 1997; Wing et al., 1984).

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 (Malapani et al., 2002a). 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 s) and a shorter (6 s) 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.

Fig. 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, panel 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 (Figs. 1B, 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 modeling 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 (1977) in 1977 and still strongly influential in the field (Gibbon et al., 1984; Church, 2003; Gibbon, 1992; Allan, 1998; Gallistel and Gibbon, 2000; Wearden, 1999). 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 (Church, 2003; Gallistel and Gibbon, 2000) and humans (Rakitin et al., 1998; Gibbon et al., 1997b; Ivy and Hazeltine, 1995; Wearden and McShane, 1988). According to this model, pulses from an internal pacemaker are integrated by an accumulator when attention is directed to time (Lejeune, 1999; Meck and Church, 1983; Zakay and Block, 1997; Rakitin, 2005). The value of the accumulator increases linearly with time, and is stored in memory upon the occurrence of reinforcement or feedback (Gibbon et al., 1984; Jones and Wearden, 2003). A ratio-based decision process then compares values of the accumulator to remembered values in order to
determine responses on future trials (Allan and Gibbon, 1991; Allan and Gerhardt, 2001). The SET model parameters can fluctuate from trial to trial, but their overall statistics are independent of the target time interval. This produces the unidirectional, proportional rightward shift seen in the OFF–ON groups by relative slowing of the pacemaker rate, and explains the fact that most experimental manipulations producing monotonic changes across a range of target intervals (Malapani and Rakitin, 2003). 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 Gibbon and Malapani, 2002 for a more detailed discussion). The migration effect, however, cannot be reconciled with SET or other similar timing theories (e.g., Zakay and Block, 1997). 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 (2003) 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 (1996), 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 modeling (cf. Seung, 1996; Seung et al., 2000; Wilson and Cowan, 1972; Usher and McClelland, 2001; Hopfield, 1984; Abbott, 1991; McClelland, 1979; Grossberg, 1988). 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 Miall, 1996), 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 (Malapani and Rakitin, 2003)—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 (Gibbon, 1992). In contrast, parameter values

<table>
<thead>
<tr>
<th>Table 1 – Clinical motor and neuropsychological profiles of the four PD groups (Malapani et al., 2002b)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>F/M</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Education</td>
</tr>
<tr>
<td>Evolution (years)</td>
</tr>
<tr>
<td>Hohen and Yahr</td>
</tr>
<tr>
<td>UPDRS-ON</td>
</tr>
<tr>
<td>Tremor</td>
</tr>
<tr>
<td>Akinesia</td>
</tr>
<tr>
<td>Rigidty</td>
</tr>
<tr>
<td>UPDRS-OFF</td>
</tr>
<tr>
<td>Tremor</td>
</tr>
<tr>
<td>Akinesia</td>
</tr>
<tr>
<td>Rigidty</td>
</tr>
<tr>
<td>% improvement</td>
</tr>
<tr>
<td>MMS</td>
</tr>
<tr>
<td>MATTISE Global</td>
</tr>
<tr>
<td>Attention</td>
</tr>
<tr>
<td>Initiation</td>
</tr>
<tr>
<td>Constancy</td>
</tr>
<tr>
<td>Concept</td>
</tr>
<tr>
<td>Memory</td>
</tr>
<tr>
<td>Grober and Booschke Encoding</td>
</tr>
<tr>
<td>Recognition</td>
</tr>
</tbody>
</table>

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.
modeling 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 Malapani and Rakitin (2003) in three additional ways. First, we use explicit solutions to our simple firing rate-based model to show the scalar property, with proportionally broader distributions at shorter intervals (in agreement with experimental data).
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., Wilson and Cowan, 1972; Fourcaud and Brunel, 2002; Nykamp and Tranchina, 2000). 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 Experimental procedures 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 Malapani and Rakitin (2003), 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 of timing models that use two distinct mechanisms to time different intervals. We also group parameters according to whether the accumulator results in storage (encoding) or decoding (decoding) of time estimates. The crucial notion is whether the accumulator results in storage (encoding) or decoding (decoding) of time estimates. The crucial notion is whether the accumulator results in storage (encoding) or decoding (decoding) of time estimates. The crucial notion is whether the accumulator results in storage (encoding) or decoding (decoding) of time estimates.

Specifically, we will show that the parameters of the accumulator (12) must change as follows among experimental conditions to reproduce experimentally observed effects on both mean time estimates and the scalar property:

\[ 0 = \theta_{\text{ON}} \leq \theta_{\text{OFF, encode}} < \theta_{\text{OFF, decode}}. \]  

\[ I_{\text{OFF, decode}} < I_{\text{ON}} \leq I_{\text{OFF, encode}}. \]

where at least one of the \( \leq \) relations must be a strict inequality. We specify a single set of parameters \((\lambda_{\text{ON}}, I_{\text{ON}})\) 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 \( \lambda \) 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 Eqs. (1)–(2) as well as the value \( \lambda_{\text{ON}} = 0 \), are the only choices that could have been made.

2.1. Trends in time estimates for the various experimental conditions

In our model, the initial slope of the firing rate \( r(t) \) is 1. Therefore, our parameter choices (1)–(2) guarantee that, for all sufficiently early times, the firing rates for the different experimental contingencies are ordered as follows:

\[ r_{\text{OFF, decode}}(t) < r_{\text{ON}}(t) < r_{\text{OFF, encode}}(t). \]  (3)

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

\[ r_{\text{ON}}(t) < r_{\text{OFF, encode}}(t) < r_{\text{OFF, decode}}(t). \]  (4)

Following Gibbon (1992), 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. Fig. 2 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 (3) holds for the short interval \( T_1 = 1 \), while Eq. (4) is valid for the long interval \( T_2 = 3 \).

Fig. 2 also demonstrates the trends in time production that result from the inequalities (3)–(4). 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 \( r_{\text{ON}}(t) \) are later decoded using \( r_{\text{OFF, decode}}(t) \) (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 \( r_{\text{OFF, encode}}(t) \) and then decoded into time productions via \( r_{\text{ON}}(t) \) and \( r_{\text{OFF, decode}}(t) \), 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 Fig. 2 and all subsequent figures, we chose \( I_{\text{ON}} = 1, \theta_{\text{OFF, encode}} = 1.25, \theta_{\text{OFF, decode}} = 0.35 \) (as per Eq. (1)), and \( \lambda_{\text{ON}} = 0, \lambda_{\text{OFF, encode}} = 0, \lambda_{\text{OFF, decode}} = 1 \) (all in accordance...
2.1.3. The scalar property and its violation

Throughout this paper, we adopt Gibbon's (1992) 'ratio rule' for determining thresholds. 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.

To see this, let $p_\theta(t)$ 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 $p(t)dt = dr(t) p_\theta(r(t))$, or

$$p(t) = \frac{dr(t)}{dt} p_\theta(r(t)).$$

(5)

For linear accumulators, $r'(t) = 1$ (where $' = (d/dt)$), 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 Fig. 3). 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 Fig. 4. 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 (Malapani et al., 2002a).

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 be remedied by assuming a broader distribution of thresholds in the OFF vs. ON L-dopa condition (specifically, with twice the standard deviation).
2.1.4. Necessity of parameter settings

We now show that the parameter values and orderings of Eqs. (1)–(2) 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 \( > I_{ON} \).

Fig. 3 – 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 is relatively compressed.

Fig. 4 – (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.
This follows from the fact that "other words, the OFF distributions will have a greater mates will be more skewed towards shorter times for all derived above, also predicts that distributions of time esti-
to the OFF – underestimation of the longer interval in the ON group (Malapani et al., 2002a), as statistical reanalysis of these data confirms.

We remark additionally that the “less than or equals” ambiguity in the relations λ_{OFF, encode} ≤ λ_{ON} and λ_{OFF, decode} ≤ λ_{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 λ_{OFF, encode} = 0 = λ_{ON} and hence λ_{OFF, decode} > λ_{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.

2.1.5. 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 Eqs. (1)–(2). 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 Eqs. (1)–(2), 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 ON–OFF 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 (Malapani et al., 2002a), as statistical reanalysis of these data confirms.

In particular, overestimation of the short interval is significantly higher in the OFF–OFF relative to the ON–OFF group (P < 0.005; F = 19.6) and likewise for the OFF–OFF relative to the ON–OFF condition (P < 0.05; 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 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 λ_{OFF, decode} > 0, which leads to rates r_{OFF, decode}(t) which accelerate in time and hence, via Eq. (5), to a compression of the tail of the distribution correspond-
ing to longer estimates. However, statistical analyses have not found this trend to be reliably present in the existing experimental data.

2.2. 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:

\[ p_{\hat{h}}(h) = \frac{1}{\sqrt{2pk^2\lambda}} \exp \left( -\frac{(h-h^*)^2}{2k^2\lambda} \right), \]

where \( k \) is the proportionality constant. Rewriting, this is

\[ p_{\hat{h}}(h) = \frac{1}{h \sqrt{2pk^2}} \exp \left( -\frac{(h/h^*)^2}{2k^2} \right) = \frac{1}{h} q \left( \frac{h}{h^*} \right). \]

Inserting this latter expression into Eq. (5), we have

\[ p(t) = r(t) \frac{1}{h} q \left( \frac{r(t)}{h} \right). \]

Then, substituting in solution (13) gives

\[ p(t) = \frac{1}{h} \exp(kt) q \left( \frac{1}{h} \left( \frac{\exp(kt) - 1}{k} \right) \right). \]

This expression reveals that densities of time estimates will be Gaussian only when \( \lambda = 0 \), and will otherwise be skewed. Additionally, Eq. (8) also shows that the densities of response times predicted by our model depend on only two quantities: \( \lambda \) 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.

2.2.1. 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 (Leon and Shadlen, 2003; Reutimann et al., 2004). 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 \( \hat{\theta} \), but two different values of \( I \) (one for each of the two intervals being timed) which are separately adjusted during encoding so that \( r(T) = \hat{\theta} \) at each target time \( T \).

As an example, Fig. 5 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 \( \bar{\theta} = 1 \). Fig. 5 (right) also shows the resulting distribution of time estimates, which exactly match those for the corresponding two-threshold model (compare with Fig. 4). 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 $T$, $r(T) = \theta$. Substituting this into Eq. (13) gives

$$
\frac{I}{h} = \frac{k}{e^{\frac{T}{\theta} - 1}}. 
$$

(9)

We emphasize that this relationship holds in both the one-threshold and two-threshold models: although the values of $I$ and $\theta$ may be different in each case, their ratio will be the same for fixed values of $I$ and $T$. This is a consequence of the simple form of Eq. (13), 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 $(I/\theta)$ is decoded to produce time estimates in the one- and two-threshold models. As discussed above, $I$ and $\theta$ 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 $(I/\theta)$ 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 $\theta$ is the same for decoding in both the one-threshold and two-threshold models, then we see that both values $(I/\theta)$ and $\lambda$, which completely determine the distribution of time productions $p(t)$ via Eq. (8), 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 Eqs. (1)-(2), as well as the predictions for behavioral data that follow from these settings.

2.2.2. An equivalent hypothesis on parameter variations between task conditions

Above, we allowed the parameters $\lambda$ 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 $\theta$ between these experimental conditions. That is, we assumed that $\theta$ is determined via Eq. (9) during the encode process, regardless of medication state, and that it is exactly this same value of $\theta$ which is later used for decoding, again regardless of medication state. However, another possibility is that $\theta$ changes between experimental condition instead of $I$, while $\lambda$ continues to vary as above. These two possibilities are equivalent if the corresponding parameters are varied inversely, because Eq. (8) shows that only the ratio $(I/\theta)$ 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 (Gibbon, 1977), 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 $\theta = br(T)$, and that during decoding, time estimates are made when $r(t) = (\theta/b)$, where the value of $b$ is
specific to task condition (encode vs. decode, OFF vs. ON). In this case, the parameter ordering

\[ b_{\text{OFF, decode}} < b_{\text{ON}} \leq b_{\text{OFF, encode}} \]  

(10)

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

3. Discussion

In this paper, we have shown how a threshold-based mechanism for interval timing, similar to that of scalar expectancy theory (Gibbon et al., 1984; Church, 2003) 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 (Malapani et al., 2002a). 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 4. 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: \( \lambda \), the net neural feedback, and \( I \) the external drive to the population. These parameters must vary among experimental conditions as in Eqs. (1)–(2) 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 Experimental procedures section, we rescaled and grouped parameters to define the neural feedback \( \beta = \beta g - 1 \) and the neural drive \( I = gc \), which completely determine the dynamics of our model. Here, \( \beta \) 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 Eqs. (1)–(2), we can assess how these values may vary in order to affect the required changes in \( \lambda \) and \( I \) among task conditions. For example, the transition from ON-drug encoding to OFF-drug decoding (in which \( \lambda \) increases and \( I \) decreases) could be caused by an increase in the weight \( \beta \) with a (smaller) decrease in gain \( g \), by an increase in gain \( g \) with a (greater) decrease in external inputs \( c \), or by other covariances in parameters. Any of these choices result 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 Seung (1996, 2003), we note that the single firing rate \( r \) may be thought of as characterizing a distinguished subpopulation of cells whose activity increases along 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., Seung, 1996; Seung et al., 2000; Goldman et al., 2003). However, as Seung (1996) and references therein have pointed out, this neural integrator property requires fine tuning of parameters (i.e., setting \( \beta \) 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 (Koulakov et al., 2002; Goldman et al., 2003). While these papers do explore some cases in which these properties enhance the robustness of neural integration, it 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 do provide evidence for linearly accumulating firing rates (e.g., Reutimann et al., 2004; Leon and Shadlen, 2003) or linearly accumulating firing rates in concert with other firing patterns (Matell et al., 2003). 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 (Gibbon et al., 1984; Gibbon, 1992). However, we believe that our model would give similar results if, as proposed by Renart and Wang (2005), 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 modeled 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 Malapani et al. (2002a), 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, for tasks in which only a single interval must be encoded and reproduced, each duration is typically overestimated (Malapani et al., 1998b). 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, \( r_{\text{OFF, decode}}(t) \) and \( r_{\text{OFF, decode}}(t) \), 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 Malapani et al., 1998b), 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 (Ivry, 1996; Gibbon et al., 1997a).

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 that produced OFF \( l \)-dopa should display relatively greater leftward skew (or less rightward skew) relative to comparable distributions that 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 Malapani and Rakitin, 2003). 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 is observed.

4. Experimental procedure

4.1. A self-excitatory firing rate model of interval timing

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

\[
\frac{dr}{dt} = -r + f(br + c),
\]

(11)

where \( r \) 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, \( \beta \) is the strength of excitatory recurrent connections, and \( c \) is the strength of inputs from areas external to the population itself (see Fig. 6). 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 modeling of, e.g., Matell and Meck, 2000; Gibbon et al., 1997b; Contreras-Vidal and Schultz, 1990, as well as architecturally-based
models that are not directly related to timing, Graybiel and Kimura, 1995; Frank et al., 2001).

We further assume that the net input \( \tilde{r} + r \) remains in a range where \( f() \) may be approximated by the linearization \( f(\delta) = \delta g \) where \( g \) is the ‘gain’ of the population, so that Eq. (11) becomes

\[
\frac{dr}{dt} = kr + I. \tag{12}
\]

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

We take initial conditions \( r(0) = 0 \) for Eq. (12), giving the solution

\[
r(t) = \frac{I}{K} \exp(\lambda t) - 1 \tag{13}
\]

Fig. 7 displays the trajectories that this solution takes for different values of the parameters \( \lambda \) and \( I \), demonstrating that each plays a distinguished role: \( \lambda \) 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 \( \theta \) is the time estimate on that trial. To time different intervals, there are two possibilities: (i) the parameters \( \lambda \) and \( I \) could differ or (ii) the thresholds \( \theta \) 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.

**Acknowledgments**

The authors thank Peter Balsam for pointing out trends in the magnitude of overestimation and underestimation effects mentioned in the main text as well as additional insights, and Stephen Fairhurst and Randy Gallistel for other contributions. E.B. was supported by a NSF Mathematical Sciences Postdoctoral Research Fellowship. We received support from grants R01-MH 54793, awarded to C. M., and AG K01-000991 from the National Institute on Aging, awarded to B.R.

**References**


Lejeune, H., 1999. Switching or gating? The attentional challenge
Leon, M., Shadlen, M., 2003. Representation of time by neurons in
Lieberman, P., Kako, E., Friedman, J., Tajchman, G., Feldman, L.S.,
Jimenez, E.B., 1992. Speech production, syntax comprehension,
Lustig, C., Meck, W., 2001. Paying attention to time as one gets
older. Psychol. Sci. 12, 478–484.
induces deficits in working memory and feedback effects of
Malapani, C., Rakitin, B., 2003. Interval timing in
the dopamine-depleted basal ganglia: from empirical data to
timing theory. In: Meck, W.H. (Ed.), Functional and Neural
Mechanisms of Interval Timing. CRC Press, Boca Raton, FL.
Malapani, C., Pillon, B., Dubois, B., Agid, Y., 1994. Impaired
simultaneous cognitive task performance in parkinsons
Malapani, C., Dubois, B., Rancurel, G., Gibbon, J., 1998a. Cerebellar
dysfunctions of temporal processing in the seconds range in
humans. NeuroReport 9, 3907–3912.
Malapani, C., Rakitin, B., Levy, R., Meck, W.H., Deweer, B., Dubois,
B., Gibbon, J., 1998b. Coupled temporal memories in Parkin-
Neurosci. 10, 316–331.
Malapani, C., Deweer, B., Gibbon, J., 2002a. Separating storage from
retrieval dysfunction of temporal memory in Parkinsons
Malapani, C., Rakitin, B.C., Fairhurst, S., Gibbon, J., 2002b.
Matell, M.S., Meck, W.H., 2000. Neurophysiological mechanisms of
Matell, M., Meck, W., Nocolelis, M., 2003. Interval timing and the
McClelland, J.L., 1979. On the time relations of mental processes:
an examination of systems of processes in cascade. Psychol.
Meck, W.H., Church, R.M., 1983. A mode control model of counting
(Eds.), Time, Internal Clocks, and Movement. Elsevier
Science.
Nakamura, R., Nagasaki, H., Harabayashi, H. Disturbances of
Characteristics of tapping response to the period signals. Brain
Lang.
Nykamp, D., Tranchina, D., 2000. A population density approach
that facilitates large-scale modeling of neural networks:
analysis and application to orientation tuning. J. Comp.
Neurosci. 8, 19–50.
O’Boyle, D.J., 1997. On the human neuropsychology of timing of
simple repetitive movements. In: Bradshaw, C.M.,
Szabadi, E. (Eds.), Time and Behaviour: Psychological and
Neurobehavioural Analyses. North-Holland/Elsevier,
Amsterdam, pp. 459–515.
Pastor, M.A., Jahanshahi, M., Artieda, J., Obeso, J.A. Performance of
repetitive wrist movements in parkinsons disease. Brain.
compatibility on choice time production accuracy and
685–702.
Rakitin, B.C., Hinton, S.C., Penney, T.B., Malapani, C., Gibbon, J.,
timing in humans. Exp. Psychol., Anim. Behav. Processes 24,
1–19.
Rakitin, B.C., Stern, Y., Malapani, C., 2005. The effects of aging
on time production in delayed free-recall. Brain Cogn. 58,
17–34.
Society for Neuroscience, San Diego.
Reutimann, J., Yakovlev, V., Fusi, S., Senn, W., 2004. Climbing
neuronal activity as an event-based cortical representation of
Roberts, S., 1981. Isolation of an internal clock. J. Exp. Psychol.,
73, 183–194.
Seung, H.S., 2003. Amplification, attenuation, and integration, In:
Adibib, M.A. (Ed.), The Handbook of Brain Theory and Neural
memory of eye position in a recurrent network of
Usher, M., McClelland, J.L., 2001. On the time course of perceptual
Impairment of temporal organization of speech in basal
Wearden, J.H., 1999. ‘Beyond the fields we know...’: exploring and
Wearden, J.H., McShane, B., 1988. Interval production as an
analogue of the peak procedure: evidence for similarity of
human and animal timing processes. Q. J. Exp. Psychol. 40B,
363–375.
Wilson, H., Cowan, J., 1972. Excitatory and inhibitory interactions
in localized populations of model neurons. Biophys. J. 12,
1–24.
Wing, A.M., Kristofferson, A.B., Response delays in the timing of
Wing, M., Keele, S.W., Margolin, D.I., 1984. Motor disorder and the
timing of repetitive movements. In: Gibbon, J., Allan, L. (Eds.),
Timing and Time Perception. New York Academy of Sciences,
Sci. 6, 12–16.