Psychological Review
2014, Vol. 121, No. 3, 422–462

© 2014 American Psychological Association
0033-295X/14/$12.00 http://dx.doi.org/10.1037/a0037012

The Ising Decision Maker: A Binary Stochastic Network for Choice Response Time

Stijn Verdonck and Francis Tuerlinckx
KU Leuven, University of Leuven

The Ising Decision Maker (IDM) is a new formal model for speeded two-choice decision making derived from the stochastic Hopfield network or dynamic Ising model. On a microscopic level, it consists of 2 pools of binary stochastic neurons with pairwise interactions. Inside each pool, neurons excite each other, whereas between pools, neurons inhibit each other. The perceptual input is represented by an external excitatory field. Using methods from statistical mechanics, the high-dimensional network of neurons (microscopic level) is reduced to a two-dimensional stochastic process, describing the evolution of the mean neural activity per pool (macroscopic level). The IDM can be seen as an abstract, analytically tractable multiple attractor network model of information accumulation. In this article, the properties of the IDM are studied, the relations to existing models are discussed, and it is shown that the most important basic aspects of two-choice response time data can be reproduced. In addition, the IDM is shown to predict a variety of observed psychophysical relations such as Piéron's law, the van der Molen-Keuss effect, and Weber's law. Using Bayesian methods, the model is fitted to both simulated and real data, and its performance is compared to the Ratcliff diffusion model.

Keywords: choice response time, statistical mechanics, diffusion models, stochastic Hopfield network, speed–accuracy tradeoff

Supplemental materials: http://dx.doi.org/10.1037/a0037012.supp

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

The speeded two-choice response time (RT) task is a well-established paradigm in experimental psychology for investigating the principles underlying simple decision making. In the psychological literature, several successful models have been proposed based on the idea of the accumulation of noisy evidence over time (Link & Heath, 1975; Ratcliff, 1978; Ratcliff & Smith, 2004; Stone, 1960; Usher & McClelland, 2001; Vickers, 1970). An important class of accumulator models, of which the drift diffusion model is the prime example, relies on a single or a few linear stochastic differential equations (SDEs). Decades of careful research resulted in excellent fits between the best accumulator models and behavioral data from speeded two-choice RT tasks. Initially, these models were conceived as abstract representations of the decision process. In the last decade however, there has been an increasing trend of investigating their neurophysiological underpinnings (Ditterich, 2010; Gold & Shadlen, 2007; Niwa & Ditterich, 2008; Ratcliff, Cherian, & Segraves, 2003; Ratcliff, Hasegawa, Hasegawa, Smith, & Segraves, 2007; Rorie, Gao, McClelland, & Newsome, 2010; Smith & Ratcliff, 2004).

More recently, the speeded two-choice RT task has raised interest in the field of computational neuroscience. In this field, models are usually defined at a basic neurocomputational level, providing a microscopic description of individual neurons and their interactions as they combine into a neural network. The type of network typically used when modeling human brain dynamics is the so-called integrate-and-fire network, which closely matches the experimentally observed pulse-firing-based neuronal communication (Amit & Brunel, 1997). Wang (2002) used this framework to design a network specifically for the speeded two-choice RT task. In their analysis, Wong and Wang (2006) reduced this high-dimensional network to fewer dimensions, eventually leading to a set of two nonlinear SDEs, which connects back to the diffusion models mentioned earlier. Interesting work on the reduction of networks to one-dimensional nonlinear SDEs has been done by Roxin and Ledberg (2008). Although sometimes referred to in psychological literature (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006), the Wong and Wang integrate-and-fire network

Stijn Verdonck and Francis Tuerlinckx, Faculty of Psychology and Educational Sciences, KU Leuven, University of Leuven.

The research leading to the results reported in this article was sponsored in part by Belgian Federal Science Policy within the framework of the Interuniversity Attraction Poles program (IAP/P7/06), by Grant GOA/15/003 from the KU Leuven, and by Grants G.0534.09 and G.0806.13 from the Fund of Scientific Research Flanders. We are indebted to Han van der Maas for constructive feedback and to Kristof Meers for his excellent assistance in computing and coding. We would also like to thank Martijn Mulder for providing his at that time unpublished data. For the simulations, we used the infrastructure of the VSC–Flemish Supercomputer Center, funded by the Hercules Foundation and the Flemish Government–Department of Economy, Science, and Innovation (EWI).

The code used to obtain the results presented in the article is hardware specific (HPC with GPU accelerators) and will not run on another hardware configuration without modifications. However, documented IDM simulation software for all combinations of CPU/NVIDIA GPU, R/Matlab, Windows/Linux, and related tools will be shared, as they become available, on the following website: http://ppw.kuleuven.be/okp/idm/

Correspondence concerning this article should be addressed to Stijn Verdonck, Faculty of Psychology and Educational Sciences, KU Leuven, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium. E-mail: stijn.verdonck@ppw.kuleuven.be

and its two-dimensional nonlinear derivates (see also Eckhoff, Wong-Lin, & Holmes, 2011) are rarely (if ever) fit to real behavioral data. According to us, one of the reasons for this restraint is that these networks are analytically overly complex and therefore not sufficiently developed as genuine candidates for practical choice RT modeling. Consequently, there seems to be a need for a network approach that is analytically less intricate than the integrate-and-fire two-choice decision network but still preserves all essential nonlinear properties of multiple attractor networks that cannot be emulated by simple linear SDE models.

In this article, we present a more abstract alternative to the above mentioned integrate-and-fire two-choice decision network. Our starting point is the relatively simple Hopfield network (Hopfield, 1982) or formally equivalent dynamic Ising model (Amit, Gutfreund, & Sompolinsky, 1985a,1985b). This network model is based on binary neurons that evolve through time according to an abstract dynamical transition rule. It stands in contrast with the integrate-and-fire network, in which the individual neuron model and its time evolution are inspired by more detailed electrochemical principles. We call our specific application of the dynamic Ising model the Ising Decision Maker (IDM). The basic architecture consists of two pools of neurons, with pairwise excitatory connections within a pool and pairwise inhibitory connections between pools. Like Wong and Wang's network, the IDM operates as a multiple attractor network. When it comes to accurately describing neural dynamics, an integrate-and-fire network is generally considered to be more appropriate because its neurons and their dynamics are more closely related to the biophysiological reality. However, features typical for multiple attractor networks in general can already be reproduced by a simple Ising model (e.g., multiple stable states). Moreover, both the integrate-and-fire network and the Ising model are used for fitting observed correlations between neurons, leading to comparable results (Cocco, Leibler, & Monasson, 2009; Tkacik, Schneidman, Berry, & Bialek, 2009). From this perspective, the Ising model supplied with a suitable dynamics could prove to be a sufficient description of the neurocomputational mechanics underlying the speeded two-choice RT task, even despite the obvious loss of biophysiological detail compared to the integrate-and-fire approach. The IDM is an abstract model of information processing that gives expression to the microscopic biological reality of correlated stochastic neurons in a statistical mechanical framework. In this sense, the IDM follows the recommendation formulated by Forstmann, Wagenmakers, Eichele, Brown, and Serences (2011), p. 276: "Abstract models should ideally start to incorporate assumptions about the neural substrate."

In what follows, we first formulate the basic principles of the IDM at an equilibrium level and then the associated dynamics. Next, we do a formal comparison with existing SDE-based psychological models. Subsequently, we show through simulations that the model is able to reproduce all qualitative aspects of speeded two-choice RT tasks currently captured by traditional diffusion models. Using Bayesian methods, we then fit the IDM to both simulated data (to investigate recovery) and real data. The IDM is compared to different versions of the Ratcliff diffusion model in a model-selection context. Finally, the IDM is shown to explain some additional phenomena, like Piéron's law (with increasing, but weak, stimulus intensity, the choice RT decreases; see, e.g., Pins & Bonnet, 1996), the van der Molen-Keuss effect (with increasing, but very large, stimulus intensity, the choice RT increases again; see Jaśkowski & Włodarczyk, 2006; van der Molen & Keuss, 1979), and a decision-level version of Weber's law.

IDM: The Ising Decision Maker

In this section, we start by explaining the general stochastic Hopfield network (Amit et al., 1985a,1985b; Hopfield, 1982). After having discussed the basic concepts of this type of network, we tailor a version providing a basic neurocomputational description of the speeded two-choice RT task and elaborate this into a workable model.

Stochastic Hopfield Network

The stochastic Hopfield network or Ising model[1] consists of N binary neurons. Each neuron i has an activity S_i, which can be either 0 (inactive) or 1 (active). The activities for all N neurons can be collected in a microscopic state vector $\boldsymbol{S} = (S_1, \ldots, S_N)$. The defining quantity of the network is the energy function $E(\boldsymbol{S})$, which associates an energy value with every one of the 2^N possible configurations of $\boldsymbol{S} = (S_1, \ldots, S_N)$:

$$E(\boldsymbol{S}) = -\sum_{i<j} w_{ij} S_i S_j + \sum_i \theta_i S_i \tag{1}$$

with parameters w_{ij} and θ_i determining the structure of the network. The weight parameter w_{ij} describes the interaction strength between neurons i and j, and θ_i describes the activity threshold for neuron i. In the first term of Equation 1, the summation runs over all distinct pairs of interacting neurons (thus, $i < j$).[2] As an example, a three-neuron version of this network is graphically represented in Figure 1.

The network is stochastic, which means that $\boldsymbol{S}$ is a binary random vector. For the equilibrium distribution of the microscopic states, the canonical ensemble is assumed. The term *canonical ensemble* has roots in statistical mechanics (e.g., Schroeder, 1999), but for all practical purposes, it means that the probability distribution of $\boldsymbol{S}$, $\Pr(\boldsymbol{S})$, is a Boltzmann distribution,

$$\Pr(\boldsymbol{S}) = Z^{-1} e^{-\beta E(\boldsymbol{S})}, \tag{2}$$

in which Z is the so-called partition function:

$$Z = \sum_{\boldsymbol{S}} e^{-\beta E(\boldsymbol{S})}. \tag{3}$$

The role of Z is that of a normalization factor, such that $\Pr(\boldsymbol{S})$ is a valid probability mass function that sums to one over all possible configurations of $\boldsymbol{S}$. The parameter $\beta \geq 0$ is traditionally referred to as thermodynamic beta or inverse temperature. If it is close to infinity, the system will be forced into those configurations with absolute lowest energy; as it decreases, the system can also occupy energetically less favorable states. At $\beta = 0$, all 2^N microscopic states are equally probable.

[1] A simple statistical mechanical model based on binary units was first explored by Ising (1925) (following an idea of his supervisor, Lenz), in the context of ferromagnetism. We have chosen to name our model the Ising Decision Maker to make the formal and historical link with statistical physics more explicit. Many of the tools we use in this article have been developed in this branch of physics, and it seems reasonable to assume that the large body of available literature could fuel further advances.

[2] Note that an alternative formulation uses activity values of -1 and 1; this leads to a formally identical energy function but with redefined w_{ij} and θ_i.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

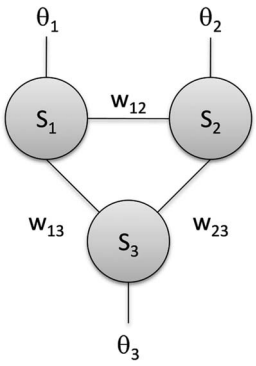


Figure 1. Graphical representation of a stochastic Hopfield net or Ising model with three neurons.

The description up until this point does not contain any information about the way the stochastic vector $\boldsymbol{S}$ evolves through time. In most cases, a sequential Metropolis algorithm is used (e.g., Newman & Barkema, 1999). Every Markov chain time step, a new candidate configuration $\boldsymbol{S}^*$ is constructed by randomly selecting a single neuron and changing its activity (0 to 1 and 1 to 0). The system will evolve from $\boldsymbol{S}$ to $\boldsymbol{S}^*$ with probability π:

$$\pi = \min\{e^{-\beta(E(\boldsymbol{S}^*)-E(\boldsymbol{S}))}, 1\}. \tag{4}$$

In the long run, the Metropolis algorithm leads to the sampling of $\boldsymbol{S}$ according to the Boltzmann distribution from Equation 2.

The IDM Design

We now construct a specific instance of the stochastic Hopfield network representing the supposed abstract neural correlate of speeded two-choice perceptual decision making. The N available neurons are divided into two disjunct pools that are self-exciting and mutually inhibiting. In addition, each pool is excited by a different aspect of an external stimulus through the interaction with a two-component external field that represents the relevant sensory data passed on by earlier predecision systems more directly connected to the physical stimulus.

Equilibrium description. Consider two pools $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ of N_1 and N_2 neurons, respectively, with $N_1 + N_2 = N$. An individual neuron from pool p ($p = 1, 2$) is represented as S_{pi} (with $i = 1, \ldots, N_p$). It is assumed that all neurons in pool p have an equal threshold $\theta_p > 0$. In addition, we define a positive time-varying two-component external field input ($b_1(t)$, $b_2(t)$), each component supplying evidence for the related pool and choice alternatives. In the absence of a stimulus, the external field is (0, 0). When a stimulus is presented, the field changes to some stimulus-specific two-component value characterizing the stimulus shown. Although one could consider stimuli gradually changing through time (e.g., Ratcliff & Rouder, 2000), we stick to the basic speeded choice experiment where a stimulus is either present or not. At any given time t, we consider the external field ($b_1(t)$, $b_2(t)$) to be either (0, 0) or some stimulus-specific two-component value.

All neurons are positively correlated with their fellow pool members, and the strength of these excitatory connections is given by w_p^+ (with $w_p^+ > 0$). Between the two pools, neurons are negatively correlated with $-w_{12}^-$ the strength of the inhibitory connections (meaning $w_{12}^- > 0$).

With all neurons interconnected, the total energy of the two-pool decision network, for a configuration $\boldsymbol{S} = (\boldsymbol{S}_1, \boldsymbol{S}_2)$, is

$$\begin{aligned} E(\boldsymbol{S}) = &-w_1^+ \sum_{i<j} S_{1i}S_{1j} - w_2^+ \sum_{i<j} S_{2i}S_{2j} + w_{12}^- \sum_{i,j} S_{1i}S_{2j} \\ &-(b_1 - \theta_1)\sum_i S_{1i} - (b_2 - \theta_2)\sum_i S_{2i}, \end{aligned} \tag{5}$$

where the time dependence of the external field is implicit. In the first two terms of Equation 5, the summation runs over all distinct pairs of interacting neurons inside a single pool (limiting the summation to distinct pairs implies taking $i < j$). Using constant interaction strengths w_1^+, w_2^+, and $-w_{12}^-$ makes all neurons inside a pool interchangeable without consequence for the system's energy. The use of constant interaction strengths can therefore be seen as an effective mean field approximation (see, e.g., Schroeder, 1999). In the context of a mean field approximation, the assumption of full connectivity is no longer relevant and can easily be departed from by lowering the interaction strengths. A graphical representation of the IDM network design is shown in Figure 2.

Next, let us make the transition from this microscopic level of description to a macroscopic level. For this purpose, we define a macroscopic vector variable $\boldsymbol{y} = (y_1, y_2)$ that represents the average activity of the neurons in each pool. This vector of mean activities $\boldsymbol{y} = (y_1, y_2)$ is a vector-valued function $\boldsymbol{f}$ of the total microscopic configuration $\boldsymbol{S}$: $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{S}) = (f_1(\boldsymbol{S}), f_2(\boldsymbol{S}))$. The mean activity in pool p ($p = 1, 2$) is defined as

$$y_p = f_p(\boldsymbol{S}) = \frac{\sum_i S_{pi}}{N_p}, \tag{6}$$

which means that $0 \leq y_p \leq 1$. Working with mean activity rather than summed activity makes it easier to compare instances of the model with a different number of neurons. Note that the mapping from $\boldsymbol{S}$ to $\boldsymbol{y} = (y_1, y_2)$ is a many-to-one mapping because several binary vectors $\boldsymbol{S}$ may lead to the same pair of mean activities $\boldsymbol{y} = (y_1, y_2)$.

In a next step, we express the properties of the system in terms of the macroscopic variable $\boldsymbol{y}$ instead of the more detailed microscopic $\boldsymbol{S}$. Expressing the energy from Equation 5 as a function of

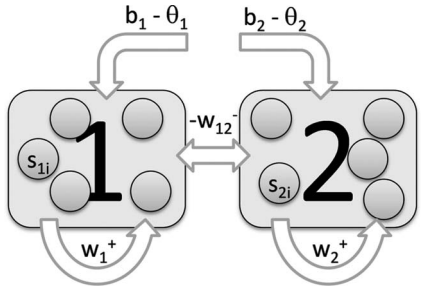


Figure 2. Graphical representation of the Ising Decision Maker network design.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

y_1 and y_2, we get the following (see Appendix A for a detailed derivation of the first equation):

$$\begin{aligned} E(\boldsymbol{S}) &= -w_1^+\left(\frac{(N_1y_1)^2 - N_1y_1}{2}\right) - w_2^+\left(\frac{(N_2y_2)^2 - N_2y_2}{2}\right) \\ &\quad + w_{12}^- N_1N_2y_1y_2 - (b_1 - \theta_1)N_1y_1 - (b_2 - \theta_2)N_2y_2 \qquad (7) \\ &= -\frac{w_1^+N_1^2}{2}y_1^2 - \frac{w_2^+N_2^2}{2}y_2^2 + w_{12}^- N_1N_2y_1y_2 \\ &\quad -\left(N_1b_1 - N_1\left(\theta_1 + \frac{w_1^+}{2}\right)\right)y_1 - \left(N_2b_2 - N_2\left(\theta_1 + \frac{w_2^+}{2}\right)\right)y_2 \\ &= -W_1^+y_1^2 - W_2^+y_2^2 + W_{12}^-y_1y_2 - (B_1 - \Theta_1)y_1 - (B_2 - \Theta_2)y_2 \\ &= E(y_1, y_2) = E(\boldsymbol{y}), \end{aligned} \qquad (8)$$

where $W_p^+ = \frac{w_p^+N_p^2}{2}$, $W_{12}^- = w_{12}^-N_1N_2$, $B_p = N_pb_p$, and $\Theta_p = N_p\left(\theta_p + \frac{W_p^+}{2}\right)$.[3] In this article, we additionally assume both pools have the same number of neurons, meaning $N_p = \frac{N}{2}$; their thresholds are the same, so $\theta_p = \theta$; and their excitatory interactions have the same strength w^+, so $w_p^+ = w^+$. This simplifies Equation 8 to

$$E(\boldsymbol{y}) = -W^+y_1^2 - W^+y_2^2 + W^-y_1y_2 - (B_1 - \Theta)y_1 - (B_2 - \Theta)y_2. \qquad (9)$$

The derivation shows that all configurations $\boldsymbol{S}$ mapping onto the same macroscopic variable $\boldsymbol{y}$ have equal energy: $E(\boldsymbol{S}) = E(\boldsymbol{y})$ for $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{S})$. If we now compute the probability of a macroscopic outcome $\boldsymbol{y}$ based on Equation 2, we find (summing over all microscopic configurations that result in $\boldsymbol{y}$)

$$\begin{aligned} \Pr(\boldsymbol{y}) &= \sum_{\boldsymbol{S}|\boldsymbol{y}=\boldsymbol{f}(\boldsymbol{S})} \Pr(\boldsymbol{S}) = \sum_{\boldsymbol{S}|\boldsymbol{y}=\boldsymbol{f}(\boldsymbol{S})} Z^{-1}e^{-\beta E(\boldsymbol{S})} \\ &= \left[\sum_{\boldsymbol{S}|\boldsymbol{y}=\boldsymbol{f}(\boldsymbol{S})} 1\right] Z^{-1}e^{-\beta E(\boldsymbol{y})} = \Omega(\boldsymbol{y})Z^{-1}e^{-\beta E(\boldsymbol{y})} \\ &= Z^{-1}e^{\Sigma(\boldsymbol{y}) - \beta E(\boldsymbol{y})} = Z^{-1}e^{-\beta(E(\boldsymbol{y}) - \beta^{-1}\Sigma(\boldsymbol{y}))} \\ &= Z^{-1}e^{-\beta F(\boldsymbol{y})}, \end{aligned} \qquad (10)$$

where $\Sigma(\boldsymbol{y})$ is the Boltzmann entropy of the system for $\boldsymbol{y}$:

$$\Sigma(\boldsymbol{y}) = \log(\Omega(\boldsymbol{y})), \qquad (11)$$

and $\Omega(\boldsymbol{y})$ is the number of configurations $\boldsymbol{S}$ that results in $\boldsymbol{y}$.[4] In the final line of the derivation, we use $F(\boldsymbol{y})$ to denote the Helmholtz free energy of the system limited to $\boldsymbol{y}$ (see, e.g., Schroeder, 1999):

$$F(\boldsymbol{y}) = E(\boldsymbol{y}) - \beta^{-1}\Sigma(\boldsymbol{y}). \qquad (12)$$

The total partition function Z can also be written in terms of $F(\boldsymbol{y})$:

$$Z = \sum_{\boldsymbol{y}} e^{-\beta F(\boldsymbol{y})},$$

allowing $\Pr(\boldsymbol{y})$ to be formulated exclusively in terms of $F(\boldsymbol{y})$ and β.

In addition, $\Omega(\boldsymbol{y})$, defined as the number of microscopic states for a given $\boldsymbol{y}$, is easily calculated from the microscopic combinatorics as follows:

$$\begin{aligned} \Omega(\boldsymbol{y}) &= \prod_p \Omega(y_p) \\ &= \prod_p \binom{N_p}{N_py_p} \\ &= \prod_p \frac{N_p!}{(N_py_p)!\,(N_p(1 - y_p))!}. \end{aligned}$$

The entropy of $\boldsymbol{y}$ can then be approximated for large N_p through Stirling's formula ($\log(n!) \approx n\log n - n$) as

$$\begin{aligned} \Sigma(\boldsymbol{y}) &= \log\left[\prod_p \frac{N_p!}{(N_py_p)!\,(N_p(1 - y_p))!}\right] \\ &\approx -\sum_p N_p(y_p\log(y_p) + (1 - y_p)\log(1 - y_p)). \end{aligned} \qquad (13)$$

Again assuming $N_p = \frac{N}{2}$, the final formula for the entropy of the system reads

$$\Sigma(\boldsymbol{y}) \approx -\frac{N}{2}\sum_p (y_p\log(y_p) + (1 - y_p)\log(1 - y_p)).$$

Combining this approximation with Equation 12 gives

$$\begin{aligned} F(\boldsymbol{y}) &= E(\boldsymbol{y}) - \beta^{-1}\Sigma(\boldsymbol{y}) \\ &= -W^+y_1^2 - W^+y_2^2 + W^-y_1y_2 - (B_1 - \Theta)y_1 - (B_2 - \Theta)y_2 \\ &\quad + \beta^{-1}\frac{N}{2}\sum_p (y_p\log(y_p) + (1 - y_p)\log(1 - y_p)). \end{aligned} \qquad (14)$$

Inserting the expression for the free energy from Equation 14 into Equation 10 gives the probability distribution of the pair of mean activities $\boldsymbol{y} = (y_1, y_2)$.

The probability distribution of $\boldsymbol{y}$ now has seven parameters: W^+, W^-, Θ, N, B_1, B_2, and β. However, not all of these parameters are uniquely identifiable because one may multiply β with a constant K and multiply the parameters W^+, W^-, Θ, B_1, and B_2 with K^{-1} without changing the probability mass function $\Pr(\boldsymbol{y})$. Hence, to identify the model, we could set β to an arbitrary value. For reasons of completeness, we leave β in the following equations.

When studying the properties of the probability mass function $\Pr(\boldsymbol{y})$, we can focus our attention on the free energy surface $F(\boldsymbol{y})$ because of the simple monotonic relationship between $\Pr(\boldsymbol{y})$ and $F(\boldsymbol{y})$. It is good to keep in mind that a minimum of $F(\boldsymbol{y})$ corresponds to a maximum in the probability distribution $\Pr(\boldsymbol{y})$.

In Figure 3, we show what the free energy surface may look like at the beginning of an IDM trial and how it deforms under the influence of the external field (B_1, B_2), increasing instantaneously from (0, 0) to some finite two-component value when

[3] The functions $E(\boldsymbol{S})$ and $E(\boldsymbol{y})$ are separate functions and can be distinguished by the notation of their argument.

[4] In physics, the right-hand side of Equation 11 is often preceded by a scaling factor known as the Boltzmann constant k_B. This constant links the microscopic energetic description of a physical system (e.g., an ideal gas) to the Kelvin temperature scale, but this has little meaning in the context of the IDM and is therefore set to 1.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

a stimulus is presented. The parameter values giving rise to these surfaces can be found in Table 1, under the header Equilibrium, and in the figure caption. Note that we use the parameter values from Table 1 as a basis for all the simulations discussed in this article (if other values are used, this is clearly indicated). In Figure 3, it can be seen that under the influence of the external field $\boldsymbol{B}$, the free energy changes from a three-minima surface before stimulus onset (one dominant spontaneous state with low activity for both pools and two less likely decision states with increased activity for one pool and slightly decreased activity for the other) to a two-minima surface during stimulus presentation where there are only two decision states, both of them showing increased activity for one pool and slightly decreased activity for the other. When the stimulus is dropped, the free energy returns to its initial input-free shape.

Although completely defined by the two variables B_1 and B_2, for some applications it is interesting to decompose the external field (B_1, B_2) into a selective (B_{s1}, B_{s2}) and nonselective (B_{ns}, B_{ns}) component:

$$B_1 = B_{s1} + B_{ns},$$
$$B_2 = B_{s2} + B_{ns}, \tag{15}$$

with B_{s1}, B_{s2}, $B_{ns} \geq 0$. Evidently, the nonselective component impacts both pools identically, whereas the selective component may have a different value per pool. Additionally, the selective component (B_{s1}, B_{s2}) can be reformulated, introducing *stimulus distinctness* C as a signed proportion (i.e., $C \in [-1, 1]$) and input strength B_s as the overall strength of the selective part of the input:

Table 1
Parameter Values of the Ising Decision Maker Used for Simulations (Unless Defined Otherwise)

Parameter	Description	Value
Equilibrium		
W^+	Self-excitation	52,500
W^-	Mutual inhibition	8,400
Θ	Internal threshold	51,450
N	Total number of neurons	2,000
$\boldsymbol{B}$	Two-component external field	Different values
β	Inverse statistical temperature	1/24
Dynamical		
σ	Collective step size	0.01
Δt	Time step	1 ms
h	Detection box size	0.4
Nondecision		
T_{er}	Nondecision time	0 ms

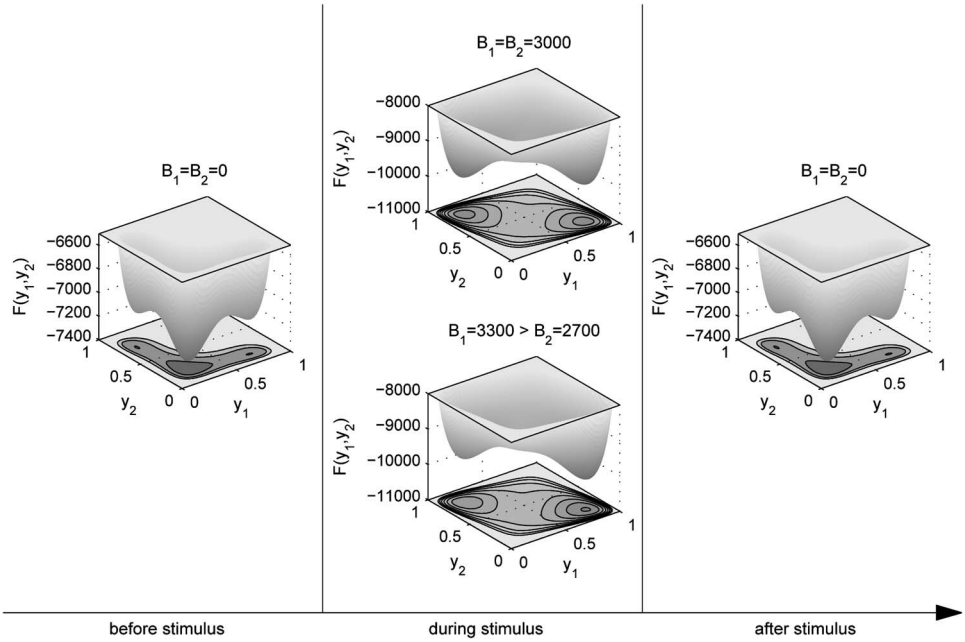


Figure 3. An example of what the free energy function $F(\boldsymbol{y})$ may look like during the time course of a single Ising Decision Maker trial (extremely high free energy function values near the boundaries have been clipped off). The left graph shows the free energy surface at the beginning of the trial (when $B_1 = B_2 = 0$): There is one dominant spontaneous state with low activity for both pools and two less likely decision states for which one of the pools is moderately active but not the other. When the external field is turned on (i.e., the stimulus is presented), the spontaneous state disappears, and only two decision minima are present, showing increased activity for one pool and slightly decreased activity for the other. Two examples are shown in the middle figures: The upper figure is caused by a stimulus with equal external field components $B_1 = B_2 = 3{,}000$, the lower figure by a stimulus with one dominant field component $B_1 = 3{,}300 > B_2 = 2{,}700$. When the external field is switched off again (i.e., the stimulus disappears), the initial free energy surface returns. The remaining parameters used here are the equilibrium parameters shown in Table 1.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

$$B_1 = B_s(1 + C) + B_{ns},$$
$$B_2 = B_s(1 - C) + B_{ns}. \quad (16)$$

A stimulus distinctness of zero means the stimulus is completely uninformative and both alternatives should be considered equally correct. In that case, there is an equal amount of evidence for both alternatives, or $B_1 = B_2$. A strictly positive value for stimulus distinctness C means $B_1 > B_2$ and assigns more evidence to the first alternative then the second, making the first alternative the objectively correct choice (unless the assignment of stimuli to categories is stochastic, as in Ratcliff & Rouder, 1998, in which case $B_1 > B_2$ can occur, while Alternative 2 is the correct one). The case of negative distinctness C, where $B_2 > B_1$ and the second alternative is the correct choice, can be reinterpreted as a case of positive distinctness by simply switching pool labels. Indeed, because both pools are identical in design (same number of neurons $\frac{N}{2}$, same W^+, same Θ), swapping B_1 and B_2 is identical to swapping y_1 and y_2. Stimulus distinctness C is not allowed to be greater than one because this would lead to a negative value for B_{s2}. Note that Figure 3 also shows that stimulus distinctness is linked with an asymmetry in the free energy surface across the main diagonal of the (y_1, y_2)-plane (lower middle plot). Nonzero stimulus distinctness (meaning that there is more evidence for one of the two response options) leads to a lower free energy and consequently higher probability of one of the two decision states. In the case of zero stimulus distinctness (upper middle plot), both decision states are equally likely as the corresponding free energy minima have the same shape and depth. In the section entitled Modes of Operation of the IDM, below, we study in greater detail the different free energy surfaces that can be generated using various parameter values.

The proposed parameterization of the external input field is a bit more sophisticated than is common for the modeling of a stimulus (e.g., a simple drift rate in the Ratcliff diffusion model or a dual drift rate in most two-dimensional diffusion models), but it provides an operational definition of the stimuli presented in a variety of experiments. We start by recognizing the fact that a stimulus has different aspects, all of them possibly having an impact on the two available decision pools. Some aspects of the stimulus, if in a suitable range, are not relevant to the discrimination task at hand (e.g., the brightness of a dot when asked to determine if it appears left or right from the middle of the screen is not relevant unless it is subthreshold). The sum of these decision-irrelevant or nonselective aspects can still result in an external field on both pools as long as it is equal for all the stimuli being discriminated. The aggregate impact of the nonselective aspects is modeled as the nonselective stimulus strength B_{ns} and is considered to be constant in a standard two-choice RT experiment. The stimulus aspects that are related to the discrimination at hand (e.g., the location of the dot when asked to determine if it appears left or right from the middle of the screen) are allowed to have a different impact on both pools, which is represented by the decision-relevant or selective evidence for both alternatives, denoted as B_{s1} and B_{s2}.

Regarding the input parameters, two additional things should be noted. First, it is not unusual to assume that the total amount of evidence supplied to the network is constant for a standard two-choice RT experiment (Ratcliff et al., 2007). This means that the selective evidence parameters B_{s1} and B_{s2} can be reformulated as B_s and C, with B_s constant for all available stimuli. The only parameter that still depends on the stimulus is distinctness C. Second, for the simple two-choice RT paradigm, B_{ns} and B_s are not uniquely defined; they can only be measured independently in a specific experimental design. In the sections on the van der Molen-Keuss effect and Weber's law (see below), we demonstrate the use of the complete input parameterization.

Dynamics of the IDM. Having explained the equilibrium properties of the IDM, we now turn to the dynamical aspects of the model. Although the underlying stochastic Hopfield network is often equipped with a single neuron updating Metropolis dynamics, the equilibrium description given above is valid for many other Markov chain dynamics. In this article, we first discuss the traditional single neuron updating case and then focus on a Markov chain defined on the macroscopic level, considering changes of $\boldsymbol{y}$ rather than $\boldsymbol{S}$. How this macroscopically defined dynamics can be interpreted as a form of collective change of microscopic activity is explained. Throughout the rest of the article, only the macroscopically defined dynamics is used.

First, in order to attach a physical time to the discrete, dimensionless Markov chain time, we quantify one Markov chain time step as Δt. We then define time t as an integer multiple n of Δt:

$$t = n\Delta t, \quad (17)$$

with n the number of Markov chain time steps completed. The parameter Δt determines the time scale of the dynamics and is one of the parameters that need to be estimated when fitting the model to data (see below).

Microscopically defined dynamics based on single neuron updates. In this paragraph, we consider the traditional Metropolis dynamics based on single neuron updates as described in the section on the stochastic Hopfield network. Because, for a large N, the effect of changing a single neuron's activity on the mean activity is small, we can locally approximate energy, entropy, and free energy as linear functions of mean activity and write the resulting dynamics as a set of diffusion equations at the macroscopic level (Kikuchi, Yoshida, Maekawa, & Watanabe, 1991). As calculated in Appendix B, the diffusion equations are

$$dy_1 = -\beta D\frac{\partial F(\boldsymbol{y})}{\partial y_1}dt + \sqrt{2D}dW_1,$$
$$dy_2 = -\beta D\frac{\partial F(\boldsymbol{y})}{\partial y_2}dt + \sqrt{2D}dW_2, \quad (18)$$

where W_1 and W_2 are two uncorrelated Wiener processes and

$$D = \frac{2}{eN^2\Delta t}.$$

Macroscopically defined dynamics based on collective neuron updates. An interesting question is what happens when instead of sequentially updating individual neurons, we allow larger sets of neurons to undergo a collective change in activity during one and the same time step. There are a number of ways this could be achieved on a microscopic level, but deriving the exact macroscopic effect is a challenging affair. However, it is easy to see that collective neuron updating can result in macroscopic changes (i.e., in y_1 and y_2) that exceed the impact of a single neuron update.

An easy and straightforward way to incorporate collective neuron updating into the IDM is by allowing larger steps of the mean

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

activities. More specifically, we assume that at every time step, a proposed change in (y_1, y_2) is sampled from a bivariate normal probability distribution $\mathcal{N}(\mathbf{0}, \sigma^2 I_2)$ (with I_2 the two-dimensional identity matrix). This proposal is then accepted or rejected based on the Metropolis acceptance probability that is connected to the free energy difference between the proposed and the current macroscopic configuration.[5] The parameter σ regulates the size of the proposed activity steps and is referred to as the *collective step size*. As a result, when σ is large, the IDM uses large activity steps and scans the free energy surface in a rough, coarse manner. We refer to such large step dynamics as *coarse-grained dynamics*. When σ is small, the IDM uses small activity steps, and the resulting dynamics is more sensitive to the local shape of the free energy surface. Regardless of the value of σ, however, in the long run, the Boltzmann distribution will be obtained.

The dynamics described above is defined on a macroscopic level and may seem disconnected from the microscopically defined sequential dynamics mentioned earlier. However, for vanishing collective step size σ and time step Δt, it connects back to the previous diffusion result (see Appendix C):

$$\begin{aligned} dy_1 &= -\beta D \frac{\partial F(\mathbf{y})}{\partial y_1} dt + \sqrt{2D} dW_1, \\ dy_2 &= -\beta D \frac{\partial F(\mathbf{y})}{\partial y_2} dt + \sqrt{2D} dW_2, \end{aligned} \tag{19}$$

where W_1 and W_2 are two uncorrelated Wiener processes and

$$\frac{\sigma^2}{2\Delta t} \rightarrow D. \tag{20}$$

Apart from the expression of the diffusion constant, this result is exactly the same as the one obtained with microscopically defined sequential dynamics.

For nonvanishing collective step size σ, however, it is no longer possible to approximate the dynamics as a diffusion equation. This deviation from local gradient behavior, brought about by collective neuron updating, has interesting properties for the modeling of instruction-related speed–accuracy tradeoff (SAT; see below). Throughout the rest of this article, we use the macroscopically defined dynamics with collective step size σ. The values of the parameters σ and Δt that are used in the simulations can be found in Table 1 (if other values are used, they are clearly indicated). Taking sufficiently small values for σ and Δt will bring the system to the diffusion limit, which it has in common with the microscopically defined dynamics. In this limiting case, the parameters σ and Δt can be replaced by a single diffusion constant D, as is shown in Equation 20.

Starting position. In contrast to traditional accumulator and diffusion models, the starting position of an IDM decision trajectory is not defined as a separate parameter. It is embedded in the free energy function itself, where the absence of an external field before stimulus presentation forces a low activity for both pools (see Figure 3, left plot). This gives rise to a stable spontaneous distribution, which serves as a random starting position for the activity at the moment the external field (B_1, B_2) (induced by the stimulus) is introduced. The introduction of the external field typically leads to two decision states (two minima) in the free energy surface (see Figure 3, middle plots) that force the decision trajectory away from its starting point and toward one of the decision states. Thus, from the very moment the external field is introduced and the spontaneous state is dissolved, the decision clock starts running ($t = 0$).

Stopping rule. The decision clock is stopped when the decision trajectory arrives at one of the decision states. This arrival is detected by means of a specific stopping rule. In this case, square detection boxes placed around the two decision minima are used (see Figure 4). After stimulus presentation, the decision part of the trial is completed as soon as the system enters a box and the choice pertaining to the associated decision minimum is made.

All existing dynamical models of decision making struggle with the stopping issue in one form or another, and several algorithms have already been proposed. By far the most popular choice for two-dimensional dynamical systems is a set of (perpendicularly intersecting) absorbing boundary lines. However, as is shown in the section on the van der Molen-Keuss effect (see below), for very high stimulus intensities, the IDM displays potentially relevant exotic behavior that can only be unlocked if the decision process is allowed to advance to high activities for both pools simultaneously so it can approach the decision state from another angle. This is not possible using a traditional absorbing boundary line, but it is possible using a detection box. For regular stimuli, however (see the simulations and the fit to a real data set below), neural activities tend to cross the absorbing boundary at low-activity rates for the competing pool, and there is almost no difference between using a line or a box. In other words, the detection box behaves like an absorbing boundary line for normal stimuli, while still allowing interesting behavior for high stimulus intensities. As indicated in Table 1, the detection box size h used in simulations is 0.4.

Nondecision time. For even the simplest of decision tasks, the two-component external field (B_1, B_2) is the result of predecision information processing that in itself takes time to complete. Furthermore, after a decision is made, a physical reaction is required to complete the speeded two-choice RT trial, which again takes time. As is common in the diffusion modeling literature (e.g., Luce, 1986; Ratcliff & Rouder, 1998), we model all this extra nondecision RT as a simple constant T_{er} that needs to be added to the IDM-generated decision time to obtain the total RT. For most of the illustrative simulations in this article, we take T_{er} to be zero. This way only the actual decision time is obtained. For the parameter recovery study in the application section, however, data are simulated with a typical nondecision time of 0.3 s.

Recapitulation: A typical decision trial. The IDM provides a dynamical description of the decision-making system for the entire duration of the choice trial: before, during, and (if desired) even after the exposure to the stimulus. As illustrated in Figure 3, the free energy surface changes dramatically between these stages. Initially, when input is absent ($B_1 = B_2 = 0$), there is a locally stable spontaneous state where the system is forced to wait in anticipation of a stimulus. When a detectable stimulus is presented to the IDM, the free energy surface is instantly deformed to a surface with two clear decision states. In an attempt to reestablish equilibrium, the system now evolves to one of the two locally

[5] To compensate for the fact that not every macroscopic state of activity represents the same number of microscopic states, free energy (and not just energy) has to be used in the Metropolis algorithm. This way, detailed balance is preserved, and the same IDM equilibrium distribution is obtained as outlined before.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

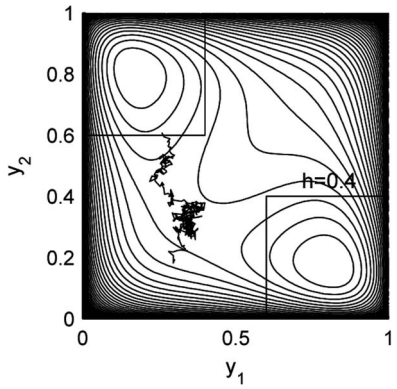


Figure 4. Illustration of the stopping rule: The square detection boxes are superimposed on the contour lines of the free energy surface. The length of a side h or box size is the only parameter, and in our simulations, it is fixed to 0.4. A choice is made as soon as the evolving activities hit one of the boxes.

stable decision states. The evolution of the mean activities of the two pools of neurons is described by a discrete-time Markov chain (i.e., using a Metropolis algorithm). The proposed activity steps are sampled from a bivariate normal distribution. Upon choice detection (when crossing the boundaries of one of the square detection boxes), the appropriate motor response is triggered. The effective decision-making process is timed between the introduction of the external field to the network and the detected arrival at a decision state. To account for the entire observed interval between stimulus presentation and final response, an amount of nondecision time can be added. In the simplest experimental paradigm, once present, the stimulus does not change until a response is given, upon which the trial is considered completed. Unless specified otherwise, we assume this to be the case.

Choice probabilities and RT distributions. For a given stimulus (with an associated external field), the IDM generates a pair of RT distributions ($\phi_1(t)$, $\phi_2(t)$), for Choices 1 (correct) and 2 (incorrect), respectively. More specifically, $\phi_p(t)$ represents the fraction of trials that arrive at the detection box corresponding with decision p at time t, and its support ranges from T_{er} (the minimum choice RT) to $+\infty$. Strictly speaking, the support of $\phi_p(t)$ is discrete (because of the use of discrete-time Markov chain dynamics; see Equation 17). However, because the time step Δt is typically several orders of magnitude smaller than the width of the actual RT distribution, the underlying discreteness has no practical consequence for the smoothness of the RT distribution. Therefore, we interpret $\phi_p(t)$ as a step-function density with a continuous support, so we can use standard calculus and terminology when handling it. The probability of decision p occurring regardless of RT is $\pi_p = \int \phi_p(t)dt$. The accompanying mean RTs are denoted as $\tau_p = \frac{1}{\pi_p}\int t\phi_p(t)dt$. The cumulative distribution is denoted as $\Phi_p(t) = \int_0^t \phi_p(u)du$.

Comparison of the IDM to Other Models

Several models for speeded two-choice RT have been formulated as a single SDE or a set of SDEs (for a comprehensive overview, see Bogacz et al., 2006). In this section, we compare the IDM with the most important existing models. As most models in the comparison are based on continuous diffusion equations, it makes sense to compare them to the continuous diffusion limit of the IDM.

Let us take as a starting point the following general multivariate SDE:

$$d\boldsymbol{y} = -\nabla U(\boldsymbol{y})dt + c d\boldsymbol{W}, \tag{21}$$

where $\boldsymbol{y}$ is a stochastic variable changing over time (the dependence on time t is suppressed in our equations), $\boldsymbol{W}$ is a vector of uncorrelated Wiener processes, $U(\boldsymbol{y})$ is the potential function, and c is the diffusion constant (in the SDE sense[6]). The dynamical system described in Equation 21 minimizes its potential function $U(\boldsymbol{y})$ stochastically by drifting along the negative gradient of $U(\boldsymbol{y})$. The fundamental idea of this section is to compare different accumulator and diffusion models with respect to the potential functions the decision maker tries to minimize.

Let us start with relating the IDM to Equation 21. If we set $c_{\text{IDM}} = \sqrt{2D}$ and

$$U_{\text{IDM}}(\boldsymbol{y}) = \beta D F(\boldsymbol{y}),$$

we get Equation 19. This shows that for the diffusion limit of the IDM, the potential function that is minimized is the (scaled) free energy. We may insert the expression of $F(\boldsymbol{y})$ in Equation 19 to obtain a more detailed expression for the IDM drift rate vector (or gradient of the potential function):

$$\begin{aligned} -\beta D\frac{\partial F(\boldsymbol{y})}{\partial y_1} &= -\beta D[-2W^+ y_1 + W^- y_2 - (B_1 - \Theta)] \\ &\quad -D\frac{N}{2}(\log(y_1) - \log(1 - y_1)), \\ -\beta D\frac{\partial F(\boldsymbol{y})}{\partial y_2} &= -\beta D[-2W^+ y_2 + W^- y_1 - (B_2 - \Theta)] \\ &\quad -D\frac{N}{2}(\log(y_2) - \log(1 - y_2)). \end{aligned} \tag{22}$$

The drift rates are coupled (i.e., y_1 and y_2 appear in both drift rates), and they are nonlinear functions of y_1 and y_2.

The Ratcliff Diffusion Model

A very popular model is the Ratcliff diffusion model (Ratcliff, 1978; Ratcliff & McKoon, 2007; Ratcliff & Rouder, 1998; Wagenmakers, 2009). The key part of the model is the following one-dimensional SDE:

$$dy = \nu dt + c dW. \tag{23}$$

Each stimulus is characterized by a specific drift rate ν that quantifies its net evidence toward one of the alternatives. For each trial of the experiment, the decision process starts at $y = z$ (with $0 < z < a$) and continues until one of two boundaries 0 or a is reached. A decision is made depending on which boundary is crossed. The potential function characterizing Equation 23 is linear and only specified up to an arbitrary constant k:

[6] All other instances of the term *diffusion constant* in the text are used in the Focker–Planck sense and are symbolized by D.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

$$U_{\text{RDM}}(y) = -\nu y + k. \tag{24}$$

A graphical representation of this potential function is shown in Figure 5a. The actual state of the system is defined by its position on the y-axis. The system tends to move in a downhill direction, that is, in a direction opposite to the gradient of the potential function U. This evolution is made stochastic by the Wiener noise. Once the system crosses one of the decision boundaries, the corresponding decision is made, and the diffusion process is stopped. Often, the model parameters are defined as realizations of random variables, which follow some distribution. Such a set-up introduces intertrial variability in the formulation of Equation 23 (see Ratcliff & Rouder, 1998, for more explanation of these variabilities).

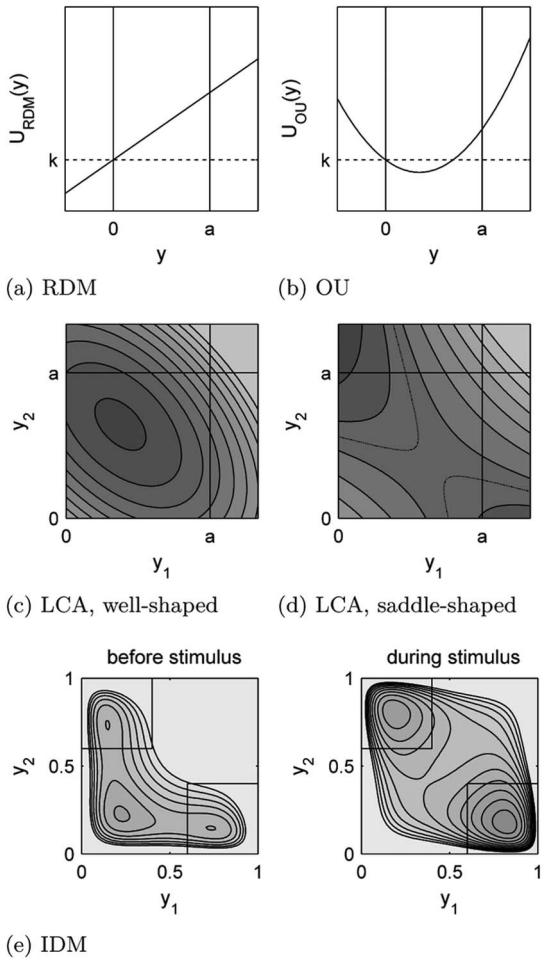


Figure 5. Potential functions of different diffusion models for speeded two-choice response time. a: The Ratcliff diffusion model (RDM; Equation 24 with $\nu = 0.66$, $a = 0.5$). b: The Ornstein-Uhlenbeck model (OU; Equation 26 with $\nu = 0.7$, $\gamma = 4$, $a = 0.5$). c: The leaky competing accumulator model (LCA), well-shaped (Equation 28 with $\nu_1 = 2.1$, $\nu_2 = 2.5$, $\gamma = 2$, $\kappa = 1$, $a = 1.5$). d: The LCA, saddle-shaped (Equation 28 with $\nu_1 = 2.1$, $\nu_2 = 2.5$, $\gamma = 2$, $\kappa = 4$, $a = 0.75$). e: The Ising Decision Maker (IDM; Equation 29 with the parameters from Table 1 and external field changing from $(B_1, B_2) = (0, 0)$ [left plot] to $(B_1, B_2) = (3{,}300, 2{,}700)$ [right plot]).

The Ornstein-Uhlenbeck Diffusion Model

Another model that received some attention in the mathematical psychology literature is based on the Ornstein-Uhlenbeck (OU) stochastic process. The OU diffusion model is sometimes referred to as the leaky accumulator model but was first applied to decision making by Busemeyer and Townsend under the name of decision field theory (Busemeyer & Townsend, 1992, 1993). The SDE reads as

$$dy = (\nu - \gamma y)dt + cdW, \tag{25}$$

with $\gamma > 0$ (such that we have a mean-reverting process). The potential function for the OU model is defined as

$$U_{\text{OU}}(y) = \frac{1}{2}\gamma y^2 - \nu y + k. \tag{26}$$

A graphical representation of this potential function is shown in Figure 5b. Boundaries are assumed at 0 and a, and the starting point of the process is $0 < z < a$. The process strives to reach the minimum of a parabolic potential function at $y_{\min} = \frac{\nu}{\gamma}$. The Wiener noise keeps the system from freezing in that minimum, making sure it eventually crosses one of the decision boundaries in case $0 < y_{min} < a$.

The Leaky Competing Accumulator Model

A more advanced version of the OU model is the leaky competing accumulator model or LCA (see Bogacz et al., 2006; Usher & McClelland, 2001). The LCA or mutual inhibition model is essentially a two-dimensional OU process with mutual inhibition. Two accumulators y_1 and y_2, starting out at the origin, evolve according to the following set of SDEs:

$$\begin{aligned} dy_1 &= (\nu_1 - \gamma y_1 - \kappa y_2)dt + cdW_1, \\ dy_2 &= (\nu_2 - \gamma y_2 - \kappa y_1)dt + cdW_2, \end{aligned} \tag{27}$$

with κ, ν_1, $\nu_2 > 0$. At a certain point in time, one of the two processes exceeds a bound a, and the corresponding decision is made. For the LCA model, the potential function depends on y_1 and y_2:

$$U_{\text{LCA}}(y_1, y_2) = \left(\frac{1}{2}\gamma(y_1^2 + y_2^2) - \nu_1 y_1 - \nu_2 y_2 + \kappa y_1 y_2\right) + k. \tag{28}$$

Note that there are only boundaries placed at $y_1 = a$ and $y_2 = a$ but no lower boundaries, meaning the accumulators y_1 and y_2 can become negative without consequence. Usher and McClelland (2001) used a threshold-linear activation function $y_p \rightarrow \max(y_p, 0)$, $p = 1, 2$, to ensure this never happens (for another account, see van Ravenzwaaij, van der Maas, & Wagenmakers, 2012). If $\kappa^2 < \gamma^2$, the potential function is in fact a simple two-dimensional well ($\gamma > 0$) or hill ($\gamma < 0$) with its minimum at $\boldsymbol{y}_{\min} = \left(\frac{\gamma\nu_1 - \kappa\nu_2}{\gamma^2 - \kappa^2}, \frac{\gamma\nu_2 - \kappa\nu_1}{\gamma^2 - \kappa^2}\right)$. The corresponding contour lines are elliptic (see Figure 5c). If $\kappa^2 > \gamma^2$, the potential function has the shape of a saddle, with its saddle-point at the same coordinates $\boldsymbol{y}_{\min}$. The contour lines are now hyperbolic. A saddle-shaped LCA potential function can be seen in Figure 5d.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

To examine the relation between the leaky accumulator model and the diffusion limit of the IDM, consider again the potential function of the IDM:

$$U_{\mathrm{IDM}}(y_1, y_2) = -\beta D W^{+}(y_1^2 + y_2^2) - \beta D(B_1 - \Theta)y_1 - \beta D(B_2 - \Theta)y_2 + \beta D W^{-} y_1 y_2 + D\frac{N}{2}\sum_p (y_p \log(y_p) + (1 - y_p)\log(1 - y_p)). \quad (29)$$

Ignoring the IDM-specific contribution of the entropy (last line in Equation 29), the two potential functions are both quadratic. The linear terms have similar roles in both models: They represent the sensory input to the decision process (expressed through the term containing the individual y_1 or y_2) and the mutual inhibition of both decision variables during evidence buildup (as expressed in the term with the cross-product y_1y_2). In the IDM interpretation, the internal threshold Θ provides a permanent contribution to the linear term, independent of stimulus (presentation). Upon stimulus presentation, this baseline will be perturbed by the stimulus-dependent external field $\boldsymbol{B}$. In the LCA, the quadratic term is allowed to be positive (leaky) as well as negative (self-amplifying; see, e.g., Usher & McClelland, 2001). In the IDM, however, the explicit quadratic term depends on the positive self-excitation W^{+} and is therefore negative by definition. Note, however, that the logarithmic function makes an additional, positive contribution to the quadratic term, as becomes clear when performing a Taylor expansion of the free energy around (0.5, 0.5) (this can be seen in Appendix D, where this expansion is carried out for another purpose). The resulting total quadratic coefficient is required to be negative to allow for multiple stable states. Despite slight differences in the interpretation of the corresponding parameters, the LCA can be seen as a second-order approximation of the IDM's dynamics during stimulus presentation.

The nonlinear part of the IDM free energy gives rise to several unique properties. In contrast to the LCA potential surface, the full nonlinear IDM potential surface can have multiple simultaneously occurring locally differentiable minima, corresponding to stable states of the underlying neural network. As already mentioned in the section Equilibrium description, one of these states is identified as the spontaneous state, a low-activity stable state that exists when no stimulus is applied. This spontaneous state gives rise to the concept of starting point in a natural way. Indeed, when a stimulus is presented, the spontaneous state disappears, forcing the system to evolve toward one of the two newly formed decision minima. Moreover, the equilibrium probability distribution introduces a variability of the starting point. In diffusion modeling practice, starting point variability is imposed to account for phenomena such as fast errors (see below). In the IDM, both the starting point and the variability around it are inherent in the model.

Other Models: The Pooled Inhibition Model and Wong and Wang's Integrate-and-Fire Network

Another model for speeded decision making is the pooled inhibition model. This model is a three-dimensional dynamical system, consisting of two decision variables, with similar roles as in the LCA model, and one extra variable representing an inhibition pool:

$$dy_1 = (\nu_1 - \gamma y_1 - \kappa y_3)dt + cdW_1, \quad (30)$$

$$dy_2 = (\nu_2 - \gamma y_2 - \kappa y_3)dt + cdW_2, \quad (31)$$

$$dy_3 = (\beta y_1 + \beta y_2 - \alpha y_3)dt, \quad (32)$$

with $\kappa, \alpha, \beta, \nu_1, \nu_2 > 0$. If the pools are leaking, then $\gamma > 0$; if they are self-amplifying, then $\gamma < 0$. Although commonly reduced to the LCA (see Bogacz et al., 2006), the pooled inhibition model has an interesting feature that deserves some attention. Pooled inhibition makes use of asymmetrical interactions: The inhibition pool y_3 is excited by the decision pools y_1 and y_2 (Equation 32 with terms βy_1 and βy_1), while y_3 itself inhibits y_1 and y_2 (Equations 30 and 31 with terms $-\kappa y_3$ and $-\kappa y_3$, respectively). This asymmetry induces oscillatory effects that prevent writing the drift vector uniquely as the gradient of a potential function. For a complete potential description, an extra vector potential has to be added that describes the rotational part of the vector field (see the Helmholtz decomposition in Griffiths, 1999).

In their reduction to the LCA, Bogacz et al. (2006) made the assumption that inhibition pool variable y_3 evolves much faster than decision pool variables y_1 and y_2, effectively flattening out the oscillations. This means that for every (y_1, y_2), the inhibition pool variable y_3 immediately reaches its stable point (it is stable because $\alpha > 0$) or $y_{3_{min}} = \frac{\beta}{\alpha}(y_1 + y_2)$. This reduces the system to

$$dy_1 = \left(\nu_1 - \left(\gamma + \frac{\kappa\beta}{\alpha}\right)y_1 - \frac{\kappa\beta}{\alpha}y_2\right)dt + cdW_1,$$

$$dy_2 = \left(\nu_2 - \left(\gamma + \frac{\kappa\beta}{\alpha}\right)y_2 - \frac{\kappa\beta}{\alpha}y_1\right)dt + cdW_2.$$

If $\gamma < \frac{k\beta}{\alpha}$, the LCA is self-amplifying. A saddle-shaped potential surface is only possible if $\left(\gamma + \frac{\kappa\beta}{\alpha}\right)^2 < \left(\frac{\kappa\beta}{\alpha}\right)^2$ or $\gamma < 0$. In other words, a saddle-shaped LCA derived from the pooled inhibition model will always be self-amplifying.

Like the pooled inhibition model, the integrate-and-fire network put forward by Wong and Wang (2006) is based on biologically more realistic asymmetrical interactions and therefore includes a separate pool of inhibitory neurons (for a detailed mathematical description of Wong and Wang's integrate-and-fire network, we refer the reader to the original paper). Also, while the IDM models the influence of third-party neurons as thermal noise expressed by the inverse statistical temperature β, Wong and Wang's network explicitly uses an extra pool of excitatory neurons. Nevertheless, Wong and Wong eventually approximated their integrate-and-fire network with a system consisting only of two pools of selective neurons with exclusively symmetric interactions (which no longer supports oscillatory behavior). This means the Wong and Wang integrate-and-fire network can be analyzed with the help of potential surfaces. Although the derivations for the integrate-and-fire network are analytically more intricate, there is a strong qualitative resemblance between its final two-dimensional representation and the IDM's free energy surfaces. This observed similarity should not come as a surprise because the resulting neural correlations in both systems (positive within pools and negative between pools) are very similar by set-up.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Modes of Operation of the IDM

For the free energy surfaces shown in Figure 3, we used parameter values giving rise to realistic behavior of the model (see Table 1). Other parameter values, however, could result in totally different behavior. The most important factor determining the model's behavior is the number of local minima present at the different stages of input. To examine this qualitative feature of the free energy surface, we systematically vary two crucial parameters and count the number of local minima of $F(\boldsymbol{y})$ for each set. The first parameter of focus is of course the external field strength $\boldsymbol{B} = (B_1, B_2)$, which quantifies the different stages of input during a trial. More specifically, upon stimulus presentation, it changes from zero to some stimulus-specific two-component value. For this analysis, the two components of the external field are taken to be equal, $B_1 = B_2 = B$, leading to a diagonal symmetry for all resulting free energy surfaces. Setting B_1 equal to B_2, we limit ourselves to stimuli carrying equal evidence for both choices. As can be seen from Equation 16, we have chosen $C = 0$ and $B = B_s + B_{ns}$. Considering stimuli with nonzero distinctness would complicate matters more than necessary for this particular analysis. In line with Wong and Wang (2006), the other parameter of focus is the self-excitation W^+. All other parameters were fixed to particular values (see Table 1, equilibrium parameters). Figure 6 consists of two panels. Figure 6a contains a partitioning of the (B, W^+)-coordinate plane in terms of the number of local minima of the corresponding free energy function.[7] The parameter B is varied from $-2{\cdot}10^4$ to $2{\cdot}10^4$ and W^+ is varied from $4{\cdot}10^4$ to $6.5{\cdot}10^4$. Figure 6b shows plots of the free energy surfaces corresponding to various positions on the (B, W^+)-coordinate plane (the positions are indicated with lowercase letters).

Looking at Figure 6a, we see regions with one up to four local minima. Not surprisingly, this partitioning is somewhat similar to the one obtained by Wong and Wang (2006) for their integrate-and-fire model. The operation of the IDM is best understood in terms of horizontal movements over the (B, W^+)-coordinate plane, as the presentation of a stimulus during a trial manifests itself through a change in parameter B. When a stimulus is presented, parameter B changes from zero to some finite value. Instantly, the free energy $F(\boldsymbol{y})$ changes its shape, depending on the values of the other parameters of the model. When the stimulus disappears, B drops to zero again, and $F(\boldsymbol{y})$ returns to its original input-free form. Note that the position of zero input is arbitrary in the sense that the horizontal axis can be shifted by changing the internal activation threshold Θ.

To structure our discussion of Figure 6, we have divided the (B, W^+)-coordinate plane into horizontal bands, denoted by I to IV, in which different modes of operation can be recognized. Band I (lowest values of W^+) is a region of pure monostability: Regardless of the external field strength B, only one minimum is found. At this single minimum, both pools are equally active. This can be seen in the five subplots on the lowest row of Figure 6b, showing the free energy surfaces for five parameter combinations in Band I. For increasing B, the single minimum travels from low to high activity, following the main diagonal line (which is the axis of symmetry).

Apart from a large region of monostability, Band II contains a small region of free energy surfaces with two minima, called the *competition region.* By competition, we mean that there is a concurrent attraction of two opposing minima, each corresponding to a high activity for one pool and a low activity for the other. These minima are called *decision states.* For increasing B, the system changes from monostability to competition and back to monostability. This evolution is shown on the second row of Figure 6b. The low-activity symmetrical state in the first region of monostability is called the *spontaneous state.* It corresponds to the starting area of the decision process (at $B = 0$). The high-activity variant in the second region of monostability is called the *startled state*. A possible implication of this startled state is studied in detail below (in the section on the van der Molen-Keuss effect).

Band III is comparable to Band II but has some additional three-state stability. For increasing B, the system passes following stages: monostability, three-state stability (a spontaneous state and two decision states), competition, three-state stability (two decision states and a startled state), and monostability. This is illustrated by the third row of Figure 6b.

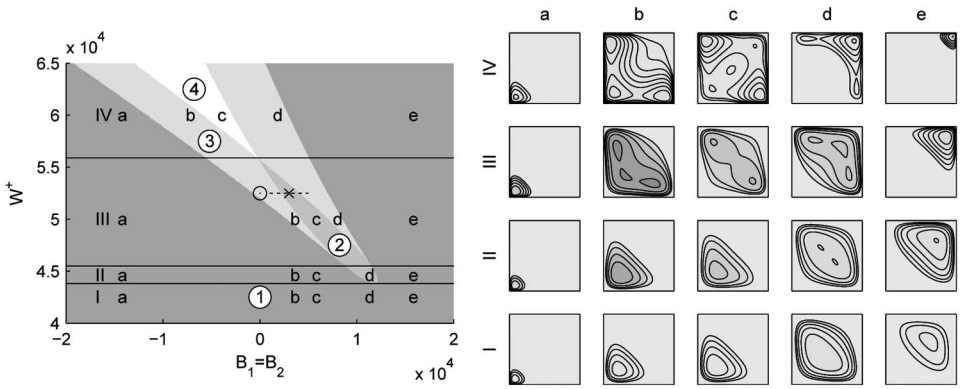
Band IV consists of free energy surfaces with two, three, and four minima (two symmetrical and two opposing minima). For increasing B, the system consequently shows monostability, three-state stability (a spontaneous state and two decision states), four-state stability, three-state stability (two decision states and a startled state), and monostability. We refer the reader to the top row of Figure 6b for an illustration.

Next, we consider which bands allow for a realistic decision-making model. The input-free offset Θ has to be chosen in such a way that at $B_1 = B_2 = 0$, a process can reside in a low-activity, nonselective state (i.e., a minimum on the main diagonal close to $\boldsymbol{y} = (0, 0)$). This setting should account for a stable baseline activity. If B (or B_1 and B_2) is increased (because of the stimulus onset), the free energy function $F(\boldsymbol{y})$ should show two decision states. This is possible in all bands but I.

In Band II, decision making can only happen in the competition region. When a stimulus is presented, the system changes from a monostable configuration with a spontaneous state to a configuration with two minima, both of them decision states. Consequently, the process is forced to evolve to one of the two decision states, thereby crossing a detection boundary and making a decision. When the input disappears again, the system is thrown back to its original monostable configuration, guiding the process back to the spontaneous state and thus forgetting the decision it made.

In Band III, there are two possible regimes of decision making. If a stimulus carries the system into the first region of three-state stability, sufficiently close to the neighboring competition region, low-lying decision states have already formed while the spontaneous state (i.e., the low-activity minimum on the main diagonal) is still present but very shallow. In this case, thermal fluctuations (i.e., random noise inherent in the IDM) may cause the process to cross the lowered free energy barrier between the spontaneous state and one of the decision states (see also Deco, Rolls, & Romo, 2009, for a similar account). For larger input strengths, the spontaneous state disappears completely, and decision making again takes place inside the competition region. The second,

[7] These minima are found by running a local minimization algorithm initiated from a sufficiently fine grid of starting points on the free energy surface.



(a) Number of local minima (b) Corresponding free energy surfaces

Figure 6. Number of local minima of the free energy surface for varying self-excitation W^+ and external field $B_1 = B_2$. The remaining model parameters are fixed as indicated in Table 1. In Panel a, presented from dark to light are regions with one to four minima. The regions are labeled with their respective number of minima (circled numbers). The same graph is divided into four bands (I–IV), representing four possible modes of operation. Panel b shows the actual free energy surface at five subsequent positions for each of the four bands. The positions are indicated on Panel a and marked with the letters a, b, c, d, and e. In Panel a, the input-free situation ($\boldsymbol{B} = (0, 0)$) and a possible input situation ($\boldsymbol{B} = (3{,}000, 3{,}000)$) for $W^+ = 52{,}500$ are marked with a circle and a cross. These correspond to the free energy surfaces shown in Figure 3 (left and upper middle ones, respectively). Upon stimulus presentation, the free energy surface is thrown from the input-free situation (circle marker) to the detectable input situation (cross marker), where it will stay as long as the stimulus is active. The spontaneous state present for the input-free situation disappears at the detectable input situation, forcing the system to evolve to one of the two decision states.

competition-based regime seems to be more plausible because, as becomes clear below, it is generally faster and leads to slow errors, a very desirable feature in decision modeling. Nevertheless, there might be experimental evidence for the first regime (a thermal escape from a locally stable spontaneous state) if one considers stimuli that are very hard to detect.

Finally, the situation in Band IV is comparable to that in Band III, except for the absence of competition regime driven decision making.

As a final remark, it is interesting to note that if one chooses the model's parameters such that $B_1 = B_2 = 0$ (i.e., no input is given) occurs in the first region of three-state stability of Band III, the IDM can sustain decision states even after the external input field has been removed. More specifically, when the stimulus disappears, the free energy function returns to an input-free shape with a spontaneous state and two weaker reminiscent decision states. A process that already evolved to a decision state before the stimulus disappeared is consequently trapped in the reminiscent decision state and can only return to the spontaneous state by accidental thermal fluctuations.

The IDM as a Model for Speeded Two-Choice RT

In this section, we examine the IDM in further detail as a model for the speeded two-choice RT task. First, we recapitulate how the IDM decision-making process works. Second, some prominent characteristics of simulated IDM RT distributions are examined (e.g., shape, relation between mean RT and standard deviation). Third, the probability latency function and the phenomenon of slow and fast errors are studied for the IDM. Then, two possible mechanisms for the SAT are proposed.

The IDM Decision-Making Process: A Time Evolution of Neural Activity

As explained before, in speeded two-choice decision making, the stimulus starts and further steers the decision-making process. For the IDM, we make the practical assumption that the stimulus is sustained long enough for the decision state to be found. At the onset of the stimulus, the free energy surface $F(\boldsymbol{y})$ changes shape such that (for the parameter values given in Table 1 and a certain range of stimulus-generated external fields) two competing decision minima arise. As a result, a systematic drift in the mean activity of the two pools (i.e., y_1 and y_2) leads the process away from the baseline level (the spontaneous state) and toward one of the two possible decision minima of the free energy surface.

All the aforementioned events are illustrated in Figure 7, where one can see the neural activity of the IDM before, during, and after the presentation of a stimulus that generates an external field of size (3,300, 2,700) and lasts for 4 s. Figures 7a and 7b essentially tell the same story, but from a different perspective: In Figure 7a, the (y_1, y_2)-trajectories of the neural activity are plotted on the free energy surface that is guiding them; in Figure 7b, their explicit time evolution is shown. The free energy surface (middle plot of Figure 7a) corresponding to the stimulus-generated external field as well as the time window during which it is applied (Figure 7b) is shaded in gray. In both panels, the different states the process visits are clearly distinguishable. First, there is spontaneous baseline activity in anticipation of the stimulus. Then, upon stimulus presentation, the external field is applied, and a coordinated rise takes place toward the free energy saddle point, shortly after which mutual inhibition drives both pools to a clearly separated sustained

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

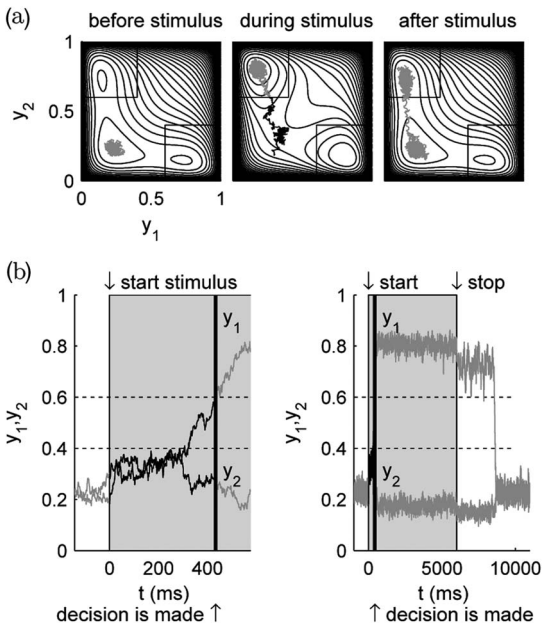


Figure 7. Mean activity of both pools during a trial. The parameter values used are given in Table 1, and the stimulus-generated external field is set at (3,300, 2,700). Panel a shows the evolution of the free energy surface $F(\mathbf{y})$ during an Ising Decision Maker decision trial and a possible decision trajectory. In Panel b, the time evolution of the neural activity of a typical process is plotted explicitly. The 4-s time interval of stimulus presentation (or more specifically, external field activation) as well as the corresponding free energy surface in Panel a are shaded in gray. In Panel b, the left graph zooms in on the decision-time-relevant part of the right graph. The decision time is the time between the presentation of the stimulus and the process crossing a detection box (this moment is indicated with a thick black vertical line), and the corresponding trajectory is indicated in black. Notice the horizontal dotted lines at activities 0.4 and 0.6. They are a translation of the two-dimensional detection boxes: If one pool's activity is above 0.6 while the other pool's activity is less than 0.4, the process is located inside a detection box.

activity (the process is trapped in a decision state). On its way to this decision state, the process crosses a detection box, giving rise to a decision and decision time. The part of the trajectory responsible for the decision time is highlighted in black.

For our current purposes, the decision-making trial ends when a decision is made and the corresponding motor response is initiated. However, it is interesting to see what happens in the case of a prolonged stimulus presentation (assuming the free energy surface does not change with detection and/or response). As long as the external field is active, the process is likely to remain in this decision state. When the external field is removed, the process is trapped in a reminiscent decision state, somewhat less active than the originating decision state. Only after a while, the process picks up enough thermal noise to cross the free energy barrier separating the reminiscent and spontaneous states, returning to the latter.

The parameter values used for the free energy surfaces in Figure 7 can be associated with Band III in Figure 6 and generate reminiscent states. Therefore, there is a delayed return to baseline activity. However, for parameter values in Band II, the return to the spontaneous state will start the instant the stimulus's external field is removed and will typically take up less time than the earlier decision-making process.

Response Time Distributions Generated by the IDM

Many typical experimental features of choice RT distributions also arise from the IDM. Detailed information on the simulations that are used to calculate the IDM probability density functions can be found in Appendix E. The IDM gives rise to a right-skewed RT distribution, as seen in Figure 8a (gray line). An alternative way of looking at the RT distribution is provided by the hazard function, which is the rate of reaching a decision at time t given no decision has been made before t:

$$h_p(t) = \frac{\phi_p(t)}{\pi_p - \Phi_p(t)}.$$

The hazard function (Figure 8a, black line) highlights certain properties of the tail of the original RT distribution. For the parameter values used in Figure 8, the IDM hazard function first increases and then drops again for high RT (the noise at the end of the curve is due to the fact that the results are based on simulations). This type of hazard function is commonly encountered in RT data and can also be obtained with the Ratcliff diffusion model (see Ratcliff & Dongen, 2011).

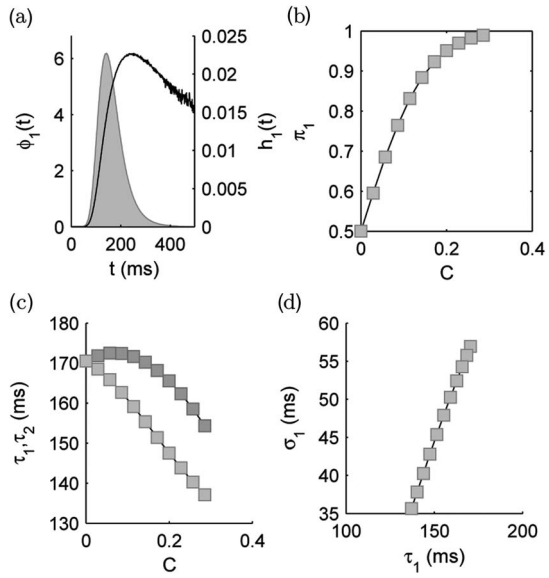


Figure 8. Basic experimental features of choice response time (RT) reproduced by the Ising Decision Maker. Parameter values are taken from Table 1, with the stimulus-generated external field defined by Equation 16 and parameters $B_{ns} = 1{,}000$, $B_s = 2{,}500$. a: The right-skewed correct RT distribution $\phi_1(t)$ for $C = 0.08$ (gray line, bordering the filled area) and the accompanying hazard function $h_1(t)$ (black line). b: Accuracy increases with decreasing stimulus difficulty (stimulus distinctness C varies from 0 to 0.4). c: Mean correct RT τ_1 (lighter gray) decreases with decreasing stimulus difficulty; mean error RT τ_2 (darker gray) first rises somewhat and then decreases. d: Close to linear relation between the mean τ_1 and the standard deviation σ_1 of the correct RT distribution.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Another trademark of choice RTs is that for decreasing stimulus difficulty, accuracy increases and mean correct RT goes down. In the IDM, stimulus distinctness C is the parameter that codes for stimulus difficulty: If $C = 0$, the stimulus is difficult, and if C is close to 1 or -1, the stimulus is easy (this is in line with Wong & Wang, 2006, and Deco et al., 2009). Figures 8b and 8c show how in the IDM the accuracy and the mean RT (for both correct and error responses) are related to increasing difficulty. More interestingly, the IDM satisfies the approximately linear relation between the mean τ_1 and the standard deviation $\sigma_1 = \sqrt{\frac{1}{\pi_1}\int_0^\infty (t-\tau_1)^2 \phi_1(t)dt}$ of the correct RT distribution (with $C > 0$, option 1 is the correct choice) for various levels of stimulus distinctness, as reported by Wagenmakers and Brown (2007). This is shown in Figure 8d.

Probability Latency Functions

In the previous section, we discussed some relatively obvious characteristics of speeded two-choice RT distributions. In this section, we look at a more challenging aspect. In the psychological literature, a typical quantity of interest in speeded two-choice RT experiments is the probability latency function (see Luce, 1986; Ratcliff & Rouder, 1998). It deals with the typical set of stimuli used in speeded choice experiments: stimuli that are comparable in nature but with different intrinsic levels of difficulty. One can consider the basic probability latency function as a parametric plot of mean RT versus accuracy with the stimulus intrinsic level of difficulty as a parameter. In the Ratcliff diffusion model, for example, the stimulus intrinsic level of difficulty is modeled by the drift rate of a one-dimensional diffusion process: It characterizes the stimulus's net evidence pointing to one of the two alternatives. In the IDM, stimulus distinctness C is used to model intrinsic stimulus difficulty and is therefore the parameter that needs to be varied to construct the IDM probability latency function.[8]

For every single value of C, there are two points on the probability latency curve: one on the right side for the correct responses and one on the left side for the incorrect responses (the abscissa points are thus $\pi_2 = 1 - \pi_1$ and π_1, respectively, where π_1 is the proportion correct). Although, experimentally, no two stimulus presentations are exactly the same, we assume that repeated stimuli and stimuli that are considered to have the same evidence ambiguity are characterized by the same distinctness C.

In short, the probability latency function connects accuracy and mean RTs through their mutual stimulus dependency. The probability latency function is a very useful tool in choice RT experiments because the plot can be constructed without specifying an actual value of C for every condition.

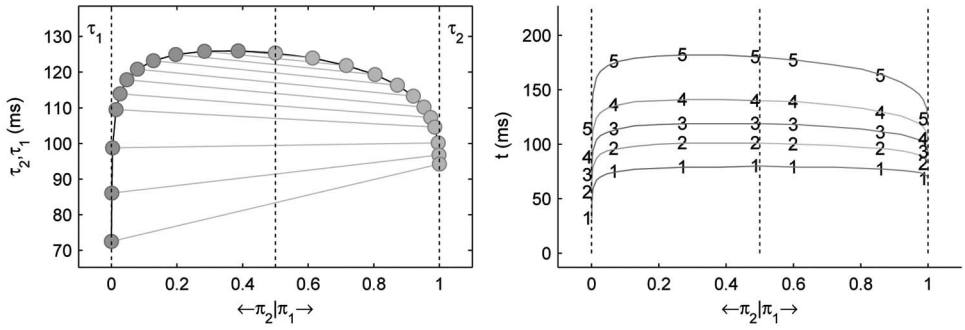
Probability latency functions for speeded two-choice RT experiments have a typical inverted U-shape. Mostly, they show some mild degree of asymmetry around the $\pi_1 = 0.5$ line. The asymmetry is caused by the fact that for a given stimulus, error responses (at the left of $\pi_1 = 0.5$) can be slower than correct responses (at the right of $\pi_1 = 0.5$) or the other way around. Both situations can occur in the same experiment and therefore in the same probability latency function (see, e.g., Ratcliff & Rouder, 1998), but slow errors will always arise at probabilities closer to the 50% accuracy mark than fast errors. The existence of slow and fast errors has been the main reason of the introduction of variability in both drift speed and starting position in the Ratcliff diffusion model (see Ratcliff & Rouder, 1998; Ratcliff, Van Zandt, & McKoon, 1999). The IDM can account for both slow and fast errors without an ad hoc extension of the model (we come back to the fast–slow error balance in the section about SAT). This is illustrated in Figure 9: Figure 9a shows a simple probability latency curve, connecting probabilities to mean RTs, and Figure 9b shows a more detailed picture, considering multiple quantiles of the RT distribution instead of a single mean RT. We discuss the genesis of slow and fast errors in the next two paragraphs.

Slow errors. In the IDM, for $C \neq 0$, slow errors can arise as a consequence of the asymmetric displacement of the free energy saddle point separating the two decision states. Indeed, nonzero distinctness pushes the saddle point toward the erroneous solution (see Figure 3 for a helpful three-dimensional view). The decision trajectories leading to incorrect answers are generally situated closer to the saddle point and are thus more subject to its local influence. This influence has two aspects. First, decision trajectories are bent toward the saddle point, prohibiting a straight approach of the (incorrect) minimum. Second, their net evolution toward the minimum is slowed down because the drift around the saddle point is close to zero. This mechanism is shown in Figure 10. The figure suggests that larger values of distinctness lead to error responses that are increasingly slower than their correct counterparts.

Fast errors. We identify three different mechanisms leading to fast errors in the IDM. First, there are the thermal fluctuations of the spontaneous state. After a short period of evolution with zero input (in anticipation of the stimulus), these random fluctuations automatically lead to a stable distribution of spontaneous activities, independent of the original starting position. The distribution of spontaneous activities at stimulus onset has a similar effect on the eventual probability density functions as the intertrial starting position variability in the Ratcliff diffusion model (e.g., Ratcliff & Rouder, 1998). However, in contrast to the starting position variability in the Ratcliff diffusion model, the IDM spontaneous activity distribution is an integral part of the original model and not an ad hoc extension (although in the Ratcliff diffusion model, the spread in initial position has been proposed as a consequence of so-called premature sampling; see, e.g., Laming, 1968; Rouder, 1996). In analogy to Figure 10, Figure 11 shows how, for the usual parameter values (see Table 1) with $B_s = 2{,}500$, $C = 0.55$, but with changing values for Θ and B_{ns}, the IDM decision process distribution evolves once the stimulus is presented. For each row, parameters Θ and B_{ns} are changed with the same amount. This leads to a different zero-input free energy surface and therefore different spontaneous distribution but leaves the stimulus-generated free energy surface unchanged. Analogous to the effect of starting point variability in the traditional diffusion model (e.g., Ratcliff & Rouder, 1998), a wider spontaneous distribution leads to an increased bias of the initial activity of correct and error decision trajectories. This can be seen in Figure 11, where the overlap between the error and correct process distributions at stimulus presentation ($t = 0$) decreases for wider sponta-

[8] A more refined picture, also taking into account stimulus related parameters B_{ns} and B_s, could be considered, but for the sake of simplicity, we stick to parameter C. The other parameters mentioned play a more prominent role in the discussion of the van der Molen-Keuss effect and Weber's law.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.



(a) Probability latency function (b) Quantiles

Figure 9. Probability latency function for an Ising Decision Maker with parameter values taken from Table 1, except for $\Theta = 51{,}050$ and $\sigma = 0.012$. Additionally, $B_{ns} = 600$, $B_s = 2{,}500$, and C is varied from 0 to 1. Plot a shows a simple probability latency curve based on mean response times. The gray lines connect error responses (π_2, τ_2) on the left (in darker gray) with their correct counterparts (π_1, τ_1) on the right (in lighter gray). For this set of parameters, both slow and fast errors occur. Plot b sketches a more detailed picture of each response time distribution by using response time quantiles (0.1, 0.3, 0.5, 0.7, 0.9).

neous distributions. For sufficiently wide spontaneous distributions, the error slowing effect described in the previous paragraph can be overcome, and error responses can become faster than their correct counterparts. Also analogous to the diffusion model, the effect is most prominent for high values of stimulus distinctness (low difficulty) because only trajectories starting out close to an improbable incorrect decision state have a chance of ever reaching it, effectively reducing their length and duration. Another consequence of the thermal fluctuations of the spontaneous state is that if the period of stimulus anticipation is too long, there is a possibility a decision trajectory will spontaneously cross the decision boundary. This is clearly the case for the time evolution shown in the third row of Figure 11, explaining the nonzero value of the error response probability distribution at $t = 0$. The same nonzero value is present for the correct response probability distribution, but due to the different scales at which error and correct response results are presented (this applies to the trajectory densities as well as the final probability distribution plots), this does not show on the plot. In this article, we choose to leave premature decisions (caused by boundary crossings before the end of a 1-s anticipation period) out of the accuracy and RT results since they are not taken based on a stimulus and typically removed from choice RT data.

A second mechanism that can account for fast errors is related to the discrete nature of the decision trajectories for nonvanishing collective step size σ. As has been explained before, at a microscopic level, we may think of this discreteness as the simultaneous updating of larger sets of neurons, instead of just one neuron at a time. Taking large activity steps while moving across the free energy surface $F(\boldsymbol{y})$ (referred to as a *coarse-grained dynamics*), can cause fast errors, even without an initial spontaneous spread. The effect of collective step size σ on the time evolution of the IDM decision process distribution is illustrated in Figure 12. Larger collective step size (more coarse-grained dynamics) leads to a decrease in accuracy and to error responses speeding up even more than correct responses. For sufficiently large collective step size, this results in error responses that are faster than their correct counterparts. Fast errors dominate at low difficulty (or high stimulus distinctness) as a consequence of the interaction between the coarse-graining mechanism and the shape of the free energy surface. Especially for high stimulus distinctness, the incorrect response side of the surface close to the starting point of the process is concave (a bump), while the correct response side is convex (a valley). The coarse-grained dynamics allows the process to punch through the flanks of the error response bump, effectively cutting corners. In the correct response valley, there are no corners to cut, so this advantage is lost. Because the explained effects are a direct consequence of the coarse-graining mechanism, they disappear in the diffusion limit. In this limit, collective step size σ is directly related to the diffusion constant D (see Equation 20). Changing D, however, is no more than a rescaling of time and does not influence accuracy or the fast–slow error balance.

A third mechanism leading to fast errors arises for high inputs only. Looking at position d of Band III (in the second region of three-state stability) and the corresponding free energy surface in Figure 6b, we find a configuration with two decision states and a nondecision state situated at high activity for both pools (a so-called startled state). Only a sufficiently high external field B can move a system in anticipation of a stimulus across the competition region into the second region of three-state stability where this configuration is found. A comparable situation arises in the integrate-and-fire context (Wong & Wang, 2006), but the specific effect of a startled state on choice RT distributions was not investigated. In Figure 13, the effect of the startled state on the IDM decision process distribution is illustrated. For sufficiently distinct (easy) stimuli, the startled state merges with the correct decision state, creating an attractive funnel toward the correct decision boundary. This funnel effectively reroutes and slows down the correct response trajectories. Decision trajectories well on their way to an error response can still fall into the funnel and will then take a much longer time to eventually end up as a correct response. For a startled state to have an impact, it is pivotal to use detection boundaries that allow decision trajectories to get close to the startled state before resulting in a choice. This is the case for the boundary boxes used in this article. Using traditional perpendicular decision boundary lines, however, will not produce this

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

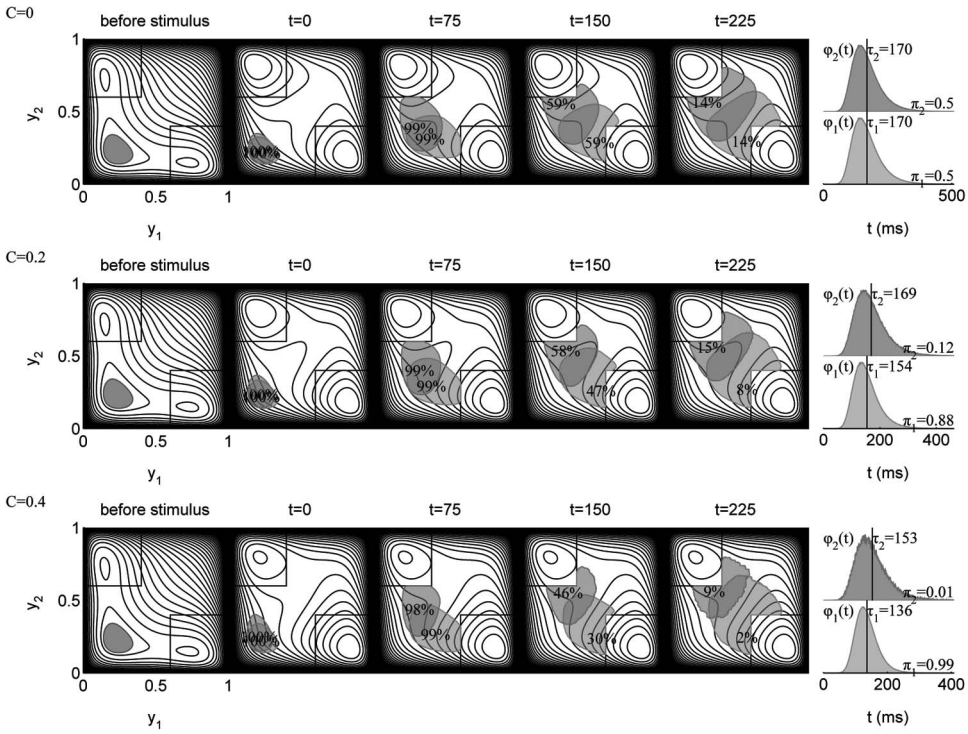


Figure 10. Genesis of slow errors. Time evolution of the ***y***-plane distribution of possible decision processes for an Ising Decision Maker operating in the middle of the competition regime (parameter values are taken from Table 1, with $B_{ns} = 1{,}000$, $B_s = 2{,}500$). The subsequent rows consider different values of stimulus distinctness *C*. At the end of each row, the resulting probability distribution ϕ_p for both correct ($p = 1$, in lighter gray) and incorrect ($p = 2$, in darker gray) responses are shown as well as their respective mean response times and choice probabilities. In subsequent time evolution graphs, the lighter gray cloud represents the normalized distribution of processes eventually leading to a correct decision (lower right corner), the darker gray cloud its incorrect counterpart (upper left corner). When processes hit the decision boundary, they are removed from their cloud. The percentages inside the clouds indicate how many of the respective processes are still running and have not yet reached their decision boundary.

type of fast errors. This third mechanism of fast errors is probably not the one typically found in common speeded two-choice RT experiments as it is only available for very high network input and is more prominent for moderately coherent stimuli than for very coherent stimuli, contrary to the typical experimental picture. However, there are experimental accounts of fast errors manifesting themselves only when their physical intensity (not their distinctness) is strongly amplified. We come back to these results and the relevance of the startled state when discussing the van der Molen-Keuss effect.

Speed–Accuracy Tradeoff

In the rich modeling tradition of the (instruction-induced) SAT, one of two routes is generally taken. SAT can be reached by fine-tuning some of the parameters of the evidence-integrating dynamical system (e.g., the boundary separation in the Ratcliff diffusion model; Ratcliff & Rouder, 1998), or it can be the result of a strategic mix of an unchanged evidence-integrating dynamical system on the one hand and a guessing system on the other hand (Yellott, 1971). Both mechanisms are however not mutually exclusive (Osman et al., 2000). In this article, we take the first route and investigate two possible mechanisms of decision-process-related SAT. We first investigate the traditional choice of detection boundary separation as a parameter for speed–accuracy manipulation. We then propose a novel mechanism of speed–accuracy manipulation, based on the idea of coarse-grained dynamics, which has already been introduced as a possible explanation of fast errors. Although there are other possible approaches to controlling the SAT in the IDM (e.g., by increasing the nonselective input, as suggested by Furman & Wang, 2008), we limit ourselves to the two approaches mentioned above.

In most current models of decision making, speed–accuracy manipulation is modeled by changes in the conditions that end the decision process. In the standard diffusion model, common practice is to change the boundary separation parameter. For the IDM, this can be easily translated as a change in box size *h*. In Figure 14, we show how changing box size impacts the probability latency functions for an IDM with parameter values from Table 1. To allow for a clean diffusion interpretation (no coarse graining), we take a diffusion limit setting with $D = 0.05$. As expected from the analogy with the diffusion model, the basic requirements of speed–accuracy are met: An increase in speed is accompanied by a decrease in accuracy and an

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

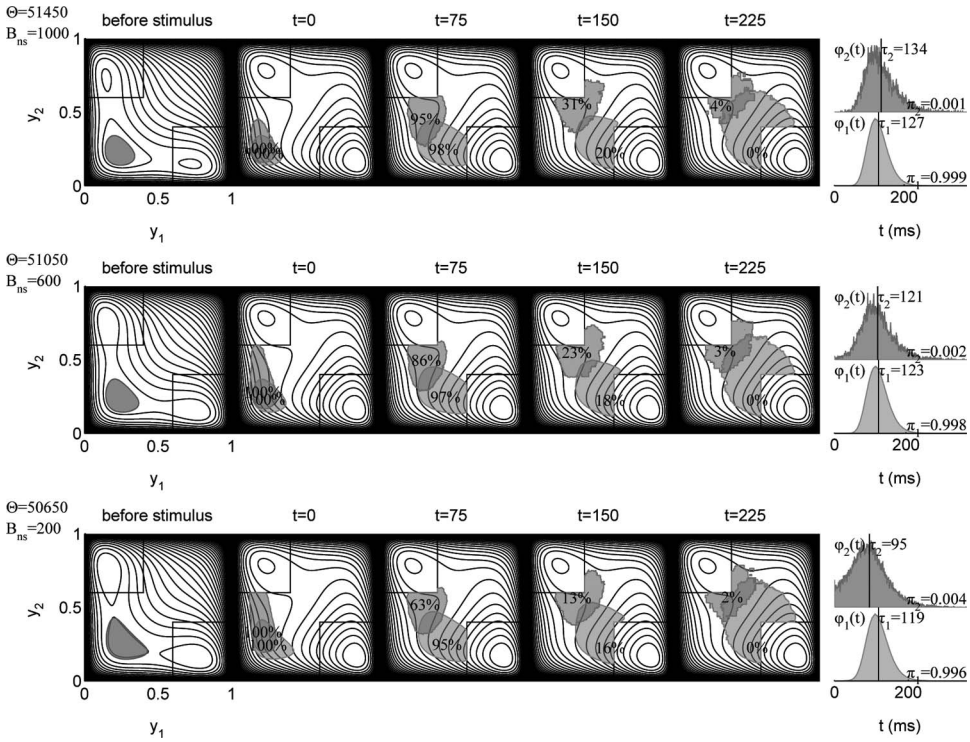


Figure 11. Fast errors as a consequence of the spontaneous distribution. Time evolution of the ***y***-plane distribution of Ising Decision Maker decision processes. Parameter values are taken from Table 1 with B_s = 2,500, C = 0.55. Different rows consider different values for the remaining parameters Θ and B_{ns} resulting in different spontaneous distributions, as can be seen on the first plot of each row. Wider spontaneous distributions clearly lead to error responses becoming faster than their correct counterparts (see the subsection on fast errors for more explanation). For more information on these graphs, we refer the reader to Figure 10.

increase in the proportion of errors that are faster than their correct counterpart. The figure also shows that all three basic probability latency patterns observed in experiments can be reproduced: curves with exclusively slow errors, curves with exclusively fast errors, and curves with a combination of both slow and fast errors.

Recent neurophysiological findings suggest, however, that the magnitude of neural activity triggering a response is in fact invariant under speed–accuracy manipulations (Heitz & Schall, 2012, p. 626). This would rule out boundary separation as a simple explanation of SAT. In what follows, we test the ability of coarse-graining parameter collective step size σ to control the SAT in the IDM in a realistic way. A SAT mechanism based on coarse-grained dynamics does respect the detection boundary invariance proposed by Heitz and Schall (2012). Figure 15 illustrates how increasing collective step size σ simultaneously increases processing speed, decreases decision accuracy, and decreases the lag of slower error responses, eventually even leading to fast errors (as has been explained already in the previous section). As it turns out, this is exactly what is found in typical speeded two-choice decision experiments (e.g., Ratcliff & Rouder, 1998). Also in this case, the figure shows that all three basic probability latency patterns observed in experiments can be reproduced: curves with exclusively slow errors, curves with exclusively fast errors, and curves with a combination of both slow and fast errors. For constant time step Δt, the increase in processing speed produced by increasing collective step size σ is clearly exaggerated. Therefore, the second dynamical parameter Δt is considered to be an additional parameter modeling SAT.

From an algorithmic viewpoint, coarse graining explains SAT in a very intuitive way. In order to solve the same decision problem (i.e., finding the correct decision state or global minimum of the free energy), one may use different dynamical strategies: an accurate but slow strategy (with small collective step size, bringing the IDM closer to local gradient behavior) versus a quick and dirty strategy (larger collective step size, effectively coarse graining). Another recent study seems to support our interest in coarse graining: Torres, Marro, Cortes, and Wemmenhove (2008) discovered that varying the fraction of simultaneously updated neurons (which also amounts to a form of coarse graining) is a way to control the efficiency of a neural searching process.

Application of the IDM to Data

In this section, how the IDM can be fitted to data is demonstrated. In line with the two different mechanisms for SAT that have been discussed above, two different versions of the IDM are considered.

The first version allows box size h to change between SAT conditions. As discussed above, box size is closely related to the standard notion of boundary separation in the traditional diffusion model. To facilitate a clear comparison with the existing class of diffusion models, we keep to the IDM diffusion limit, meaning it suffices to estimate a diffusion constant D instead of both dynamical parameters

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

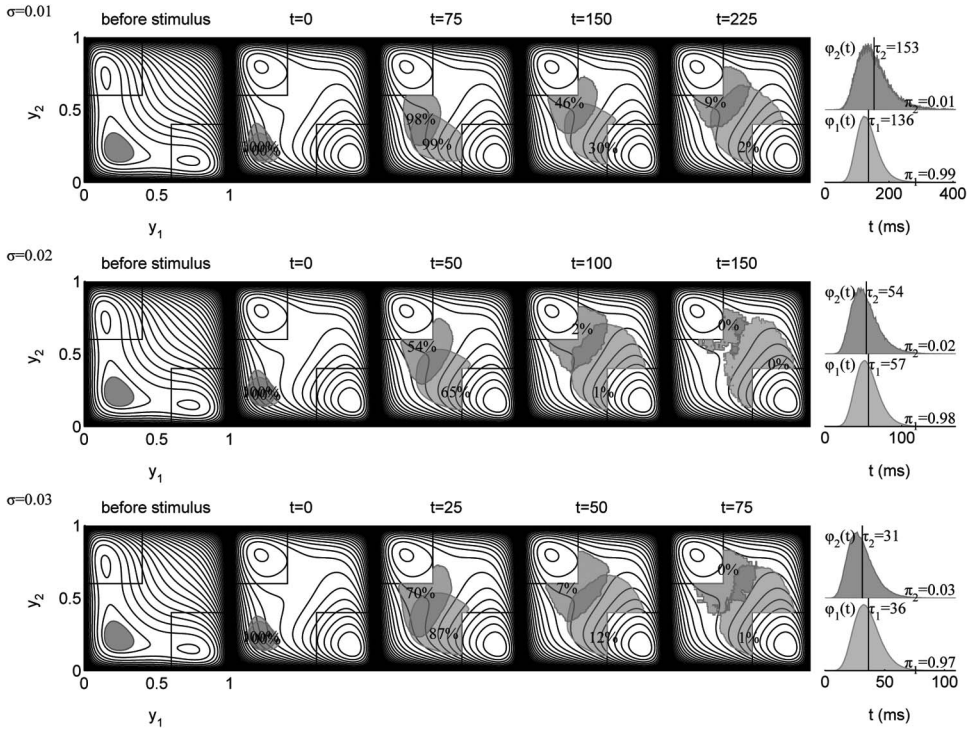


Figure 12. Fast errors as a consequence of coarse-grained dynamics. Time evolution of the ***y***-plane distribution of Ising Decision Maker decision processes. Parameter values are taken from Table 1 with $B_{ns} = 1{,}000$, $B_s = 2{,}500$, $C = 0.4$, and changing σ. Different rows consider different collective step sizes σ, arranged from small to large. Increasing collective step size (using more coarse-grained dynamics) leads to a decrease in accuracy, faster responses in general, and error responses speeding up even more than their correct counterparts, eventually leading to error responses that are faster than the corresponding correct responses. For more information on these graphs, we refer the reader to Figure 10.

Δt and σ (see Equation 20). We call this version IDM BS (from boundary separation or box size).

The second version of the IDM allows the dynamical parameters of the IDM, namely, σ and Δt, to change between SAT conditions. We call it IDM CG, referring to the coarse-graining mechanism that is used for SAT manipulation. Evidently, the diffusion limit is parted from in this case. The detection box size, on the other hand, is fixed to 0.4 in the IDM CG.

In addition, we want to illustrate that (both versions of) the model can compete with more traditional models, in particular the Ratcliff diffusion model.

Statistical Inference

Our statistical inferential framework is Bayesian, a choice we made for two reasons. First, recent research has made a strong case

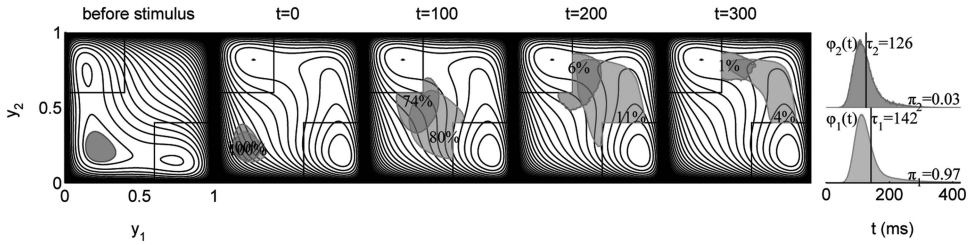


Figure 13. Fast errors as a consequence of a startled state. Time evolution of the ***y***-plane distribution of Ising Decision Maker decision processes under influence of a startled state merged with the correct decision state. Parameter values are taken from Table 1 with $B_{ns} = 2{,}500$, $B_s = 2{,}500$, $C = 0.4$. A substantial part of the correct decision processes is caught by a funnel, delaying their passage into the detection box. This results in slower correct responses or, relatively speaking, faster errors. For more information on these graphs, we refer the reader to Figure 10.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

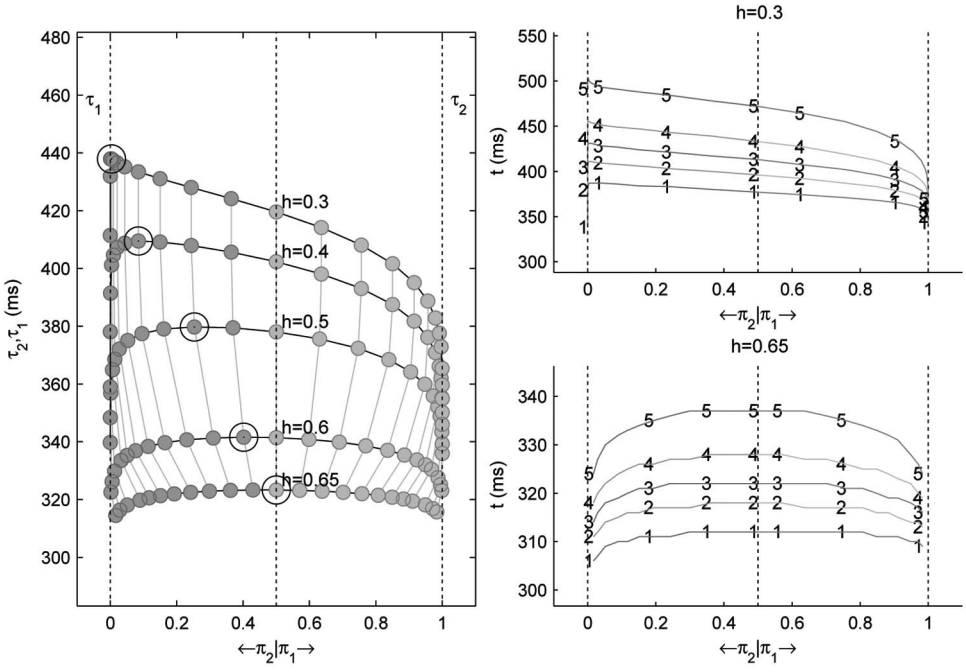


Figure 14. Probability latency curves for increasing box size h (left panel). The other parameter values are taken from Table 1, except for T_{er}, which is taken to be 0.3 s. Additionally, to allow for a clean diffusion interpretation (no coarse graining), we take a diffusion limit setting with $D = 0.05$. For the input, we take $B_{ns} =$ 1,000, $B_s = 2{,}500$, and C is varied from 0 to 0.4 with steps of 0.05 and further to 0.7 with steps of 0.1. Speed increases and accuracy decreases with increasing box size h. Additionally, errors that are slower than their correct counterparts are gradually transformed into errors that are faster, while the curve's maximum (indicated with an extra circle) moves from left to right. On the right panels, quantile probability latency representations of two of the curves on the left panel ($h = 0.4$, $h = 0.65$) are shown, providing some additional detail.

in favor of Bayesian statistics (see, e.g., Gelman, Carlin, Stern, & Rubin, 2003; Kruschke, 2011; Lee & Wagenmakers, 2013). Second, modern computational methods that can be applied for Bayesian statistical inference facilitate the fitting of complex models (Gelman et al., 2003), such as the IDM.

Let us collect all parameters from the IDM in a parameter vector $\boldsymbol{\alpha}$, and $\boldsymbol{Y}$ are the data. Then, at the heart of Bayesian statistics is Bayes's theorem:

$$p(\boldsymbol{\alpha} \mid \boldsymbol{Y}) = \frac{p(\boldsymbol{Y} \mid \boldsymbol{\alpha})p(\boldsymbol{\alpha})}{p(\boldsymbol{Y})}, \tag{33}$$

where $p(\boldsymbol{\alpha} \mid \boldsymbol{Y})$ is the posterior distribution, $p(\boldsymbol{\alpha})$ is the prior distribution, and $p(\boldsymbol{Y} \mid \boldsymbol{\alpha})$ is the likelihood. The denominator $p(\boldsymbol{Y})$ is called the marginal likelihood.

Fitting the IDM. The goal of a Bayesian analysis is to explore and summarize the posterior distribution $p(\boldsymbol{\alpha} \mid \boldsymbol{Y})$. This requires the clarification of three necessary aspects: the likelihood $p(\boldsymbol{Y} \mid \boldsymbol{\alpha})$, the prior distribution $p(\boldsymbol{\alpha})$, and the methods to explore the posterior distributions (the marginal likelihood $p(\boldsymbol{Y})$ is not necessary for the exploration of the posterior distribution). These three aspects are discussed in turn.

First, with respect to the likelihood of the IDM, we build on the work of Heathcote, Brown, and Mewhort (2002), who devised a quantile maximum likelihood method or quantile maximum product method (see also Heathcote & Brown, 2004; Speckman & Rouder, 2004). Instead of the likelihood of the raw data (choices and RTs), the likelihood of a set of order statistics of the data is considered because it leads to more robust estimation algorithms. More specifically and applied to the problem at hand, we start by calculating the quantiles corresponding to the five cumulative probabilities .1, .3, .5, .7, and .9 for the correct response data of a single experimental condition. Denote the calculated quantiles with the vector $\boldsymbol{q}$. Let vectors $\boldsymbol{n}$ and $\boldsymbol{m}$ contain the number of, respectively, correct and error responses observed between each set of consecutive quantiles. Thus, the data we consider are $\boldsymbol{Y} = (\boldsymbol{q}, \mathbf{n}, \mathbf{m})$. Note that the quantiles used for the binning of the error RT distribution originate from the accompanying correct RT distribution. Applying the correction of Speckman and Rouder (2004) to the original formula proposed in Heathcote et al., the quantile likelihood is then defined as

$$\begin{aligned} p(\boldsymbol{Y} \mid \boldsymbol{\alpha}) = {} & \Phi_1(q_1; \boldsymbol{\alpha})^{n_1 - 1} \cdot \phi_1(q_1; \boldsymbol{\alpha})[\Phi_1(q_2; \boldsymbol{\alpha}) - \Phi_1(q_1; \boldsymbol{\alpha})]^{n_2 - 1} \\ & \cdot \ldots \cdot \phi_1(q_5; \boldsymbol{\alpha})[1 - \Phi_1(q_5; \boldsymbol{\alpha})]^{n_5} \\ & \cdot \Phi_2(q_1; \boldsymbol{\alpha})^{m_1}[\Phi_2(q_2; \boldsymbol{\alpha}) - \Phi_2(q_1; \boldsymbol{\alpha})]^{m_2} \\ & \cdot \ldots \cdot [1 - \Phi_2(q_5; \boldsymbol{\alpha})]^{m_5}, \end{aligned}$$

where $\Phi_1(t; \boldsymbol{\alpha})$ and $\Phi_2(t; \boldsymbol{\alpha})$ are the cumulative distribution functions of respectively correct and error RTs for parameter vector $\boldsymbol{\alpha}$ and $\phi_1(t; \boldsymbol{\alpha})$ is the probability density function of the correct RTs. Note that the correction of Speckman and Rouder only applies to

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

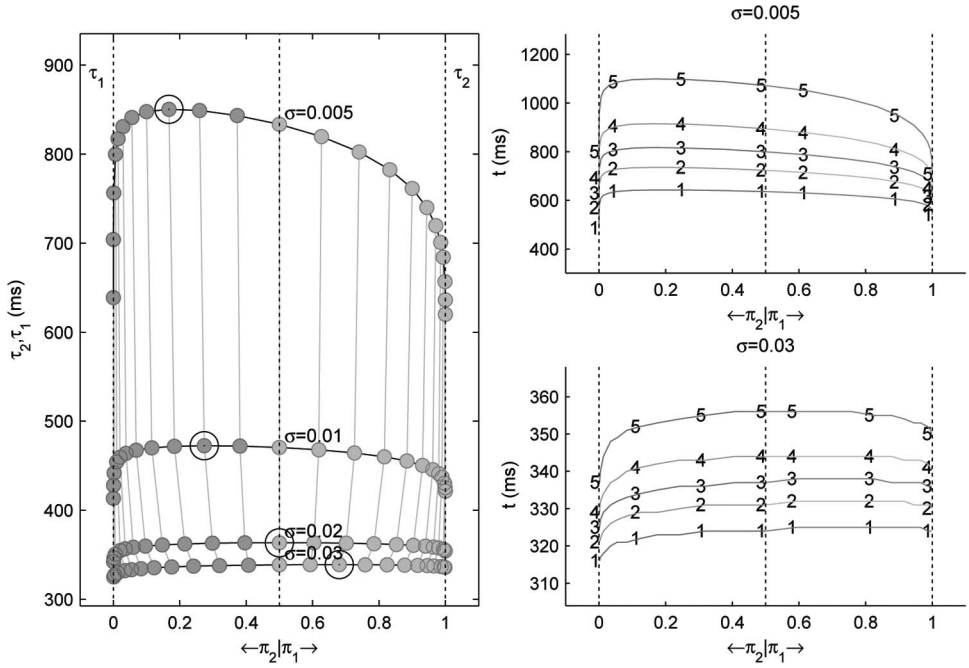


Figure 15. Probability latency curves for increasing collective step size σ (left panel). The other parameter values are taken from Table 1, except for $T_{er} = 0.3$ and $\Theta = 51{,}050$. For the input, we take $B_{ns} = 1{,}000$, $B_s = 2{,}500$, and C is varied from 0 to 0.4 with steps of 0.05 and further to 0.7 with steps of 0.1. Speed increases and accuracy decreases with increasing collective step size σ. Additionally, errors that are slower than their correct counterparts are gradually transformed into errors that are faster, while the curve's maximum (indicated with an extra circle) moves from left to right. On the right panels, quantile probability latency representations of two of the curves on the left panel ($\sigma = 0.005$, $\sigma = 0.03$) are shown, providing some additional detail.

the correct RT distribution because, for the error RT distribution, we reuse the quantiles determined from the correct RT distribution. Unfortunately, for the IDM, there is no analytical expression for the density or distribution functions. They can however be obtained through simulations. A more detailed explanation of the simulations and the construction of the quantile likelihood can be found in Appendix E. More conditions (stimuli or other) can be incorporated by simple multiplication.

Second, with respect to the prior, we assume a uniform distribution for the distinctness parameter C on the interval [–1, 1] and, in line with their interpretation, a strictly positive improper uniform distribution for all other parameters. For collective step size σ, the uniform distribution was assumed on the level of its quadrature σ^2.[9] The resulting prior distribution of the diffusion constant $D \sim \frac{\sigma^2}{\Delta t}$ when explicitly considering the diffusion limit is also uniform (strictly positive and improper).

Third, regarding the methods for exploring the posterior distribution, we want to draw L sample parameter vectors from the posterior distribution $p(\boldsymbol{\alpha} \mid \boldsymbol{Y})$. Using the sampled parameter vectors $\boldsymbol{\alpha}^1, \boldsymbol{\alpha}^2, \ldots, \boldsymbol{\alpha}^L$, we can plot histograms (to get an idea of the univariate marginal distributions), construct scatterplots (to investigate the posterior parameter correlations), and calculate sample statistics such as the mean (i.e., an estimate of the posterior mean) and the posterior credibility intervals.

However, due to the complexity of the problem (a high-dimensional posterior distribution with no closed-form analytical expression), there is no simple and straightforward sampling method. We have chosen an adaptation of the traditional Metropolis-Hastings method: the delayed rejection adaptive Metropolis sampling algorithm (Haario, Laine, Mira, & Saksman, 2006). The algorithm differs in two crucial aspects from the Metropolis-Hastings sampler: The rejection of a proposal is delayed while other proposals are tested, and the variance of the proposal distribution is updated every so many steps. The latter adaptation makes the algorithm non-Markovian, but it has been proven by Haario et al. (2006) that a sufficiently large set of samples still converges to the posterior distribution. As is the case for all Markov chain Monte Carlo techniques, the samples generated with the delayed rejection adaptive Metropolis algorithm are not independent of one another. In comparison to the traditional Metropolis-Hastings method, however, the delayed rejection adaptive Metropolis algorithm has a better mixing rate (i.e., a smaller autocorrelation in its successive samples), resulting in faster convergence to the posterior.

In order to provide the delayed rejection adaptive Metropolis sampler with a suitable starting point, it is preceded by a stochastic mode seeking algorithm, the differential evolution optimizer (Storn & Price, 1997). The differential evolution optimizer is run

[9] In other words, the IDM is actually parameterized with σ^2 rather than σ. The fact that the correlation between σ^2 and Δt in the diffusion limit is exactly linear (see Equation 20) and not curved as it would be for σ and Δt makes the resulting posterior parameter distribution a lot easier to sample.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

10 times, and the best solution is then used as a starting point for the sampler. This way, the chain starts in a region of substantial posterior mass, effectively reducing the start-up overhead (burn-in) of the sampler. Next, the delayed rejection adaptive Metropolis algorithm is initiated and run for a burn-in period (the length of which depends on the problem at hand, but in this case, it seems to be of the order of 10^4 to 10^5). After burn-in, the delayed rejection adaptive Metropolis is run for another several 10^4 or 10^5 iterations. The chains are then inspected visually to assess convergence. For the calculations of the posterior quantities, the burn-in iterations are ignored.

Identification of the IDM parameters. The parameter vector $\boldsymbol{\alpha}$ contains the free parameters of the model. The final number of free parameters varies with the experimental design. The real data discussed below originate from two typical two-choice RT experiments with six different stimulus-difficulty levels (in one of the two experiments, a broader range of stimulus difficulties was reduced to six levels to simplify the analysis), accompanied by a speed or accuracy instruction. This means we estimate six separate distinctness parameters C and two sets of SAT-sensitive parameters, one set for the speed and one set for the accuracy instruction.

However, for identification reasons, two parameters have to be fixed. First, as mentioned above, the nonselective stimulus strength B_{ns} cannot be disentangled from B_s and C in a standard two-choice RT experiment. Therefore, we fix it to zero without loss of generality. Second, in a series of pilot studies, a large posterior correlation between some of the parameters was noticed. More specifically, there was an almost exact linear relationship between N, W^+, and Θ. The reason for this near-identification problem is further investigated in Appendix D. Although not an exact tradeoff, it results in large posterior correlations and slow convergence. Because the issue of slow convergence should be dealt with, we fix Θ to the value used for the simulations (see Table 1), solving the problem. However, it should be noted that because the identification problem is not exact, using more data or other experimental manipulations (selectively targeted at Θ) may remove the need for fixing Θ.

As a result, the IDM BS has a total of 14 free parameters: N, W^+, W^-, B_s, D, h (accuracy), h (speed), C_i (with i ranging from 1 to 6), and a constant nondecision time T_{er}. The IDM CG has a total of 15 free parameters: N, W^+, W^-, B_s, Δt (accuracy), σ (accuracy), Δt (speed), σ (speed), C_i (with i ranging from 1 to 6), and a constant nondecision time T_{er}.

Fitting the Ratcliff diffusion model. The parameter estimation methodology described above is also used for fitting the Ratcliff diffusion model. Again, the differential evolution algorithm is used to generate a suitable starting point for the delayed rejection adaptive Metropolis method, which is then used to sample from the posterior. For the Ratcliff diffusion model, we do not have to resort to simulations when evaluating the cumulative distribution functions in the predetermined quantiles as several of its mathematical properties can be and have already been exploited for fast numerical implementations: for example, the analytical expression of Tuerlinckx (2004) and the Kolmogorov backward approach of Voss and Voss (2008). In our calculations, we use the code of Voss and Voss. Three different variants of the Ratcliff diffusion model will be fitted (denoted as RDM 1, RDM 2, and RDM 3). These three versions differ with respect to which parameters are affected by a speed–accuracy manipulation. The first version (RDM 1) is the standard Ratcliff diffusion model, with variability in starting position and drift speed but without variability in T_{er} (as is the case for the IDM). Boundary separation is the only parameter that is affected by the speed–accuracy instruction. In the second version (RDM 2), variability is added for T_{er}. In the third version (RDM 3), variability for T_{er} is again included, and all parameters except for the stimulus drift rates may be affected by the speed–accuracy instruction.

Model selection. Given several competing models, one may ask which one fits the data best. A model that is too complex may be overfitting the data, thereby not generalizing very well to future data. Thus, because goodness of fit in itself is not a sufficient criterion to decide between competing models, a good balance between goodness of fit and complexity is desirable (see, e.g., Pitt & Myung, 2002). In this article, we use the deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & Van Der Linde, 2002), which is a compromise between (Bayesian) fit on the one hand and model complexity on the other hand.

Simulated Data

To check the validity of our parameter estimation methodology, we apply the method to simulated data sets. For both the BS and CG versions of the IDM, data are simulated based on the parameter values in Table 1 (nondecision time T_{er} is taken to be 0.3 s). In addition, we assume that six different stimulus difficulties are presented under two speed–accuracy conditions. This leads to six parameter values for C (ranging from 0 to 0.25 in steps of 0.05). We take $B_{ns} = 0$ as discussed in the section about parameter identification, and $B_s = 2{,}500$. The BS version of the IDM is considered in the diffusion limit, so the dynamical parameters σ and Δt are replaced by a diffusion constant D (see Equation 20). We take $D = 0.05$. The BS approach to SAT requires two different box sizes: We take $h = 0.25$ for the accuracy condition and $h = 0.45$ for the speed condition. For the CG version of the model, the SAT-sensitive parameters are σ and Δt. For the accuracy condition, we take $\Delta t = 6.25 \cdot 10^{-6}$ and $\sigma = 5 \cdot 10^{-4}$; for the speed condition, we take $\Delta t = 6.25 \cdot 10^{-3}$ and $\sigma = 5 \cdot 10^{-2}$. For both versions of the IDM, we simulate a realistic number of 300 trials per stimulus condition.

In Figure 16, we show the evolution of the delayed rejection adaptive Metropolis chain after burn-in for the parameters T_{er} (for IDM BS) and N (for IDM CG). The other parameters show similar behavior, which makes it reasonable to assume that the chains have converged. In Figure 17, the upper line of each parameter graph (labeled as SD, from simulated data) compares the true parameter value of the simulated data (open circles) with the corresponding posterior means (crosses) and 95% credibility interval (line segments). As can be seen, in the vast majority of cases, the 95% posterior credibility interval encompasses the true value. Recovery appears to be somewhat better for the IDM CG than for the IDM BS. Additional analyses (not reported) showed that setting the detection box as a free parameter significantly increases the amount of data needed for good recovery. For 3,000 simulated data points per condition, we do find a perfect recovery for the IDM BS. Considering the number of trials per stimulus condition for the real data (often in the range of 300 to 500), it must be said that the estimated parameter values (and resulting free energy surfaces) should be interpreted with some caution in the case of the IDM BS.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

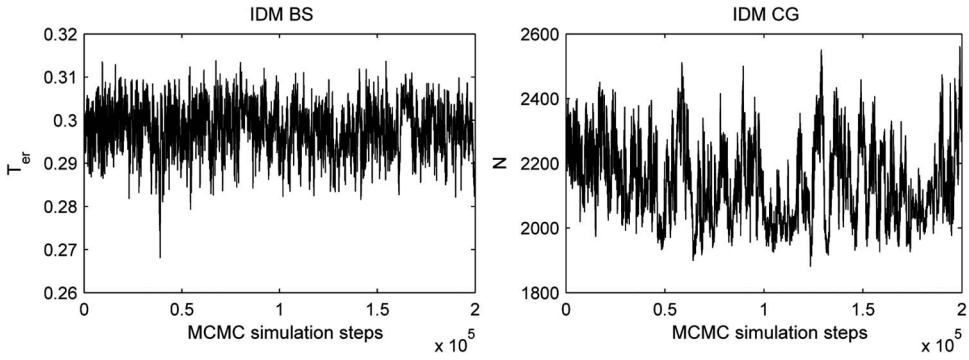


Figure 16. IDM Markov chain after burn-in for parameters T_{er} (IDM BS) and N (IDM CG) based on simulated data. The generating parameter values are indicated in Figure 17 as open circles. IDM = Ising Decision Maker; BS = boundary separation; CG = coarse graining; MCMC = Markov chain Monte Carlo.

Second, we simulate data for the three versions of the Ratcliff diffusion model (RDM 1, RDM 2, and RDM 3, as defined above). The top row in the different panels of Figure 18 displays again the true value (circle), the posterior mean (cross), and the 95% credibility interval (line segments). For all three versions of the Ratcliff diffusion model, the recovery is excellent.

Real Data

The two versions of the IDM and the three versions of the Ratcliff diffusion model described above are fitted to real data collected from two different experiments. The first experiment is the brightness discrimination task by Ratcliff and Rouder (1998),

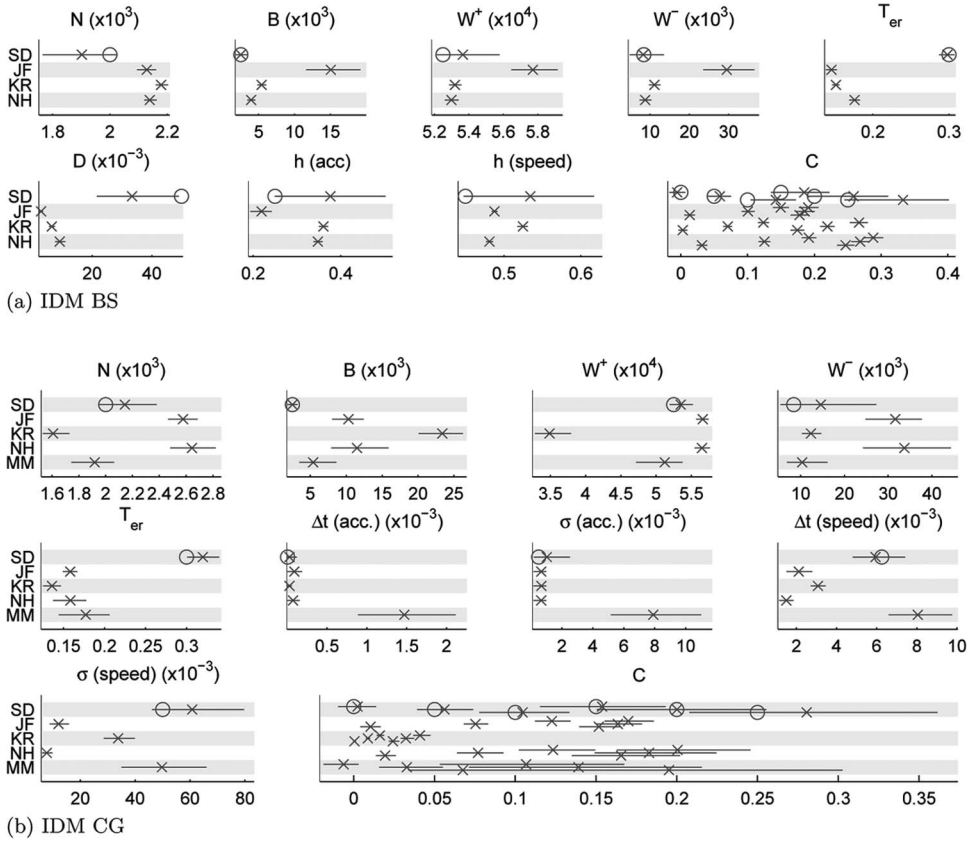


Figure 17. Summary of the parameter estimates of IDM BS and IDM CG based on simulated and real data (six difficulty levels, two speed–accuracy instructions). Every graph displays the estimates of one parameter. Estimates are shown for the simulated data set SD (based on the values in Table 1), participants JF, KR, MH (Ratcliff & Rouder, 1998), and MM (Mulder et al., 2013). The crosses indicate the posterior means, and the line segments indicate the 95% posterior credibility intervals. IDM = Ising Decision Maker; BS = boundary separation; CG = coarse graining.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

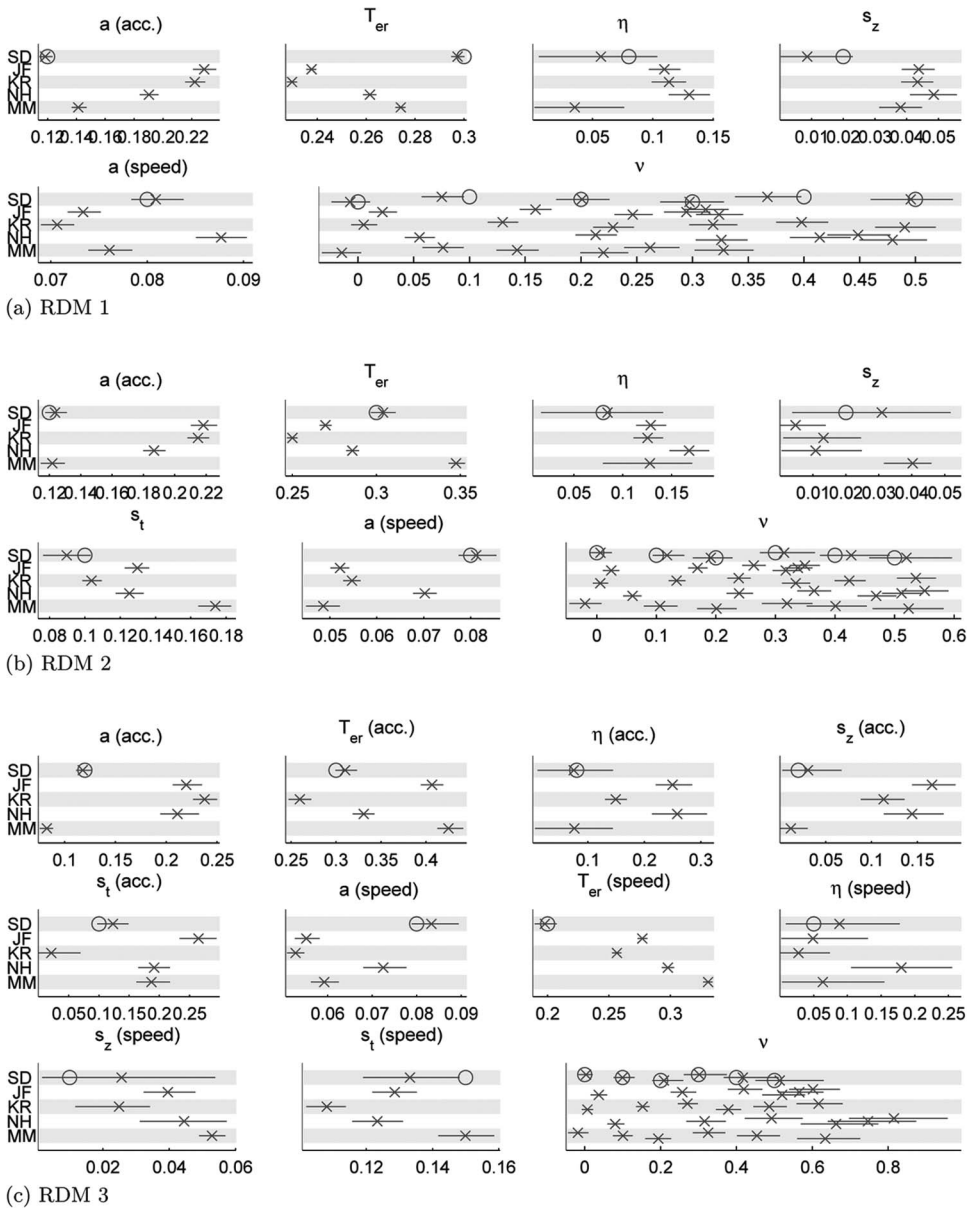


Figure 18. Summary of the parameter estimates of different versions of the Ratcliff diffusion model (RDM) based on simulated and real data (six difficulty levels, two speed–accuracy instructions). Every graph displays the estimates of one parameter. Parameter nomenclature is taken from Ratcliff and Rouder (1998). In all models, starting point $z = a/2$. Estimates are shown for the simulated data set SD (the generating values are indicated with open circles), participants JF, KR, NH (Ratcliff & Rouder, 1998), and MM (Mulder et al., 2013). The crosses indicate the posterior means, and the line segments indicate the 95% posterior credibility intervals.

which has been used as an empirical benchmark before (Brown & Heathcote, 2008). Three participants (JF, KR, and NH) had to classify patches of pixels as either *bright* or *dark*. The brightness of the pixel patch was manipulated in 33 levels, from very dark to very bright. We have binned the 33 levels into six stimuli. The correct classification of the stimulus entailed some degree of randomness: For very bright or very dark stimuli, the classification into the *bright* or *dark* category, respectively, was almost sure, while for stimuli halfway between dark and bright, the classification was fully random. All stimuli were presented under two speed–accuracy manipulations: In one manipulation, there was more emphasis on speed, while in the other, there was more emphasis on accuracy. This manipulation was implemented through the instructions but also through direct feedback (e.g., telling the participants that their response was too slow under the speed manipulation or that an error was made in the accuracy manipulation).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

The other experiment is a recently published random-dot motion task (Mulder et al., 2013). During the experiment, six different stimulus-difficulty levels were presented (i.e., six different motion-coherence levels: 0%, 5%, 10%, 20%, 40%, and 80%), and there was also a speed–accuracy manipulation, which was implemented again through a combination of instructions and feedback (see above). For this experiment, each condition (a crossing of a difficulty level and an instruction) contained 300 trials. In total, there were 1,800 trials. For this experiment, we only analyze the data of one participant (denoted as MM).

The chains generated by the delayed rejection adaptive Metropolis algorithm were similar to the chains found for the simulated data, all showing good convergence, except for the IDM BS parameter estimates of participant MM. The nonconverging estimates correspond to a monostable free energy surface, both before and during stimulus presentation. When the stimulus is presented, the minimum moves to slightly higher values of mean activity for both pools (cf. Band I in Figure 6). The detection box size for both the speed and accuracy conditions are very close to 0.5. This type of free energy surface can be sufficiently described by an ordinary well-shaped LCA and does not require higher order entropy terms. This creates a parameter redundancy for the IDM that could be addressed by fixing another parameter, effectively reducing the IDM to an LCA-type model. As we have no valid parameter estimates for the BS version of the IDM combined with participant MM, this particular combination is left out of the further analysis. An overview of all other parameter estimates can be found in Figure 17. The parameter estimates belonging to the different versions of the Ratcliff diffusion model are shown in Figure 18.

A visual comparison of IDM and Ratcliff diffusion model fits of participant JF (from Ratcliff & Rouder, 1998) is shown in Figure 19. The quantile probability plots show that all models have relatively good fits but that both versions of the IDM and RDM 3 are performing best (the two other Ratcliff diffusion model variants show some degree of misfit).

Next, we rank the models according to the DIC model-selection criterion, taking into account both goodness of fit and model complexity. The results are given in Table 2. Both versions of the IDM always outperform RDM 1 (the traditional version, as proposed in Ratcliff & Rouder, 1998) and RDM 2 (allowing for additional uniform variability of the nondecision time, which is not included in the IDM). RDM 3 (an extension of RDM 2, for which every parameter except the drift rates may change under the speed–accuracy manipulation), however, is always at least as good as both versions of the IDM. For two out of three participants, IDM CG outperforms IDM BS.

Piéron's Law, the van der Molen-Keuss Effect, and Weber's Law

In this section, we use the IDM to explain some additional empirical findings from two-choice RT tasks. First, we link the IDM to the lesser known Piéron's law and van der Molen-Keuss effect. In addition, the IDM is able to reproduce Weber's law, a classical psychophysical relation.

Piéron's Law and the van der Molen-Keuss Effect

Piéron's law for speeded two-choice RT (Pins & Bonnet, 1996; see also Smith & Ratcliff, 2009) and the van der Molen-Keuss effect (Jaśkowski & Włodarczyk, 2006; van der Molen & Keuss, 1979) are two firmly replicated experimental findings. Piéron's law was originally formulated in the context of signal identifica-

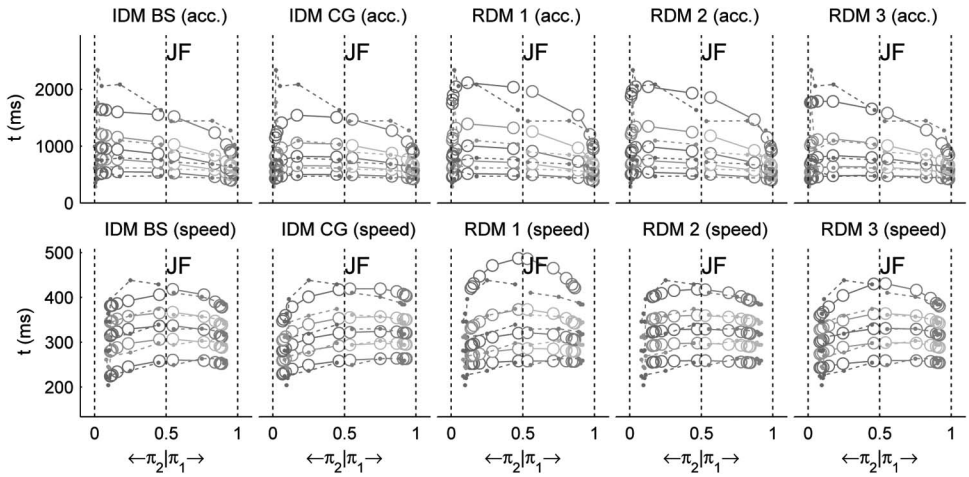


Figure 19. Quantile probability plots for data and fits of participant JF from Ratcliff and Rouder (1998). The top row displays the results for the accuracy (acc.) condition, the bottom row for the speed condition. In each row, from left to right are shown IDM BS, IDM CG, RDM 1, RDM 2, and RDM 3. Each response time (RT) distribution (two per stimulus, one for corrects and one for errors) is shown as a set of RT quantiles (at .1, .2, .5, .7, and .9), plotted vertically above the total probability value associated to that distribution. Data quantiles are indicated with full dots, and model fits are indicated with open circles. Plots for the other participants from Ratcliff and Rouder and participant MM from Mulder et al. (2013) can be found in Appendix F. IDM = Ising Decision Maker; BS = boundary separation; CG = coarse graining; RDM = Ratcliff diffusion model.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2
Model Selection Results

Participant	IDM BS	IDM CG	RDM 1	RDM 2	RDM 3
NH	32,364 (3)	32,348 (2)	32,783 (5)	32,421 (4)	32,202 (1)
KR	33,477 (2)	33,501 (3)	33,970 (5)	33,540 (4)	33,464 (1)
JF	33,847 (3)	33,657 (2)	34,696 (5)	34,072 (4)	33,623 (1)
MM	—	16,525 (2)	17,789 (4)	16,911 (3)	16,510 (1)

Note. Deviance information criterion values are presented for all five models and all four participants. The selection ranks of the models per participant are shown in parentheses. The dash indicates that there is no result. IDM = Ising Decision Maker; BS = boundary separation; CG = coarse graining; RDM = Ratcliff diffusion model.

tion and states that RT decreases as stimulus intensity increases (e.g., brightness or loudness). The observed decrease in mean RT is assumed to follow a negative power law (Piéron, 1913). As indicated by van Maanen, Grasman, Forstmann, and Wagenmakers (2012), two different interpretations of Piéron's law for speeded two-choice RT exist. The first interpretation considers a task in which the varied stimulus intensity is not the relevant dimension of the two-choice experiment (i.e., a nonselective feature in the terminology of this article), and the actual choice is made based on some other, more complex feature of the same stimulus. As an example, consider a face/artifact discrimination task with pictures varying in contrast (such that contrast is the irrelevant dimension). The second version considers a task in which the stimulus intensity is the actual discriminating factor of the two-choice experiment. In this article, we stick to the first interpretation of Piéron's law, which is in line with the work of Pins and Bonnet (1996). In addition to the suggested decrease in RT, recent research by Jaśkowski and Włodarczyk (2006) suggests that there is an accompanying decrease in accuracy as well.

The van der Molen-Keuss effect refers to what happens for very large stimulus intensities: Very intense stimuli lead to increasing RTs, and in addition, error responses can become faster than correct responses.

One may consider Piéron's law for speeded two-choice RT and the van der Molen-Keuss effect as two sides of the same coin. Both effects describe a change in RT for a so-called changing stimulus intensity. While Piéron's law describes the behavior at lower stimulus intensities (but still above detection threshold), the van der Molen-Keuss effect occurs at very high stimulus intensities. Interestingly, none of the existing theoretical models for decision making have been linked to these two effects. It is clear that some of the simpler models (e.g., Ratcliff diffusion model) are not able to explain these empirical findings without additional ad hoc assumptions.

It is important to stress that in the experiments we are referring to, stimulus intensity does not pertain to the actual evidence for the choice but to the strength of other, nonselective features of the stimulus. For instance, when categorizing a white contour line on a black background as a face or an artifact, the brightness of the line (and therefore the contrast of the picture) does not directly influence its form, which is the selective or evidence part of the stimulus. Of course, we assume that all experiments are carried out well above detection threshold because, when brightness falls to zero, no form can be distinguished at all.

In this section, we study how the IDM may explain both Piéron's law and the van der Molen-Keuss effect. Figures 20a and 20b, respectively, show the accuracy and RT results for varying nonselective input strength B_{ns} and a high value of stimulus distinctness ($C = 0.4$). For the other parameter values of the IDM, we refer the reader to Table 1. Experimental data for this high accuracy regime were collected by Jaśkowski and Włodarczyk

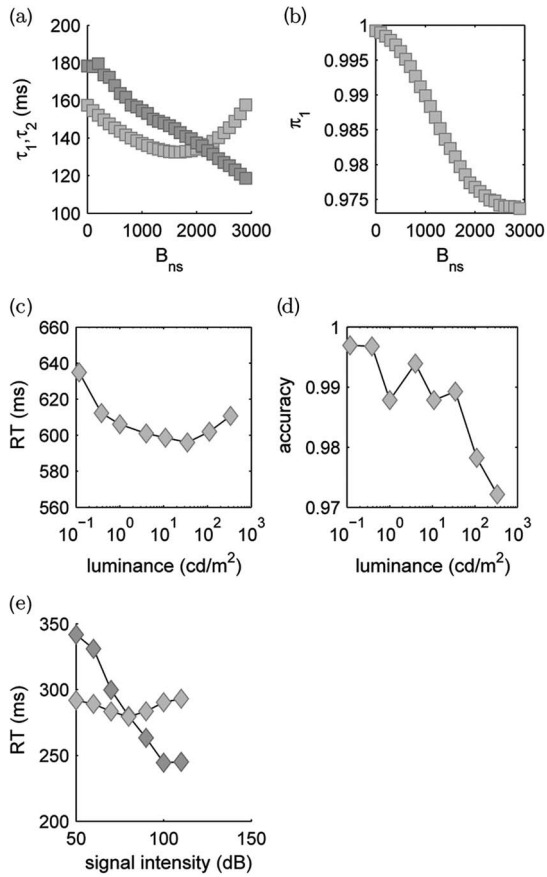


Figure 20. Piéron's law and the van der Molen-Keuss effect. Panel a shows Ising Decision Maker response time (RT) results for varying nonselective input strength B_{ns}: τ_1 (lighter gray) represents the correct responses, τ_2 (darker gray) represents the incorrect responses. The parameters used are taken from Table 1 with $B_s = 2{,}500$ and a high value of stimulus distinctness, $C = 0.4$. Panel b shows the corresponding accuracy results. Panels c and d show the experimental RTs (mainly correct responses) obtained by Jaśkowski and Włodarczyk (2006) for a visual cue Simon task for varying brightness. Panel e shows the experimental RTs for both correct (lighter gray squares) and error responses (darker gray squares) obtained by van der Molen and Orlebeke (1980) for a pitch discrimination task for varying loudness. The data shown here are based on the graphs in Jaśkowski and Włodarczyk and in van der Molen and Orlebeke. Panels a, c, and e all show a U-shaped relation between RT and stimulus intensity for correct responses. As can be seen in Panels a and e, error RT decreases monotonically with increasing stimulus intensity. Error RTs were not reported separately by Jaśkowski and Włodarczyk and are therefore not shown in Panel c. Panels b and d both show a decreasing accuracy for increasing stimulus intensity. In both experiments, stimulus intensity refers to an aspect of the stimulus (i.e., loudness and brightness, respectively) that is not immediately connected to the task's complexity but is essential in carrying the problem to the decision maker.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

(2006), and the results are shown in Figures 20c and 20d. One of the first accounts of the van der Molen-Keuss effect was by van der Molen and Orlebeke (1980) and is shown in Figure 20e. Although an exact relation between physical stimulus intensity and nonselective IDM input B_{ns} is not specified, both quantities are expected to result in comparable RT behavior (in terms of increase and decrease) under the plausible assumption of a monotonously increasing relation between the two. When B_{ns} is small but increasing, a drop in RT is indeed discovered, similar to Piéron's law. At the same time, this is accompanied by a decrease in accuracy, in accordance with the results of Jaśkowski and Włodarczyk. For very large values of B_{ns}, RT rises again. More specifically, the correct response RT rises, while the error response RT stays low. This is in agreement with the findings of van der Molen and Orlebeke.

The increase of RTs at large values of B_{ns}, especially for the correct responses, is due to the influence of a startled state for very high input strengths. Such a startled state corresponds to a high activity in both decision pools (and is therefore indecisive). This is illustrated in Figure 13, in which the time evolution of decision trajectories subject to the influence of a startled state merging with the correct decision state is presented. Once a large nonselective input is provided, the process is drawn to the startled state before it evolves generally to the correct decision state.

Two-Stimulus Comparison: Weber's Law

In all previous sections of the article, we have considered the standard version of a speeded choice task: A single stimulus is presented, and a fast decision has to be made. In the IDM, a stimulus supplies a certain amount of selective strength to each of the response alternatives (apart from a nonselective component B_{ns}). This selective part of the two-component external input field was conveniently parameterized through a distinctness variable C and a supposedly constant selective input strength B_s (see Equation 16). The fact that the IDM actually has two inputs allows us to model another type of speeded choice task in which a person is asked to compare a pair of stimuli on a certain scale. Questions of this type are "Which of two grays is lighter?" or "Which of two lines is longer?" In such a situation, we may relate the IDM to Weber's law, as has been done before by Deco, Scarano, and Soto-Faraco (2007) for an integrate-and-fire network.

Although Weber's law relates the physical dimension to the sensation, it is important to notice that in the case of the IDM, input strengths B_{s1} and B_{s2} do not coincide with the actual physical quantities. They are presented to the IDM only after being processed by sensory neurons and possibly a series of other predecision processes, most likely in a nonlinear fashion (Copelli, Roque, Oliveira, & Kinouchi, 2002; Yang & Wu, 1997). In what follows, we do not consider these sensory and preprocessing stages and focus solely on what happens after the input strengths have been computed. To ensure the IDM's decision about which of the two stimuli is greater corresponds to the original physical ordering, there should at least be a monotonic relation between the actual physical stimulus size and IDM input strength.

A key observable for experiments in which stimuli are compared is the just noticeable difference or *jnd*. The *jnd* refers to the minimal stimulus difference required to successfully discriminate between two stimuli (e.g., Luce, Bush, & Eugene, 1963). A *jnd* can be defined on the input scale of the IDM. Consider the input strengths of the selective (task-relevant) part of the external field B_{1s} and B_{2s}, with average input strength $B_s = \frac{B_{1s}+B_{2s}}{2}$ and difference $\Delta B_s = B_{1s} - B_{2s} = 2CB_s$. The $jnd_\alpha(B_s)$ for a certain average input strength B_s is equal to the minimal half-difference $\frac{\Delta B_s}{2}$ that surpasses a predefined accuracy level α. The $jnd_\alpha(B_s)$ is therefore also defined by the desired accuracy, giving a precise meaning to the term *noticeably different*. Based on the alternative formulation of the selective part of the external field given in Equation 16, the *jnd* can also be defined through the notion of just noticeable distinctness $C_\alpha(B_s)$:

$$jnd_\alpha(B_s) = \frac{\Delta_\alpha(B_s)}{2} = B_s C_\alpha(B_s). \tag{34}$$

Weber's law states that the *jnd* of some observed physical quantity is proportional to that quantity (e.g., Luce et al., 1963). In terms of Equation 34, Weber's law states that the ratio of the $jnd_\alpha(B_s)$ to the stimulus strength B_s, the so-called Weber fraction, is a constant. In the current framework, the Weber fraction is identical to $C_\alpha(B_s)$.

Although some reservation toward its generality is in place (Masin, 2009), it is supported by many experiments in the midranges of a great number of basic perceptive stimuli such as auditory pitch, luminance, and so on (see, e.g., Woodworth & Schlosberg, 1954). It is also found in more advanced tasks such as, for example, line length and number estimation (Cantlon, Platt, & Brannon, 2009; see also Smith & Ratcliff, 2009). Weber's law could emerge at different stages of processing: Automatic data compression at early sensory stages has been suggested as a possible explanation (Copelli et al., 2002), but also the discrimination/decision process itself has been put forward (Deco et al., 2007). In light of this last possibility, we check if it is possible to have Weber-like behavior in the IDM.

In Figure 21, we study the Weber fractions $C_\alpha(B_s)$ generated by the IDM as a function of B_s. For the chosen parameter values of the IDM, it can be seen that Weber's law approximately holds for a range of values of B_s situated in the competition regime (see Figure 6a, where the dotted line coincides with the B_s range shown in Figure 21).[10]

Conclusion

In this article, we have developed a prototypical neural calculator, the Ising Decision Maker or IDM, as a model for the speeded two-choice RT task. The underlying Ising model consists of two pools of binary stochastic neurons with pairwise interactions. Inside each pool, neurons excite each other; between pools, neurons inhibit each other. The perceptual input, finally, is represented by a two-component external field. Each component excites the

[10] At first glance, this appears to contradict the claim made by Deco et al. (2007) that a constant Weber fraction can only be found in the region of three-state stability of the integrate-and-fire network proposed by Wong and Wang (2006). There are many possible reasons for this discrepancy: the dependence of the phenomenon on the specific parameter values of both models, the introduction of a nonselective part to the external field of the IDM, or a more intrinsic difference between the IDM and the integrate-and-fire model.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

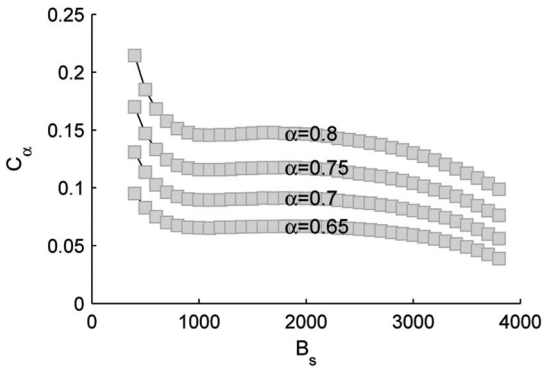


Figure 21. Weber fractions $C_\alpha(B_s)$ versus selective input strength B_s for different values of desired accuracy (α = 0.65, 0.70, 0.75, 0.80). The parameter values are taken from Table 1, with B_{ns} = 1,000. For these particular parameter values, Weber's law seems to hold (a constant Weber fraction) throughout the competition regime (approximately stretching from B_s = 1,000 to B_s = 3,000). The fractions were calculated by simulating the full range of stimulus distinctness C for every B_s. This allowed us to pinpoint the just noticeable distinctness $C_\alpha(B_s)$ (which is identical to the Weber fraction) leading to a performance of, respectively, 65%, 70%, 75%, and 80%.

neurons of one pool and in doing so provides evidence for the related choice. From this microscopic description, a discrete two-dimensional stochastic model is derived, describing the time evolution of the mean neural activity per pool. A decision is triggered when the system enters the stable state related to a high neural activity of the corresponding pool. A formal connection was established with traditional diffusion models by calculating the continuous limit of the model's discrete Markov chain. The IDM diffusion limit formally reduces to an LCA model with additional nonlinear terms, directly connected to the system's entropy.

The IDM produces realistic, right-skewed RT distributions. Under the assumption of constant total input, stimulus distinctness can be used to parameterize the evidence embedded in the stimulus. Assuming evidence in favor of the first alternative, higher stimulus distinctness leads to increased accuracy and generally faster RTs. For varying distinctness, the mean correct response RTs are approximately linearly proportional to the corresponding standard deviation, as reported by Wagenmakers and Brown (2007). The probability latency functions generated by the IDM have the typically observed inverted U-shape (e.g., Ratcliff & Rouder, 1998). Also in line with experiments, the left–right symmetry of the probability latency functions is slightly broken: Error responses occur later than their correct counterparts for stimuli of intermediate to high distinctness. For stimuli with even higher distinctness, error responses can become faster than their correct counterparts.

On the level of effective neural correlations, the IDM is very similar to the integrate-and-fire network proposed by Wong and Wang (2006). Although the IDM is a lot less intricate, the essential properties of the aforementioned integrate-and-fire network are reproduced: a spontaneous state, decision states (which can be reminiscent), and slow errors. On the microscopic level, the Ising model provides only a minimal description of the decision-making part of the brain, based on elementary neural correlations rather than the complicated biophysiology giving rise to these correlations (integrate-and-fire networks with dedicated excitatory and inhibitory neurons). Rather than an attempt at an exact biophysiological replica, the IDM is an abstract mathematical description of how the brain implements the decision process on a low level of information processing. The neurons in the IDM should be considered as functional units, not necessarily corresponding to individual physical neurons of any specific type, that process information by means of a dynamical updating rule with an equally abstract interpretation. Consequently, there is no problem with the difference in time scale between a single neurophysiological spike (physiological properties impose severe constraints on the time scale of neuronal firing) and a single Metropolis neuron update in the IDM (which happens instantaneously and has no temporal constraints otherwise). Both the spike and the neuron update serve as microscopic carriers of correlations across their respective decision networks, and there is no reason they should have anything more in common. As a result, both modeling approaches come with their own advantages. The advantage of integrate-and-fire networks is their explicit connection to the detailed mechanics of the brain's neurons. In particular, the time scale on which mean activity evolves can be successfully inferred from realistic firing rates. Because they adopt the interplay between excitatory and inhibitory neurons, integrate-and-fire networks can also accommodate oscillatory behavior. The advantage of the IDM approach, however, is its simplicity (only the most essential parts of a network are retained) and analytical elegance. In the IDM's case, the time scale on which mean activity evolves is simply a parameter of the model. Applied to human decision making, both models actually greatly simplify the biophysiological details of the brain. With the rise of the neurophysiologically inspired integrate-and-fire networks, Ising models have become less popular as a modeling tool in brain science and decision making in particular. However, it is our conviction that they are still of great value to the field as they offer a unique combination of the simplicity typical of the abstract diffusion tradition and the nonlinearity typical of multiple attractor networks. The results in this article testify to this fruitful middle ground.

In terms of neural plausibility, the IDM distinguishes itself from traditional linear diffusion models by demonstrating global multiple attractor network behavior. Regardless of the plausibility of the underlying neural rationale (Smith, 2010; Smith & McKenzie, 2011), traditional linear diffusion models are limited to a local description of evidence accumulation once the stimulus is presented. Because of its nonlinearity, the IDM describes the system during the entire trial: before, during, and after stimulus presentation. At the same time, this description includes neurally plausible concepts like the spontaneous state and decision states.

Viewing decision making as an abstract stochastic minimization algorithm of a potential energy function inspires an alternative way of parameterizing SAT than the traditional changing boundary (box size) assumption, namely, coarse graining. Increasing the collective step size of the Markov chain dynamics that stochastically minimizes the IDM's free energy function during a trial results in faster but less accurate decisions. At the same time, this quick and dirty or coarse-grained dynamics diminishes the lag of

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

error responses, eventually introducing error responses that are faster than their correct counterparts. This seems to be in agreement with what is observed experimentally. As a matter of convenience, the dynamics describing coarse graining was directly defined on the macroscopic level of mean activities. However, it would be interesting to investigate which microscopic algorithms effectively reproduce the SAT effects obtained with this macroscopic approach. A possible source of inspiration for this is more exotic Ising model dynamics like the Wolff algorithm (see, e.g., Newman & Barkema, 1999), that allow for the simultaneous updating of larger sets of neurons to increase equilibrium sampling efficiency. Another microscopically defined method of coarse graining that could be considered is the parallel updating algorithm of Torres et al. (2008). An additional feature of parallel updating lies in the introduction of oscillating (and even chaotic) behavior that may address oscillating aspects of brain functioning during decision making that cannot be modeled within the current IDM set-up.

Additionally, three major speeded two-choice experiment related effects described in psychological literature can be identified through IDM simulations: Piéron's law, the van der Molen-Keuss effect, and an account of Weber's law at the decision-making level. Breaking up the network input into a stimulus selective term (containing evidence concerning the choice) and a common nonselective term (containing decision-irrelevant information) provides a workable model for the stimuli presented in these experiments. Evidently, this input parameterization is not restricted to the IDM: It can be applied to any diffusion model with dual input (drift rate). For low to moderate nonselective input strengths, mean IDM RTs drop with increasing nonselective input strength. This behavior is referred to as the speeded two-choice version of Piéron's law and has been experimentally verified by, for instance, Pins and Bonnet (1996). For extremely high nonselective input strengths, correct RTs increase again, while error RTs keep decreasing monotonically, eventually leading to error RTs shorter than their correct counterparts. This behavior is called the van der Molen-Keuss effect (van der Molen & Keuss, 1979). Finally, a decision-level version of Weber's law was discovered stretching over most of the competition region. It should be noted that various other ways of input parameterization have been proposed and investigated, sometimes incorporating time-dependent evidence (see, e.g., Teodorescu & Usher, 2013; Tsetsos, Usher, & McClelland, 2011). In the context of the IDM, time dependency of the input could offer a unique perspective on the nature of decision states, more specifically regarding the ability to attract decision trajectories after the evidence has been removed or even reversed.

Two important issues in the study of decision processes were not directly addressed in the main body of this article and are discussed here: bias and confidence. First, let us address the issue of bias (see, e.g., Diederich & Busemeyer, 2006; Leite, 2011; Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012; Ratcliff, 1985; Ratcliff & Starns, 2009). In the derivation of the IDM, starting from Equation 9, we have assumed that the number of neurons N_p, the excitatory interaction strength W_p^+, and the internal threshold Θ_p are equal for both pools. Bias can be introduced to the network by allowing any of these parameters to become pool specific. Abandoning the parametric symmetry between pools results in an asymmetric free energy surface both before and after stimulus presentation. From the traditional Ratcliff diffusion perspective, this implies an at least slightly biased starting position, regardless of which of the three parameters was used. Introducing bias specifically through the internal thresholds (i.e., Θ) impacts the linear part of the IDM and is therefore closely related to assuming an a priori bias in drift rate in the Ratcliff diffusion model. A choice regarding the description of bias is required before considering more advanced problems such as optimal decision making (Bogacz et al., 2006). Second, there is the issue of decision confidence. To address this, one could consider using postdecision activity, as proposed by Pleskac and Busemeyer (2010).

There are many ways the IDM can be further explored and extended. One way is to generalize the two-pool model discussed in this article to a *k*-pool version (*k* being a strictly positive natural number) capable of dealing with speeded *k*-choice RT data. One can then verify if the model is able to reproduce well-established experimental phenomena like Hick's law (Hick, 1952; McMillen & Holmes, 2006). (Note that with the possibility of this extension in mind, the proofs in Appendixes B and C are already given for the more general case of *k* pools.) Another interesting issue that can be addressed is the presence of correlations between closely consecutive trials. Within the IDM framework, reminiscent states could have a significant impact on the next trial. This may prove useful in determining which of the two is more likely: decision making with or without reminiscent states (Band II and Band III, respectively; see the section entitled Modes of Operation of the IDM). Also, as a neural-network-based model, the IDM presents itself as an interesting theoretical framework for studying learning effects in speeded choice tasks. Concepts such as Hebbian learning (Hebb, 1949) can be integrated into the model in a natural way (a stochastic Hopfield network equipped with Hebbian learning is called a Boltzmann machine; see, e.g., Ackley, Hinton, & Sejnowski, 1985).[11] As a Hopfield derivate, the IDM provides a link between decision making on the one hand and the Hopfield network's original field of application, memory, on the other. From this perspective, the IDM could also prove useful in memory-related research. Finally, as the IDM is derived from a theory that is closely related to the biophysics of neural decision making and has clearly defined macroscopic neural quantities, it can be considered a good candidate for simultaneously describing behavioral data and a number of aspects of macroscopic neural data (e.g., event-related potentials, as in Philiastides, Ratcliff, & Sajda, 2006) measured during speeded choice decision tasks.

Fitting experimental data to the IDM is shown to be a realistic endeavor. More so for the data under consideration, the statistical model-selection procedure showed that the IDM can compete with the traditional Ratcliff diffusion model. This may of course not be the case for other data sets, and therefore, more systematic comparisons are needed. The introduction of selective (evidence-

[11] Using Hebbian learning in the context of a decision model has also been proposed by Anderson, Silverstein, Ritz, and Jones (1977) in the context of their brain-state-in-a-box (BSB) model. The BSB model itself has some affinity with the models discussed in this article. It is a system of linear differential equations, but nonstochastic. Noise only enters the process at the level of the input. It is therefore neither a stochastic network nor a diffusion model.

related) and nonselective (circumstantial) input provides a modeling framework for an all-in-one speeded two-choice experiment in which, for one and the same discrimination task, many different aspects are studied simultaneously: probability latency functions for different selective and nonselective input strengths, Piéron's law, the van der Molen-Keuss effect, and Weber's law at the decision level. Such experiments could also explore experimentally thus far uncharted terrain. For instance, no experimental data are available on the van der Molen-Keuss effect on stimuli resulting in an accuracy closer to chance level. The IDM's predictions for this effect challenge experimental psychologists to investigate this hidden area of speeded choice RTs. More generally, using the input model described above, the IDM and all other dual input diffusion models are theoretically equipped to model an overarching speeded two-choice experiment. Their different ability to fit all observed phenomena into one set of parameter values, as well as cope with more extreme experimental conditions (e.g., the van der Molen-Keuss effect), will dominate any model-selection criterion and hopefully provide a more fundamental classification of the currently competing speeded two-choice RT models.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

References

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science, 9,* 147–169.

Amit, D. J., & Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex, 7,* 237–252. doi:10.1093/cercor/7.3.237

Amit, D. J., Gutfreund, H., & Sompolinsky, H. (1985a). Spin-glass models of neural networks. *Physical Review A, 32,* 1007–1018. doi:10.1103/PhysRevA.32.1007

Amit, D. J., Gutfreund, H., & Sompolinsky, H. (1985b). Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters, 55,* 1530–1533. doi:10.1103/PhysRevLett.55.1530

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review, 84,* 413–451. doi:10.1037/0033-295X.84.5.413

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113,* 700–765. doi:10.1037/0033-295X.113.4.700

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153–178. doi:10.1016/j.cogpsych.2007.12.002

Brown, S. D., Ratcliff, R., & Smith, P. L. (2006). Evaluating methods for approximating stochastic differential equations. *Journal of Mathematical Psychology, 50,* 402–410. doi:10.1016/j.jmp.2006.03.004

Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences, 23,* 255–282. doi:10.1016/0165-4896(92)90043-5

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459. doi:10.1037/0033-295X.100.3.432

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences, 13,* 83–91. doi:10.1016/j.tics.2008.11.007

Cocco, S., Leibler, S., & Monasson, R. (2009). Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. *PNAS: Proceedings of the National Academy of Sciences, USA, 106,* 14058–14062. doi:10.1073/pnas.0906705106

Copelli, M., Roque, A. C., Oliveira, R. F., & Kinouchi, O. (2002). Physics of psychophysics: Stevens and Weber-Fechner laws are transfer functions of excitable media. *Physical Review E, 65,* Article 060901. doi:10.1103/PhysRevE.65.060901

Deco, G., Rolls, E. T., & Romo, R. (2009). Stochastic dynamics as a principle of brain function. *Progress in Neurobiology, 88,* 1–16. doi:10.1016/j.pneurobio.2009.01.006

Deco, G., Scarano, L., & Soto-Faraco, S. (2007). Weber's law in decision making: Integrating behavioral data in humans with a neurophysiological model. *Journal of Neuroscience, 27,* 11192–11200. doi:10.1523/JNEUROSCI.1072-07.2007

Diederich, A., & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound-change, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics, 68,* 194–207. doi:10.3758/BF03193669

Ditterich, J. (2010). A comparison between mechanisms of multi-alternative perceptual decision making: Ability to explain human behavior, predictions for neurophysiology, and relationship with decision theory. *Frontiers in Neuroscience, 4,* Article 184. doi:10.3389/fnins.2010.00184

Eckhoff, P., Wong-Lin, K., & Holmes, P. (2011). Dimension reduction and dynamics of a spiking neural network model for decision making under neuromodulation. *SIAM Journal on Applied Dynamical Systems, 10,* 148–188. doi:10.1137/090770096

Forstmann, B. U., Wagenmakers, E.-J., Eichele, T., Brown, S., & Serences, J. T. (2011). Reciprocal relations between cognitive neuroscience and formal cognitive models: Opposites attract? *Trends in Cognitive Sciences, 15,* 272–279. doi:10.1016/j.tics.2011.04.002

Furman, M., & Wang, X.-J. (2008). Similarity effect and optimal control of multiple-choice decision making. *Neuron, 60,* 1153–1168.

Gardiner, C. W. (2004). *Handbook of stochastic methods for physics, chemistry, and the natural sciences* (3rd ed.). New York, NY: Springer.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2003). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30,* 535–574. doi:10.1146/annurev.neuro.29.051605.113038

Griffiths, D. J. (1999). *Introduction to electrodynamics*. Upper Saddle River, NJ: Prentice Hall.

Haario, H., Laine, M., Mira, A., & Saksman, E. (2006). DRAM: Efficient adaptive MCMC. *Statistics and Computing, 16,* 339–354.

Heathcote, A., & Brown, S. (2004). Reply to Speckman and Rouder: A theoretical basis for QML. *Psychonomic Bulletin & Review, 11,* 577–578. doi:10.3758/BF03196614

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9,* 394–401. doi:10.3758/BF03196299

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York, NY: Wiley.

Heitz, R. P., & Schall, J. D. (2012). Neural mechanisms of speed-accuracy tradeoff. *Neuron, 76,* 616–628. doi:10.1016/j.neuron.2012.08.030

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4,* 11–26.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *PNAS: Proceedings of the National Academy of Sciences, USA, 79,* 2554–2558.

Ising, E. (1925). Beitrag zur theorie des ferromagnetismus [Contribution to the theory of ferromagnetism]. *Zeitschrift für Physik, 31,* 253–258. doi:10.1007/BF02980577

Jaśkowski, P., & Włodarczyk, D. (2006). Task modulation of the effects of brightness on reaction time and response force. *International Journal of Psychophysiology, 61,* 98–112. doi:10.1016/j.ijpsycho.2005.07.010

Karlin, S., & Taylor, H. M. (1981). *A second course in stochastic processes*. New York, NY: Academic Press.

Kikuchi, K., Yoshida, M., Maekawa, T., & Watanabe, H. (1991). Metropolis Monte Carlo method as a numerical technique to solve the Fokker–Planck equation. *Chemical Physics Letters, 185,* 335–338. doi:10.1016/S0009-2614(91)85070-D

Kloeden, P. E., & Platen, E. (2011). *Numerical solution of stochastic differential equations*. New York, NY: Springer.

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science, 6,* 299–312.

Laming, D. (1968). *Information theory of choice-reaction times*. Oxford, England: Academic Press.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge, England: Cambridge University Press.

Leite, F. P. (2011). What cognitive processes drive response biases? A diffusion model analysis. *Judgment and Decision Making, 6,* 651–687.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40,* 77–105. doi:10.1007/BF02291481

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.

Luce, R. D., Bush, R. R., & Eugene, G. (Eds.). (1963). *Handbook of mathematical psychology*. Oxford, England: Wiley.

Masin, S. C. (2009). The Weber's law that never was. In *Proceedings of the Twenty-Fifth Annual Meeting of the International Society for Psychophysics* (pp. 441–446). Galway, Ireland: International Society for Psychophysics.

McMillen, T., & Holmes, P. (2006). The dynamics of choice among multiple alternatives. *Journal of Mathematical Psychology, 50,* 30–57. doi:10.1016/j.jmp.2005.10.003

Mulder, M. J., Keuken, M. C., van Maanen, L., Boekel, W., Forstmann, B. U., & Wagenmakers, E.-J. (2013). The speed and accuracy of perceptual decisions in a random-tone pitch task. *Attention, Perception, & Psychophysics, 75,* 1048–1058. doi:10.3758/s13414-013-0447-8

Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience, 32,* 2335–2343. doi:10.1523/JNEUROSCI.4156-11.2012

Newman, M. E. J., & Barkema, G. T. (1999). *Monte Carlo methods in statistical physics*. Oxford, England: Clarendon Press.

Niwa, M., & Ditterich, J. (2008). Perceptual decisions between multiple directions of visual motion. *Journal of Neuroscience, 28,* 4435–4445. doi:10.1523/JNEUROSCI.5564-07.2008

Osman, A., Lou, L., Muller-Gethmann, H., Rinkenauer, G., Mattes, S., & Ulrich, R. (2000). Mechanisms of speed–accuracy tradeoff: Evidence from covert motor processes. *Biological Psychology, 51,* 173–199. doi:10.1016/S0301-0511(99)00045-9

Philiastides, M. G., Ratcliff, R., & Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: A timing diagram. *Journal of Neuroscience, 26,* 8965–8975. doi:10.1523/JNEUROSCI.1655-06.2006

Piéron, H. (1913). Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices [Research on the variation of sensory latency as a function of stimulus intensity]. *L'année psychologique, 20,* 17–96. doi:10.3406/psy.1913.4294

Pins, D., & Bonnet, C. (1996). On the relation between stimulus intensity and processing time: Piéron's law and choice reaction time. *Perception & Psychophysics, 58,* 390–400. doi:10.3758/BF03206815

Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences, 6,* 421–425. doi:10.1016/S1364-6613(02)01964-2

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review, 117,* 864–901. doi:10.1037/a0019737

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108. doi:10.1037/0033-295X.85.2.59

Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review, 92,* 212–225. doi:10.1037/0033-295X.92.2.212

Ratcliff, R., Cherian, A., & Segraves, M. (2003). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. *Journal of Neurophysiology, 90,* 1392–1407. doi:10.1152/jn.01049.2002

Ratcliff, R., & Dongen, H. P. A. V. (2011). Diffusion model for one-choice reaction-time tasks and the cognitive effects of sleep deprivation. *PNAS: Proceedings of the National Academy of Sciences, USA, 108,* 11285–11290. doi:10.1073/pnas.1100483108

Ratcliff, R., Hasegawa, Y. T., Hasegawa, R. P., Smith, P. L., & Segraves, M. A. (2007). Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *Journal of Neurophysiology, 97,* 1756–1774. doi:10.1152/jn.00393.2006

Ratcliff, R., & McKoon, G. (2007). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922. doi:10.1162/neco.2008.12-06-420

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9,* 347–356. doi:10.1111/1467-9280.00067

Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 127–140. doi:10.1037/0096-1523.26.1.127

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111,* 333–367. doi:10.1037/0033-295X.111.2.333

Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review, 116,* 59–83. doi:10.1037/a0014086

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review, 106,* 261–300. doi:10.1037/0033-295X.106.2.261

Rorie, A. E., Gao, J., McClelland, J. L., & Newsome, W. T. (2010). Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. *PLoS ONE, 5*(2), Article e9308. doi:10.1371/journal.pone.0009308

Rouder, J. N. (1996). Premature sampling in random walks. *Journal of Mathematical Psychology, 40,* 287–296. doi:10.1006/jmps.1996.0030

Roxin, A., & Ledberg, A. (2008). Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Computational Biology, 4*(3), Article e1000046. doi:10.1371/journal.pcbi.1000046

Schroeder, D. V. (1999). *An introduction to thermal physics*. San Francisco, CA: Addison Wesley.

Smith, P. L. (2010). From Poisson shot noise to the integrated Ornstein–Uhlenbeck process: Neurally principled models of information accumulation in decision-making and response time. *Journal of Mathematical Psychology, 54,* 266–283. doi:10.1016/j.jmp.2009.12.002

Smith, P. L., & McKenzie, C. R. L. (2011). Diffusive information accumulation by minimal recurrent neural models of decision making. *Neural Computation, 23,* 2000–2031. doi:10.1162/NECO_a_00150

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences, 27,* 161–168. doi:10.1016/j.tins.2004.01.006

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review, 116,* 283–317. doi:10.1037/a0015156

Speckman, P. L., & Rouder, J. N. (2004). A comment on Heathcote, Brown, and Mewhort's QMLE method for response time distributions. *Psychonomic Bulletin & Review, 11,* 574–576. doi:10.3758/BF03196613

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B. Statistical Methodology, 64,* 583–639. doi:10.1111/1467-9868.00353

Stone, M. (1960). Models for choice-reaction time. *Psychometrika, 25,* 251–260. doi:10.1007/BF02289729

Storn, R., & Price, K. (1997). Differential evolution: A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization, 11,* 341–359.

Teodorescu, A. R., & Usher, M. (2013). Disentangling decision models: From independence to competition. *Psychological Review, 120,* 1–38. doi:10.1037/a0030776

Tkacik, G., Schneidman, E., Berry, M. J., II, & Bialek, W. (2009). *Spin glass models for a network of real neurons.* arXiv:0912.5409

Torres, J., Marro, J., Cortes, J., & Wemmenhove, B. (2008). Instabilities in attractor networks with fast synaptic fluctuations and partial updating of the neurons activity. *Neural Networks, 21,* 1272–1277. doi:10.1016/j.neunet.2008.07.002

Tsetsos, K., Usher, M., & McClelland, J. L. (2011). Testing multi-alternative decision models with non-stationary evidence. *Frontiers in Neuroscience, 5,* Article 63. doi:10.3389/fnins.2011.00063

Tuerlinckx, F. (2004). The efficient computation of the cumulative distribution and probability density functions in the diffusion model. *Behavior Research Methods, 36,* 702–716.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592. doi:10.1037/0033-295X.108.3.550

van der Molen, M. W., & Keuss, P. J. G. (1979). The relationship between reaction time and intensity in discrete auditory tasks. *Quarterly Journal of Experimental Psychology, 31,* 95–102. doi:10.1080/14640747908400709

van der Molen, M. W., & Orlebeke, J. F. (1980). Phasic heart rate change and the U-shaped relationship between choice reaction time and auditory signal intensity. *Psychophysiology, 17,* 471–481. doi:10.1111/j.1469-8986.1980.tb00186.x

van Maanen, L., Grasman, R. P. P. P., Forstmann, B. U., & Wagenmakers, E.-J. (2012). Piéron's law and optimal behavior in perceptual decision-making. *Frontiers in Neuroscience, 5,* Article 143. doi:10.3389/fnins.2011.00143

van Ravenzwaaij, D., van der Maas, H. L. J., & Wagenmakers, E.-J. (2012). Optimal decision making in neural inhibition models. *Psychological Review, 119,* 201–215. doi:10.1037/a0026275

Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics, 13,* 37–58. doi:10.1080/00140137008931117

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology, 52,* 1–9. doi:10.1016/j.jmp.2007.09.005

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology, 21,* 641–671. doi:10.1080/09541440802205067

Wagenmakers, E.-J., & Brown, S. (2007). On the linear relation between the mean and the standard deviation of a response time distribution. *Psychological Review, 114,* 830–841. doi:10.1037/0033-295X.114.3.830

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron, 36,* 955–968. doi:10.1016/S0896-6273(02)01092-9

Wong, K.-F., & Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience, 26,* 1314–1328. doi:10.1523/JNEUROSCI.3733-05.2006

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology.* New York, NY: Holt.

Yang, X.-L., & Wu, S. M. (1997). Response sensitivity and voltage gain of the rod- and cone-bipolar cell synapses in dark-adapted tiger salamander retina. *Journal of Neurophysiology, 78,* 2662–2673.

Yellott, J. I. (1971). Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology, 8,* 159–199. doi:10.1016/0022-2496(71)90011-3

(*Appendices follow*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A

Derivation of Equation 7

To appreciate Equation 7, we note that $\sum_{i<j} S_{1i}S_{1j}$ from Equation 5 counts the number of jointly active neurons in pool 1 (i.e., both neurons i and j have an activity of 1). Take as an example a pool with $N = 8$ nodes that have the following configuration: (0, 1, 1, 0, 1, 1, 0, 0) (hence, $y_1 = 0.5$). We may construct an 8×8 cross table, with the configuration in the margins. The cells of the table are the multiplication of the corresponding marginal row and column element as is done in Table A1. In the calculation of $\sum_{i<j} S_{1i}S_{1j}$, we only have to consider the lower triangular part of the matrix (because we want to count each distinct pair (i, j) only once). From Table A1, it can be seen that the number of ones in the lower triangular part is directly related to the number of ones in the configuration (i.e., $Ny_1 = 8 \cdot 0.5 = 4$). It is easier to start with considering all elements of the full table, and then, the number of cells with ones is $(Ny_1)^2 = 16$. From this, we have to subtract the number of ones on the diagonal, that is $Ny_1 = 4$. Next, the result has to be halved such that only the lower triangular part is considered (the matrix is symmetric), and this gives the final result: $\sum_{i<j} S_{1i}S_{1j} = \frac{1}{2}((Ny_1)^2 - Ny_1)$. The calculation of the other sums in Equation 5 is comparable.

Table A1
Table With the Example Configuration (0, 1, 1, 0, 1, 1, 0, 0) in Rows and Columns and the Elementwise Multiplication in the Cells

	S_{1j}							
S_{1i}	0	1	1	0	1	1	0	0
0								
1	$1 \cdot 0$							
1	$1 \cdot 0$	$1 \cdot 1$						
0	$0 \cdot 0$	$0 \cdot 1$	$0 \cdot 1$					
1	$1 \cdot 0$	$1 \cdot 1$	$1 \cdot 1$	$1 \cdot 0$				
1	$1 \cdot 0$	$1 \cdot 1$	$1 \cdot 1$	$1 \cdot 0$	$1 \cdot 1$			
0	$0 \cdot 0$	$0 \cdot 1$	$0 \cdot 1$	$0 \cdot 0$	$0 \cdot 1$	$0 \cdot 1$		
0	$0 \cdot 0$	$0 \cdot 1$	$0 \cdot 1$	$0 \cdot 0$	$0 \cdot 1$	$0 \cdot 1$	$0 \cdot 0$	

Appendix B

Diffusion Equations for Single Neuron Updating (Equation 18)

In what follows, we derive the Ising Decision Maker (IDM) diffusion equations associated with single neuron updating Metropolis dynamics. We immediately consider a more general proof, involving k pools instead of the currently required two.

Consider k pools $p = 1, \ldots, k$ of each N_p binary neuron S_{pi} with $i = 1, \ldots, N_p$. Let $E(\boldsymbol{S})$ be the system's energy function, with $\boldsymbol{S}$ the vector containing the values of all $N = \sum_p N_p$ neurons of the system. Assume the energy function is invariant under the switching of neuron activities within a pool. This means the mean activity vector $\boldsymbol{y} = (y_1, \ldots, y_p, \ldots, y_k)$ with $y_p = \frac{\sum_i S_{pi}}{N_p}$ (as in Equation 6) is sufficient information to know the system's energy. As a consequence, we can make use of a k-dimensional scalar function $E(\boldsymbol{y}) = E(\boldsymbol{S})$, with $y_p = \frac{\sum_i S_{pi}}{N_p}$. We allow $E(\boldsymbol{y})$ to have the same name as $E(\boldsymbol{S})$, as the difference will always be clear through the notation of their arguments.

Now consider a single neuron updating algorithm with Metropolis transition probability. Let us first discuss the proposal mechanism of this algorithm. The proposal mechanism randomly selects a neuron of the system and proposes a change in its activity (0 to 1 or 1 to 0). Evidently, the probability of this neuron belonging to pool p is $\frac{N_p}{N}$. The activity change of a neuron in pool p can be seen as a change in y_p of N_p^{-1} if one chooses a neuron that has activity 0 or $-N_p^{-1}$ if one chooses a neuron that has activity 1. When selecting from pool p, choosing a neuron with activity 0 has probability $1 - y_p$; choosing a neuron with activity 1 has probability y_p.

This leads to a discrete Markov chain in the k-dimensional space of mean activity $\boldsymbol{y}$. Each time step, an N_p^{-1}-sized step (the net result of changing one neuron's activity) is proposed in one of the two directions of y_p, with p the pool of the neuron that was (randomly) selected. The proposed change in mean activity $\boldsymbol{y}$ is denoted by a stochastic vector $\boldsymbol{s}$. It has $2k$ possible realizations (two possibilities for each pool), which are indexed with a superscript: $\boldsymbol{s}^1 = (N_1^{-1}, 0, 0, \ldots)$, $-\boldsymbol{s}^1 = (-N_1^{-1}, 0, 0, \ldots)$, $\boldsymbol{s}^2 = (0, N_2^{-1}, 0, \ldots)$, $-\boldsymbol{s}^2 = (0, -N_2^{-1}, 0, \ldots), \ldots$, with proposal probabilities $\frac{N_1}{N}(1 - y_1)$, $\frac{N_1}{N}y_1$, $\frac{N_2}{N}(1 - y_2)$, $\frac{N_2}{N}y_2, \ldots$, respectively. Seeing the chain is implicitly indexed with an integer time n, we define real time as $t = n\,\Delta t$.

After an activity step has been proposed, its acceptance needs to be decided. Given an energy surface $E(\boldsymbol{y})$, a particular proposed activity step $\boldsymbol{s}^p$ is accompanied by a change in energy $\Delta E = E(\boldsymbol{y} + \boldsymbol{s}^p) - E(\boldsymbol{y})$. In accordance with the microscopic Metropolis acceptance rule (see Equation 4), a move is made to the candidate if $\Delta E \leq 0$ or with a probability $e^{-\beta \Delta E}$ if $\Delta E > 0$.

(*Appendices continue*)

The Fokker–Planck equation for the time evolving probability density function $p(\boldsymbol{y}, t)$ (indicating the density of $\boldsymbol{y}$ at time t) for a k-dimensional diffusion can be written as

$$\frac{\partial p(\boldsymbol{y}, t)}{\partial t} = -\sum_{p=1}^{k} \frac{\partial}{\partial y_p} A_p(\boldsymbol{y}, t) p(\boldsymbol{y}, t) + \frac{1}{2} \sum_{p=1}^{k} \sum_{q=1}^{k} \frac{\partial^2}{\partial y_p \partial y_q} G_{pq}(\boldsymbol{y}, t) p(\boldsymbol{y}, t) \quad \text{(B1)}$$

(see, e.g., Gardiner, 2004), where $A_p(\boldsymbol{y}, t)$ is the pth component of the drift rate vector and G_{pq} is the (p, q)th element of the diffusion matrix G.

For a small time interval Δt (see, e.g., Karlin & Taylor, 1981), $A_p(\boldsymbol{y}, t) \approx \frac{\langle s_p \rangle}{\Delta t}$ (i.e., the drift rate is the expected change in $\boldsymbol{y}$ in a very small time Δt), and $G_{pq}(\boldsymbol{y}, t) \approx \frac{\langle s_p s_q \rangle}{\Delta t}$ (i.e., the diffusion matrix is the expected squared change in $\boldsymbol{y}$ in a very small time Δt). Please note that the expectation values $\langle s_p \rangle$ and $\langle s_p s_q \rangle$ refer to the actual change in activity (i.e., after acceptance or rejection of the proposed activity step) and not to the expectation value of the proposal distribution. Before calculating the drift vector and diffusion matrix emerging from the proposed Metropolis Markov chain, we first compute two useful intermediate results.

First, we approximate the change in energy ΔE that accompanies a proposed activity step. In the first order of N_p^{-1}, a change of $\boldsymbol{s}^p$ in $\boldsymbol{y}$ results in a change in energy of

$$\begin{aligned} \Delta E &= E(\boldsymbol{y} + \boldsymbol{s}^p) - E(\boldsymbol{y}) \\ &\approx E(\boldsymbol{y}) + N_p^{-1} \frac{\partial E}{\partial y_p} - E(\boldsymbol{y}) \\ &= N_p^{-1} \frac{\partial E}{\partial y_p}. \end{aligned} \quad \text{(B2)}$$

Evidently, a change of $-\boldsymbol{s}^p$ in $\boldsymbol{y}$ brings about a change in energy of $\Delta E = -N_p^{-1} \frac{\partial E}{\partial y_p}$.

Second, we can additively decompose the expression for the IDM's entropy $\Sigma(\boldsymbol{y})$ of Equation 13 into $\Sigma_I(\boldsymbol{y})$ and $\Sigma_{II}(\boldsymbol{y})$ as follows:

$$\begin{aligned} \Sigma(\boldsymbol{y}) &= -\sum_p N_p (y_p \log(y_p) + (1 - y_p) \log(1 - y_p)) \\ &= \underbrace{-\sum_p N_p y_p \log(y_p)}_{\Sigma_I(\boldsymbol{y})} \underbrace{-\sum_p N_p (1 - y_p) \log(1 - y_p)}_{\Sigma_{II}(\boldsymbol{y})}. \end{aligned}$$

Taking first derivatives of $\Sigma_I(\boldsymbol{y})$ and $\Sigma_{II}(\boldsymbol{y})$ and multiplying with $-N_p^{-1}$ gives

$$-N_p^{-1} \frac{\partial \Sigma_I}{\partial y_p} = 1 + \log(y_p) \quad \text{(B3)}$$

and

$$-N_p^{-1} \frac{\partial \Sigma_{II}}{\partial y_p} = -1 - \log(1 - y_p). \quad \text{(B4)}$$

We now have the necessary ingredients to compute the expectation values $\langle s_p \rangle$ and $\langle s_p s_q \rangle$. First, consider the expected change in activity $\langle s_p \rangle$. Assuming $\frac{\partial E(\boldsymbol{y})}{\partial y_p} \geq 0$, we get (approximating up to the first order of N_p^{-1} and using the results of Equations B2, B3, B4, and 12)

$$\begin{aligned} \langle s_p \rangle &= \frac{N_p}{N} \left[-y_p N_p^{-1} + (1 - y_p) e^{-\beta \Delta E} N_p^{-1} \right] \\ &= \frac{1}{N} \left[-e^{\log(y_p)} + e^{-\beta \Delta E + \log(1 - y_p)} \right] \\ &= \frac{1}{Ne} \left[-e^{\log(y_p) + 1} + e^{-\beta \Delta E + \log(1 - y_p) + 1} \right] \\ &\approx \frac{1}{Ne} \left[-e^{-N_p^{-1} \frac{\partial \Sigma_I}{\partial y_p}} + e^{-\beta N_p^{-1} \frac{\partial E}{\partial y_p} + N_p^{-1} \frac{\partial \Sigma_{II}}{\partial y_p}} \right] \\ &\approx \frac{1}{Ne} \left[-1 + N_p^{-1} \frac{\partial \Sigma_I}{\partial y_p} + 1 - \beta N_p^{-1} \frac{\partial E}{\partial y_p} + N_p^{-1} \frac{\partial \Sigma_{II}}{\partial y_p} \right] \\ &\approx \frac{1}{Ne} \left[-\beta N_p^{-1} \frac{\partial (E - \beta^{-1} (\Sigma_I + \Sigma_{II}))}{\partial y_p} \right] \\ &= -\frac{\beta}{N N_p e} \frac{\partial F}{\partial y_p}. \end{aligned}$$

The same result is obtained for $\frac{\partial E(\boldsymbol{y})}{\partial y_p} < 0$. This means

$$A_p(\boldsymbol{y}, t) = \frac{\langle s_p \rangle}{\Delta t} = -\frac{\beta}{N N_p e \Delta t} \frac{\partial F}{\partial y_p}.$$

For the expected squared change in activity $\langle s_p^2 \rangle$,

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

we consider a similar approximation. Assuming $\frac{\partial E(\mathbf{y})}{\partial y_p} \geq 0$, we get (again with an approximation until the first order of N_p^{-1} and using the results of Equations B2, B3, and B4)

$$\begin{aligned}
\langle s_p^2 \rangle &= \frac{N_p}{N}\left[y_p N_p^{-2} + (1-y_p)e^{-\beta\Delta E} N_p^{-2}\right] \\
&= \frac{1}{N}\left[e^{\log(y_p)} + e^{-\beta\Delta E + \log(1-y_p)}\right] N_p^{-1} \\
&= \frac{1}{Ne}\left[e^{\log(y_p)+1} + e^{-\beta\Delta E + \log(1-y_p)+1}\right] N_p^{-1} \\
&= \frac{1}{Ne}\left[e^{N_p^{-1}\frac{\partial \Sigma_I}{\partial y_p}} + e^{-\beta N_p^{-1}\frac{\partial E}{\partial y_p} - N_p^{-1}\frac{\partial \Sigma_{II}}{\partial y_p}}\right] N_p^{-1} \\
&\approx \frac{1}{Ne}\left[1 + N_p^{-1}\frac{\partial \Sigma_I}{\partial y_p} + 1 - \beta N_p^{-1}\frac{\partial E}{\partial y_p} - N_p^{-1}\frac{\partial \Sigma_{II}}{\partial y_p}\right] N_p^{-1} \\
&\approx \frac{1}{Ne}\left[1+1\right] N_p^{-1} \\
&= \frac{2}{NN_p e}.
\end{aligned}$$

The same result is obtained for $\frac{\partial E(\mathbf{y})}{\partial y_p} < 0$.

Evidently, for the off-diagonal elements $G_{pq}(\mathbf{y}, t)$ (with $p \neq q$), it holds that $\langle s_p s_q \rangle = 0$ (because a change in one dimension means no change in any other). As a consequence, we find

$$G_{pp} = \frac{2}{NN_p e \Delta t}$$

and

$$G_{pq} = 0,$$

with $p \neq q$.

Let us define a common diffusion constant D_p as follows:

$$D_p = \frac{G_{pp}}{2} = \frac{1}{NN_p e \Delta t}.$$

We can now simplify the drift rate to

$$A_p(\mathbf{y}) = -\beta D_p \frac{\partial F(\mathbf{y})}{\partial y_p},$$

resulting in

$$dy_p = -\beta D_p \frac{\partial F(\mathbf{y})}{\partial y_p} dt + \sqrt{2D_p} dW_p,$$

with all W_p independent Wiener processes.

If all pools have equal size $N_p = \frac{N}{k}$, then D_p reduces to

$$D_p = D = \frac{k}{N^2 e \Delta t},$$

and the Fokker–Planck equation for $p(\mathbf{y}, t)$ is equivalent to the following stochastic differential equation:

$$d\mathbf{y} = -\beta D \frac{\partial F(\mathbf{y})}{\partial y_p} dt + \sqrt{2D} d\boldsymbol{W},$$

with $\boldsymbol{W}$ a vector of k independent Wiener processes.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix C

Diffusion Equations for Collective Neuron Updating (Equation 19)

In what follows, we derive the diffusion equations that are associated with macroscopically defined dynamics for the Ising Decision Maker (IDM), which mimics collective neuron updating on a microscopic level. We immediately consider a more general proof, involving k pools instead of the currently required two. For the proposal distribution of the activity steps, we take an isotropic k-dimensional normal distribution $\mathcal{N}_k(\boldsymbol{s}; \mathbf{0}, \sigma^2 I_k) = \prod_{p=1}^{k} \mathcal{N}(s_p; 0, \sigma^2)$, with I_k the k-dimensional identity matrix.

Consider a discrete Markov chain in a k-dimensional space. Each time step Δt, an activity step $\boldsymbol{s}$ is drawn from $\mathcal{N}_k(\mathbf{0}, \sigma^2 I_k)$ and accepted or rejected according to the Metropolis algorithm. Seeing the chain is implicitly indexed with an integer time n, we define real time as $t = n\Delta t$.

Given a free energy surface $F(\boldsymbol{y})$, a proposed activity step $\boldsymbol{s}$ is accompanied by a change in free energy $\Delta F = F(\boldsymbol{y} + \boldsymbol{s}) - F(\boldsymbol{y})$. Applying a Metropolis acceptance rule (cf. Equation 4, but now on the level of the free energy), a move is made to the candidate if $\Delta F \leq 0$ or with a probability $e^{-\beta \Delta F}$ if $\Delta F > 0$.

The Fokker–Planck equation for the time evolving probability density function $p(\boldsymbol{y}, t)$ (indicating the density of $\boldsymbol{y}$ at time t) for a k-dimensional diffusion can be written as in Equation B1 with $A_p(\boldsymbol{y}, t)$ again the pth component of the drift rate vector and G_{pq} the (p, q)th element of the diffusion matrix G.

For a small time interval Δt (see, e.g., Karlin & Taylor, 1981), $A_p(\boldsymbol{y}, t) \approx \frac{\langle s_p \rangle}{\Delta t}$ (i.e., the drift rate is the expected change in $\boldsymbol{y}$ in a very small time Δt), and $G_{pq}(\boldsymbol{y}, t) \approx \frac{\langle s_p s_q \rangle}{\Delta t}$ (i.e., the diffusion matrix is the expected squared change in $\boldsymbol{y}$ in a very small time Δt). Please note that the expectation values $\langle s_p \rangle$ and $\langle s_p s_q \rangle$ refer to the actual change in activity (i.e., after acceptance or rejection of the proposed activity step) and not to the expectation value of the proposal distribution.

Before calculating the drift vector and diffusion matrix emerging from the proposed Metropolis Markov chain, we first compute a useful intermediate result. The change in free energy ΔF that accompanies a proposed activity step $\boldsymbol{s}$ can be linearly approximated as

$$\begin{aligned} \Delta F &= F(\boldsymbol{y} + \boldsymbol{s}) - F(\boldsymbol{y}) \\ &\approx F(\boldsymbol{y}) + \boldsymbol{s} \cdot \nabla F(\boldsymbol{y}) - F(\boldsymbol{y}) \\ &= \boldsymbol{s} \cdot \nabla F(\boldsymbol{y}). \end{aligned} \tag{C1}$$

Under this approximation, the condition $\Delta F > 0$ simplifies to $\boldsymbol{s} \cdot \nabla F(\boldsymbol{y}) > 0$.

Averaging the acceptance rule over all possible for $\boldsymbol{s}$, we get

$$\langle \boldsymbol{s} \rangle = \int_{\boldsymbol{s}\cdot\nabla F(\boldsymbol{y})>0} \mathcal{N}_k(\boldsymbol{s}; \mathbf{0}, \sigma^2 I_k) e^{-\beta \Delta F} \boldsymbol{s} d\boldsymbol{s} + \int_{\boldsymbol{s}\cdot\nabla F(\boldsymbol{y})\leq 0} \mathcal{N}_k(\boldsymbol{s}; \mathbf{0}, \sigma^2 I_k) \boldsymbol{s} d\boldsymbol{s}.$$

In what follows, we denote condition $\boldsymbol{s}\cdot\nabla F(\boldsymbol{y}) > 0$ with index $+$ to the integral and $\boldsymbol{s}\cdot\nabla F(\boldsymbol{y}) \leq 0$ with index $-$. Applying the linear approximation in Equation C1, we get

$$\langle \boldsymbol{s} \rangle \approx \int_{+} \mathcal{N}_k(\boldsymbol{s}; \mathbf{0}, \sigma^2 I_k) e^{-\beta \boldsymbol{s}\cdot\nabla F(\boldsymbol{y})} \boldsymbol{s} d\boldsymbol{s} + \int_{-} \mathcal{N}_k(\boldsymbol{s}; \mathbf{0}, \sigma^2 I_k) \boldsymbol{s} d\boldsymbol{s}. \tag{C2}$$

Applying a coordinate transformation $\boldsymbol{u} = \frac{\boldsymbol{s}}{\sigma}$, we get

$$\langle \boldsymbol{s} \rangle \approx \sigma \left[\int_{+} e^{-\beta\sigma \boldsymbol{u}\cdot\nabla F(\boldsymbol{y})} \mathcal{N}_k(\boldsymbol{u}; \mathbf{0}, I_k) \boldsymbol{u} d\boldsymbol{u} + \int_{-} \mathcal{N}_k(\boldsymbol{u}; \mathbf{0}, I_k) \boldsymbol{u} d\boldsymbol{u} \right].$$

We now expand the exponential to the first nonzero order in σ:

$$\begin{aligned} \langle \boldsymbol{s} \rangle &\approx \sigma \Big[\int_{+} \mathcal{N}_k(\boldsymbol{u}; \mathbf{0}, I_k)[1 - \beta\sigma \boldsymbol{u} \cdot \nabla F(\boldsymbol{y})] \boldsymbol{u} d\boldsymbol{u} \\ &\quad + \int_{-} \mathcal{N}_k(\boldsymbol{u}; \mathbf{0}, I_k) \boldsymbol{u} d\boldsymbol{u} \Big] \\ &= \sigma \Big[-\beta\sigma \int_{+} \mathcal{N}_k(\boldsymbol{u}; \mathbf{0}, I_k) \boldsymbol{u} \cdot \nabla F(\boldsymbol{y}) \boldsymbol{u} d\boldsymbol{u} \\ &\quad + \int_{+\cup -} \mathcal{N}_k(\boldsymbol{u}; \mathbf{0}, I_k) \boldsymbol{u} d\boldsymbol{u} \Big] \\ &= -\beta\sigma^2 \int_{+} \mathcal{N}_k(\boldsymbol{u}; \mathbf{0}, I_k) \boldsymbol{u} \cdot \nabla F(\boldsymbol{y}) \boldsymbol{u} d\boldsymbol{u} \\ &= -\beta\sigma^2 \Big[\int_{+} \mathcal{N}_k(\boldsymbol{u}; \mathbf{0}, I_k) r(\boldsymbol{u}) \boldsymbol{u} d\boldsymbol{u} \Big] \|\nabla F(\boldsymbol{y})\|, \end{aligned} \tag{C3}$$

with $r(\boldsymbol{u}) = \boldsymbol{u} \cdot \frac{\nabla F(\boldsymbol{y})}{\|\nabla F(\boldsymbol{y})\|}$ the length of the projection of $\boldsymbol{u}$ on the unit vector along $\nabla F(\boldsymbol{y})$. The component perpendicular to the unit vector $\frac{\nabla F(\boldsymbol{y})}{\|\nabla F(\boldsymbol{y})\|}$ is canceled out when integrating $r(\boldsymbol{u})\boldsymbol{u}$. Therefore, the integral vector is oriented along unit vector $\frac{\nabla F(\boldsymbol{y})}{\|\nabla F(\boldsymbol{y})\|}$. The length of the integrands projection on the unit vector is $r(\boldsymbol{u})\boldsymbol{u} \cdot \frac{\nabla F(\boldsymbol{y})}{\|\nabla F(\boldsymbol{y})\|} = r(\boldsymbol{u})^2$. This allows us to rewrite the integral as

$$\int_{+} \mathcal{N}_k(\boldsymbol{u}; \mathbf{0}, I_k) r(\boldsymbol{u}) \boldsymbol{u} d\boldsymbol{u} = \Big[\int_0^{\infty} \mathcal{N}(r; 0, 1) r^2 dr \Big] \frac{\nabla F(\boldsymbol{y})}{\|\nabla F(\boldsymbol{y})\|}, \tag{C4}$$

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

where factor $\mathcal{N}(r; 0, 1)$ equals the density of vectors $\boldsymbol{u}$ with projection r. Combining Equations C3 and C4 results in

$$\langle \boldsymbol{s} \rangle = -\frac{\beta\sigma^2}{2} \nabla F(\boldsymbol{y}).$$

Connecting back to Equation B1, we get

$$A_p(\boldsymbol{y}, t) = -\frac{\beta\sigma^2}{2\Delta t} \frac{\partial F(\boldsymbol{y})}{\partial y_p}.$$

For the expected squared displacement, we consider a similar approximation. Up to the first nonzero order of σ (the third-order term is dropped after the fourth line), we get

$$\begin{aligned}
\langle s_p^2 \rangle &\approx \int_+ \mathcal{N}_k(\boldsymbol{s}; \boldsymbol{0}, \sigma^2 I_k) e^{-\beta \boldsymbol{s} \cdot \nabla F(\boldsymbol{y})} s_p^2 d\boldsymbol{s} \\
&\quad + \int_- \mathcal{N}_k(\boldsymbol{s}; \boldsymbol{0}, \sigma^2 I_k) s_p^2 d\boldsymbol{s} \\
&= \sigma^2 \Big[\int_+ \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) e^{-\beta\sigma \boldsymbol{u} \cdot \nabla F(\boldsymbol{y})} u_p^2 d\boldsymbol{u} \\
&\quad + \int_- \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) u_p^2 d\boldsymbol{u} \Big] \\
&\approx \sigma^2 \Big[\int_+ \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) \big[1 - \beta\sigma \boldsymbol{u} \cdot \nabla F(\boldsymbol{y}) \big] u_p^2 d\boldsymbol{u} \\
&\quad + \int_- \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) u_p^2 d\boldsymbol{u} \Big] \\
&= \sigma^2 \Big[-\beta\sigma \int_+ \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) \boldsymbol{u} \cdot \nabla F(\boldsymbol{y}) u_p^2 d\boldsymbol{u} \\
&\quad + \int \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) u_p^2 d\boldsymbol{u} \Big] \\
&\approx \sigma^2 \Big[\int \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) u_p^2 d\boldsymbol{u} \Big] \\
&= \sigma^2 \Big[\int_{-\infty}^{\infty} \mathcal{N}(r; 0, \sigma) r^2 dr \Big] \\
&= \sigma^2.
\end{aligned}$$

For the off-diagonal elements ($p \neq q$), it holds that

$$\begin{aligned}
\langle s_p s_q \rangle_{p \neq q} &\approx \int_+ \mathcal{N}_k(\boldsymbol{s}; \boldsymbol{0}, \sigma^2 I_k) e^{-\beta \boldsymbol{s} \cdot \nabla F(\boldsymbol{y})} s_p s_q d\boldsymbol{s} \\
&\quad + \int_- \mathcal{N}_k(\boldsymbol{s}; \boldsymbol{0}, \sigma^2 I_k) s_p s_q d\boldsymbol{s} \\
&= \sigma^2 \Big[\int_+ \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) e^{-\beta\sigma \boldsymbol{u} \cdot \nabla F(\boldsymbol{y})} u_p u_q d\boldsymbol{u} \\
&\quad + \int_- \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) u_p u_q d\boldsymbol{u} \Big] \\
&\approx \sigma^2 \Big[\int_+ \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) \big[1 - \beta\sigma \boldsymbol{u} \cdot \nabla F(\boldsymbol{y}) \big] u_p u_q d\boldsymbol{u} \\
&\quad + \int_- \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) u_p u_q d\boldsymbol{u} \Big] \\
&= \sigma^2 \Big[-\beta\sigma \int_+ \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) \boldsymbol{u} \cdot \nabla F(\boldsymbol{y}) u_p u_q d\boldsymbol{u} \\
&\quad + \int \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) u_p u_q d\boldsymbol{u} \Big] \\
&\approx \sigma^2 \Big[\int \mathcal{N}_k(\boldsymbol{u}; \boldsymbol{0}, I_k) u_p u_q d\boldsymbol{u} \Big] \\
&= 0.
\end{aligned}$$

This means that

$$G_{pp} = \frac{\sigma^2}{\Delta t}$$

and

$$G_{pq} = 0,$$

with $p \neq q$.

Let us define a common diffusion constant D as follows:

$$D = \frac{G_{pp}}{2} = \frac{\sigma^2}{2\Delta t}.$$

Now, we can simplify the drift rate to

$$A_p(\boldsymbol{y}, t) = -\beta D \frac{\partial F(\boldsymbol{y})}{\partial y_p}.$$

Because the Fokker–Planck equation for $p(\boldsymbol{y}, t)$ is equivalent to a stochastic differential equation, we may also write for y_p,

$$dy_p = -\beta D \frac{\partial F(\boldsymbol{y})}{\partial y_p} dt + \sqrt{2D} dW_p,$$

with all W_p independent Wiener processes, and for $\boldsymbol{y}$,

$$d\boldsymbol{y} = -\beta D \frac{\partial F(\boldsymbol{y})}{\partial y_p} dt + \sqrt{2D} d\boldsymbol{W},$$

with $\boldsymbol{W}$ a vector of k independent Wiener processes.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix D

Tradeoffs Between Parameters in the IDM

The Ising Decision Maker (IDM) parameter identification issues can best be studied after a change of variables. We use a polynomial expansion of the free energy F around the midway mean activities (0.5, 0.5). This is an acceptable approximation because the decision trajectories tend to spend most of their time in this region. Define $\boldsymbol{x} = \boldsymbol{y} - (0.5, 0.5)$, and approximate the free energy $F(x_1, x_2)$ by a Taylor expansion around (0, 0):

$$\begin{aligned} F(x_1, x_2) &= -\sum_{p=1,2} W^+ x_p^2 + W^- x_1 x_2 \\ &\quad - \sum_{p=1,2} \left(B_p - \Theta + W^+ - \frac{W^-}{2} \right) x_p \\ &\quad + \frac{\beta^{-1} N}{2} \sum_{p=1,2} \left[2x_p^2 + \frac{4}{3} x_p^4 + \frac{32}{15} x_p^6 + \ldots \right] \\ &= \sum_{p=1,2} (\beta^{-1} N - W^+) x_p^2 + W^- x_1 x_2 \\ &\quad - \sum_{p=1,2} \left(B_p - \Theta + W^+ - \frac{W^-}{2} \right) x_p \\ &\quad + \frac{\beta^{-1} N}{2} \sum_{p=1,2} \left[\frac{4}{3} x_p^4 + \frac{32}{15} x_p^6 + \ldots \right]. \end{aligned} \tag{D1}$$

As this is a free energy, constant terms are irrelevant and can be dropped (because only differences or derivatives drive the dynamics). Although there are no obvious exact tradeoffs, there are some nearly exact tradeoffs that are only broken by higher powers of x.

A first approximate parameter tradeoff occurs up to the fourth order in $\boldsymbol{x}$ of the free energy surface. Consider another free energy function, $\tilde{F}(x_1, x_2)$, which is a rescaled (miniaturized) version of $F(x_1, x_2)$ such that $\tilde{F}(x_1, x_2) = F(\lambda x_1, \lambda x_2)$ with $\lambda > 1$. Up to the fourth order, $\tilde{F}(x_1, x_2)$ can be approximated as

$$\begin{aligned} \tilde{F}(x_1, x_2) &\approx \sum_{p=1,2} (\beta^{-1} N - W^+)(\lambda x_p)^2 + W^-(\lambda x_1)(\lambda x_2) \\ &\quad - \sum_{p=1,2} \left(B_p - \Theta + W^+ - \frac{W^-}{2} \right)(\lambda x_p) + \frac{2\beta^{-1} N}{3} \sum_{p=1,2} (\lambda x_p)^4. \end{aligned}$$

Simultaneously rescaling detection box size h and collective step size parameter σ results in a dynamics that is equivalent to the original system, but on a smaller scale. Consequently, the resulting reaction time distributions are the same. Up to the fourth order, the same rescaling can be achieved by tuning the original system's parameters:

$$\begin{aligned} \tilde{F}(x_1, x_2) &\approx \lambda^2 \sum_{p=1,2} (\beta^{-1} N - W^+) x_p^2 + \lambda^2 W^- x_1 x_2 \\ &\quad - \lambda \sum_{p=1,2} \left(B_p - \Theta + W^+ - \frac{W^-}{2} \right) x_p + \lambda^4 \frac{2\beta^{-1} N}{3} \sum_{p=1,2} x_p^4 \\ &= \sum_{p=1,2} (\beta^{-1} \tilde{N} - \tilde{W}^+) x_p^2 + \tilde{W}^- x_1 x_2 \\ &\quad - \sum_{p=1,2} \left(\tilde{B}_p - \tilde{\Theta} + \tilde{W}^+ - \frac{\tilde{W}^-}{2} \right) x_p + \frac{2\beta^{-1} \tilde{N}}{3} \sum_{p=1,2} x_p^4, \end{aligned}$$

with

$$\begin{aligned} \tilde{N} &= \lambda^4 N, \\ \tilde{W}^- &= \lambda^2 W^-, \\ \tilde{W}^+ &= \beta^{-1} \tilde{N} - \lambda^2 [\beta^{-1} N - W^+], \\ \tilde{\Theta} &= -\lambda \left(0 - \Theta + W^+ - \frac{W^-}{2} \right) + \tilde{W}^+ - \frac{\tilde{W}^-}{2} + 0, \\ \tilde{B}_p &= \lambda \left(B_p - \Theta + W^+ - \frac{W^-}{2} \right) - \tilde{W}^+ + \frac{\tilde{W}^-}{2} + \tilde{\Theta} = \lambda B_p. \end{aligned}$$

These new parameters, together with the rescaled s and σ, form a parameter set that leads to approximately the same reaction time distributions as the original system. For every value of $\lambda > 1$, there is a parameter set producing the same reaction time distributions as the original set. This situation can be regarded as a one-dimensional parameter redundancy or tradeoff, albeit an intricate one, as all parameters are involved.

A second approximate parameter tradeoff holds up to the second order. Looking at the second-order terms of the free energy as given in Equation D1, there is a clear linear tradeoff between N and W^+. A change in W^+ also has an impact on the first-order coefficient of the free energy that can in turn be compensated by a change in Θ.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix E

Simulation Details

Because a closed form solution is unavailable, we rely on more intricate numerical methods to evaluate the Ising Decision Maker's (IDM's) choice response time (RT) distribution functions. A straightforward approach is to simulate a large number of decision trajectories and use the generated choices and RTs to construct both error and correct RT distributions. The IDM is supposed to start out from a stable spontaneous state. Hence, before each decision trajectory, we reset the system to its spontaneous state and let it relax for 1 s without exposure to a stimulus. If the system escapes the spontaneous state (crosses a decision boundary) before the end of this 1-s stimulus anticipation period, the trajectory is labeled as a premature decision. As the number of premature decisions typically observed in standard two-choice RT experiments is very few, we make sure that its proportion never exceeds 0.005 when fitting the IDM to real data. After simulating the IDM trajectories, the RTs pertaining to each choice are binned into an approximate RT distribution. For enough trajectories and sufficiently narrow bins, these RT distributions will converge to the actual IDM RT distributions. The bin width or time resolution is set to 1 ms, and the binning is limited to RTs under 3 s.

For vanishing collective step size (we take $\sigma < 0.001$), we can reduce the IDM to a local gradient system and profit from the increased simulation speed of the Euler-Maruyama gradient approximation method (see, e.g., Kloeden & Platen, 2011). The time step precision in the Euler-Maruyama method is set to 1 ms, coinciding with the previously mentioned time resolution. Probability distributions do not change much for higher precisions, which is in accordance with the findings of Brown, Ratcliff, and Smith (2006).

These approximate distribution functions can now be used for calculating the quantile likelihood of the parameter values used to simulate the trajectories, given a certain set of choice RT data. The quantile likelihood method requires these distributions to be integrated into coarser (data-based) bins, reducing the simulation error on the resulting likelihood function. Through extensive pilot testing, we have found that, depending on the data, 100,000 to 400,000 simulations per stimulus generate sufficiently stable evaluations of the likelihood for successful Markov chain Monte Carlo sampling. Because the simulations have to be repeated each time the likelihood function is evaluated in another parameter vector, efficient computing is pivotal. For this reason, we made use of a state-of-the-art accelerated video card (NVIDIA Tesla K20Xm). The GPUs on this kind of high-performance video card are specially designed for massively parallel computation. This makes them perfect for the calculation of thousands of independent decision trajectories.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix F

Model Fits

Model fits for the simulated data SD, participants NH and KR from Ratcliff and Rouder (1998) and participant MM from Mulder et al. (2013), are shown in Figures F1, F2, and F3, respectively.

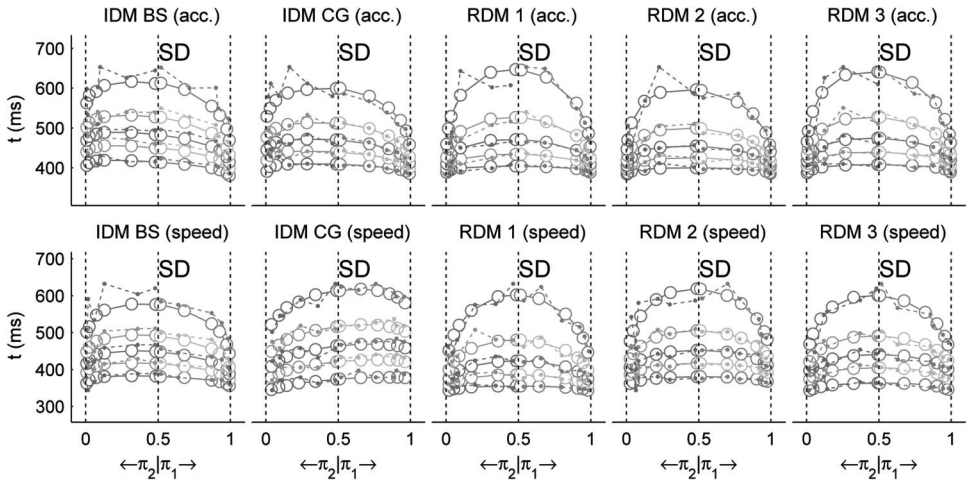
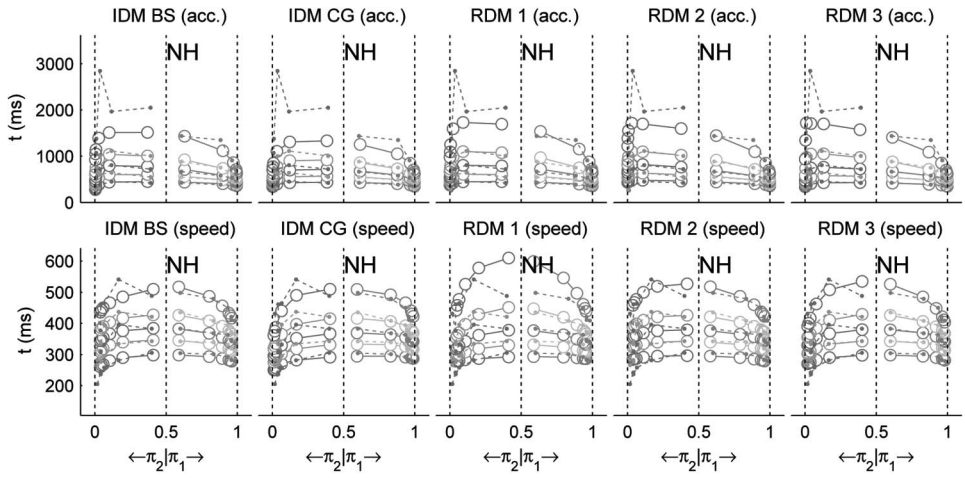


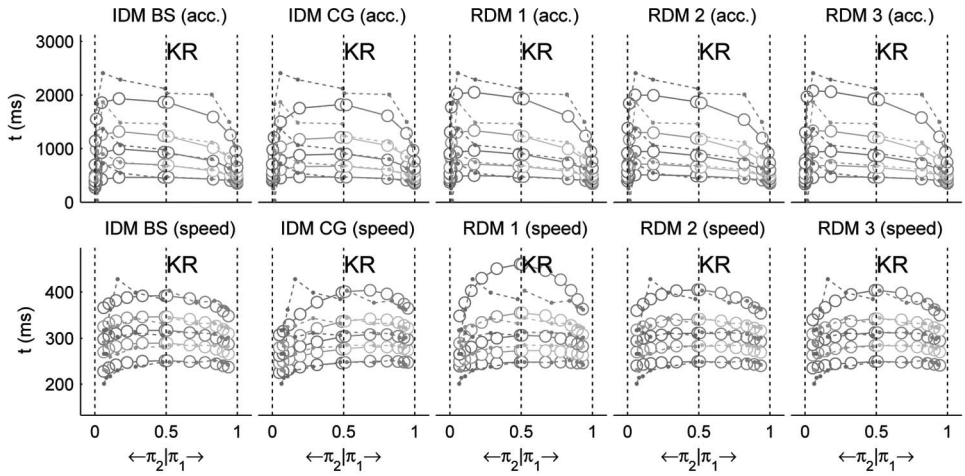
Figure F1. Quantile probability plots for simulated data and fits of, respectively, IDM BS, IDM CG, RDM 1, RDM 2, and RDM 3. The top row displays the results for the accuracy (acc.) condition, the bottom row for the speed condition. Each response time (RT) distribution (two per stimulus, one for corrects and one for errors) is shown as a set of RT quantiles (at .1, .2, .5, .7, and .9), plotted vertically above the total probability value associated to that distribution. Data quantiles are indicated with full dots, and model fits are indicated with open circles. IDM = Ising Decision Maker; BS = boundary separation; CG = coarse graining; RDM = Ratcliff diffusion model.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.



(a) NH



(b) KR

Figure F2. Quantile probability plots for data and fits of participants NH and KR from Ratcliff and Rouder (1998). Each panel represents a participant. For each panel, the top row displays the results for the accuracy (acc.) condition, the bottom row for the speed condition. In each row, from left to right are shown IDM BS, IDM CG, RDM 1, RDM 2, and RDM 3. Each response time (RT) distribution (two per stimulus, one for corrects and one for errors) is shown as a set of RT quantiles (at .1, .2, .5, .7, and .9), plotted vertically above the total probability value associated to that distribution. Data quantiles are indicated with full dots, and model fits are indicated with open circles. IDM = Ising Decision Maker; BS = boundary separation; CG = coarse graining; RDM = Ratcliff diffusion model.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

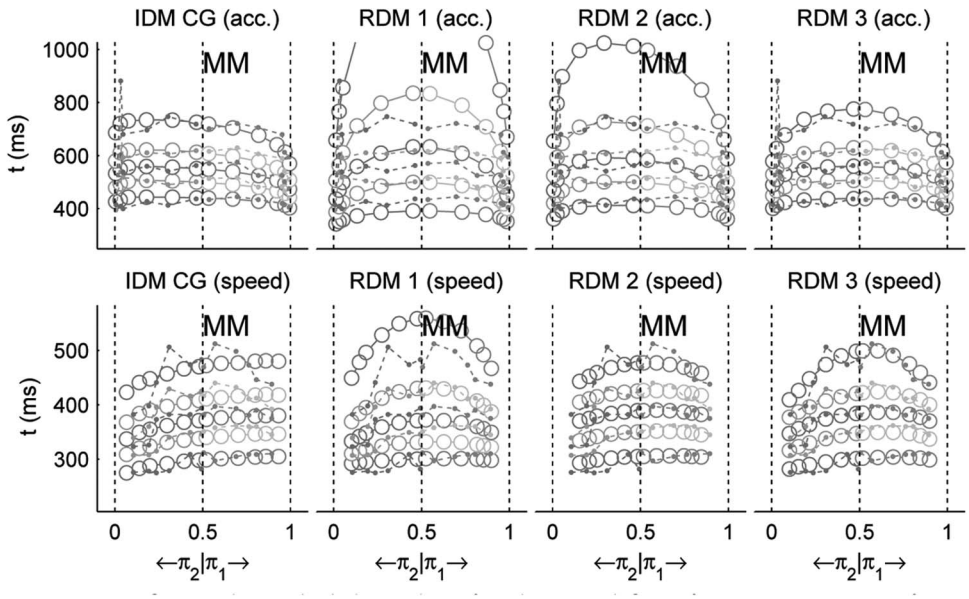


Figure F3. Quantile probability plots for data and fits of participant MM from Mulder et al. (2013). The top row displays the results for the accuracy (acc.) condition, the bottom row for the speed condition. In each row, from left to right are shown IDM CG, RDM 1, RDM 2, and RDM 3. Each response time (RT) distribution (two per stimulus, one for corrects and one for errors) is shown as a set of RT quantiles (at .1, .2, .5, .7, and .9), plotted vertically above the total probability value associated to that distribution. Data quantiles are indicated with full dots, and model fits are indicated with open circles. IDM = Ising Decision Maker; CG = coarse graining; RDM = Ratcliff diffusion model.

Received February 28, 2012
Revision received February 13, 2014
Accepted February 28, 2014 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Correction to Verdonck and Tuerlinckx (2014)

In the article "The Ising Decision Maker: A Binary Stochastic Network for Choice Response Time" by Stijn Verdonck and Francis Tuerlinckx (*Psychological Review*, 2014, Vol. 121, No. 3, pp. 422–462, http://dx.doi.org/10.1037/a0037012), an inaccurate assumption in Appendix B (see below) led to an oversimplified result in Equation 18 (the diffusion equations associated with the microscopically defined dynamics). The authors sincerely thank Rani Moran for making them aware of the problem. Only the expression of the diffusion coefficient D is incorrect, and should be changed into

$$D(\boldsymbol{y}) = \frac{\left[1 - sgn\left(\frac{\partial E(\boldsymbol{y})}{\partial y_p}\right)(1 - 2y_p)\right]}{N^2 \Delta t}.$$

It can be seen that the diffusion coefficient is no longer constant with respect to mean activity. However, the general form of Equation 18 remains unaltered. Moreover, at $y_p = 0.5$ (for all pools p), the diffusion coefficient of the microscopic dynamics equals the diffusion constant obtained with the macroscopic dynamics and a collective step size of $\sigma = \frac{\sqrt{2}}{N}$.

Because from that point onward, the alternative route of the macroscopically defined dynamics was taken, this altered result has no impact on the rest of the article or any of its conclusions.

Cause of the Problem

The derivation leading to Equation 18, given in Appendix B, is based on the assumption that the individual entropy terms $\Sigma_I(\boldsymbol{y})$ and $\Sigma_{II}(\boldsymbol{y})$ (see Equation B3 and above), as well as the energy $E(\boldsymbol{y})$ vary only slightly with the change in mean activity $\boldsymbol{y}$, caused by changing the activity of a single neuron.

However, a change in y_p of $-N_p^{-1}$ (caused by changing a neuron's activity from 1 to 0 in pool p) leads to a change in Σ_I of

$$-N_p^{-1}\frac{\partial \Sigma_I}{\partial y_p} = 1 + \log(y_p).$$

This means that the approximation that was used,

$$e^{-N_{1p}^{-1}\frac{\partial \Sigma_I}{\partial y_p}} \approx 1 - N_p^{-1}\frac{\partial \Sigma_I}{\partial y_p},$$

is only valid around $y_p = \frac{1}{e}$. For Σ_{II} the approximation that was used is only valid around $y_p = 1 - \frac{1}{e}$.

Solution

A better approach is to assume that the total free energy function $F(\boldsymbol{y})$ (Equation 14) is slowly varying, instead of the individual energy and entropy parts. A similar assumption is made in the case of the macroscopically defined dynamics.

The text in Appendix B, starting from the first paragraph after Equation B2 until the end, should be replaced with the following:

Second, Equation 13 implies that

$$N_p^{-1}\frac{\partial \Sigma}{\partial y_p} = \log\left(\frac{1 - y_p}{y_p}\right). \quad \text{(B3)}$$

We now have the necessary ingredients to compute the expectation values $\langle s_p \rangle$ and $\langle s_p s_q \rangle$. First, consider the expected change in activity $\langle s_p \rangle$. Assuming $\frac{\partial E(\boldsymbol{y})}{\partial y_p} \geq 0$, we get (approximating up to the first order of N_p^{-1} and using the results of Equations B2, B3, and 14):

$$\begin{aligned}\langle s_p \rangle &= \frac{N_p}{N}\left[-y_p N_p^{-1} + (1-y_p)e^{-\beta\Delta E} N_p^{-1}\right] = \frac{y_p}{N}\left[-1 + \frac{1-y_p}{y_p}e^{-\beta\Delta E}\right] \\ &= \frac{y_p}{N}\left[-1 + e^{-\beta\Delta E + N_p^{-1}\frac{\partial\Sigma}{\partial y_p}}\right] = \frac{y_p}{N}\left[-1 + e^{-\beta N_p^{-1}\frac{\partial E}{\partial y_p} + N_p^{-1}\frac{\partial\Sigma}{\partial y_p}}\right] = \frac{y_p}{N}\left[-1 + e^{-\beta\frac{\partial F}{\partial y_p}N_p^{-1}}\right] \\ &\approx \frac{y_p}{N}\left[-1 + 1 - \beta\frac{\partial F}{\partial y_p}N_p^{-1}\right] = -\frac{\beta y_p}{NN_p}\frac{\partial F}{\partial y_p}.\end{aligned}$$

Regardless of the value of N_p, the free energy surfaces $F(\boldsymbol{y})$ that are of interest to us, balance out energy and entropy, resulting in a rich layout of saddle points and minima. For these surfaces, the gradient components $\frac{\partial F(\boldsymbol{y})}{\partial y_p}$ are small —around the saddle points as well as the minima—justifying the use of the exponential approximation in the derivation above. For $\frac{\partial E(\boldsymbol{y})}{\partial y_p} < 0$, we obtain $\langle s_p \rangle = -\frac{\beta(1-y_p)}{NN_p}\frac{\partial F(\boldsymbol{y})}{\partial y_p}$. This means

$$A_p(\boldsymbol{y}, t) = \frac{\langle s_p \rangle}{\Delta t} = -\frac{\beta}{2NN_p\Delta t}\left[1 - sgn\left(\frac{\partial E(\boldsymbol{y})}{\partial y_p}\right)(1-2y_p)\right]\frac{\partial F(\boldsymbol{y})}{\partial y_p}.$$

For the expected squared change in activity $\langle s_p^2 \rangle$, we consider a similar approximation. Assuming $\frac{\partial E(\boldsymbol{y})}{\partial y_p} \geq 0$, we get (again with an approximation until the first order of N_p^{-1} and using the results of Equations B2, B3, and 14):

$$\begin{aligned}\langle s_p^2 \rangle &= \frac{N_p}{N}\left[y_p N_p^{-2} + (1-y_p)e^{-\beta\Delta E} N_p^{-2}\right] = \frac{y_p}{NN_p}\left[1 + \frac{1-y_p}{y_p}e^{-\beta\Delta E}\right] \\ &= \frac{y_p}{NN_p}\left[1 + e^{-\beta\Delta E + N_p^{-1}\frac{\partial\Sigma}{\partial y_p}}\right] = \frac{y_p}{NN_p}\left[1 + e^{-\beta N_p^{-1}\frac{\partial E}{\partial y_p} + N_p^{-1}\frac{\partial\Sigma}{\partial y_p}}\right] = \frac{y_p}{NN_p}\left[1 + e^{-\beta\frac{\partial F}{\partial y_p}N_p^{-1}}\right] \\ &\approx \frac{y_p}{NN_p}\left[1 + 1 - \beta\frac{\partial F}{\partial y_p}N_p^{-1}\right] = \frac{2y_p}{NN_p}.\end{aligned}$$

Regardless of the value of N_p, the free energy surfaces $F(\boldsymbol{y})$ that are of interest to us, balance out energy and entropy, resulting in a rich layout of saddle points and minima. For these surfaces, the gradient components $\frac{\partial F(\boldsymbol{y})}{\partial y_p}$ are small—around the saddle points as well as the minima—justifying the use of the exponential approximation in the derivation above. For $\frac{\partial E(\boldsymbol{y})}{\partial y_p} < 0$, we obtain $\langle s_p^2 \rangle = \frac{2(1-y_p)}{NN_p}$.

Evidently, for the off diagonal elements $G_{pq}(\boldsymbol{y}, t)$ (with $p \neq q$), it holds that $\langle s_p s_q \rangle$ (because a change in one dimension means no change in any other). As a consequence, we find:

$$G_{pp}(\boldsymbol{y}, t) = \frac{\left[1 - sgn\left(\frac{\partial E(\boldsymbol{y})}{\partial y_p}\right)(1-2y_p)\right]}{NN_p\Delta t}$$

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

and

$$G_{pq}(\boldsymbol{y}, t) = 0,$$

with $p \neq q$.

Let us define a common diffusion constant D_p as follows:

$$D_p(\boldsymbol{y}) = \frac{G_{pp}(\boldsymbol{y})}{2} = \frac{\left[1 - sgn\left(\frac{\partial E(\boldsymbol{y})}{\partial y_p}\right)(1 - 2y_p)\right]}{2NN_p\Delta t}.$$

We can now simplify the drift rate to

$$A_p(\boldsymbol{y}) = -\beta D_p(\boldsymbol{y})\frac{\partial F(\boldsymbol{y})}{\partial y_p},$$

resulting in

$$dy_p = -\beta D_p(\boldsymbol{y})\frac{\partial F(\boldsymbol{y})}{\partial y_p}dt + \sqrt{2D_p(\boldsymbol{y})}dW_p,$$

with all W_p independent Wiener processes.

If all pools have equal size $N_p = \frac{N}{k}$, then $D_p(\boldsymbol{y})$ reduces to:

$$D_p(\boldsymbol{y}) = D(\boldsymbol{y}) = \frac{k\left[1 - sgn\left(\frac{\partial E(\boldsymbol{y})}{\partial y_p}\right)(1 - 2y_p)\right]}{2N^2\Delta t}$$

and the Fokker-Planck equation for $p(\boldsymbol{y}, t)$ is equivalent to the following stochastic differential equation:

$$d\boldsymbol{y} = -\beta D(\boldsymbol{y})\frac{\partial F(\boldsymbol{y})}{\partial y_p}dt + \sqrt{2D(\boldsymbol{y})}d\boldsymbol{W},$$

with $\boldsymbol{W}$ a vector of k independent Wiener processes.

http://dx.doi.org/10.1037/a0038536

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.