Computational Brain & Behavior (2022) 5:279–301
https://doi.org/10.1007/s42113-022-00143-4

ORIGINAL PAPER

Relating a Spiking Neural Network Model and the Diffusion Model of Decision-Making

Akash Umakantha[1,2] · **Braden A. Purcell**[3] · **Thomas J. Palmeri**[4,5]

Accepted: 26 May 2022 / Published online: 13 June 2022

Abstract
Many models of decision-making assume accumulation of evidence to threshold as a core mechanism to predict response probabilities and response times. A spiking neural network model (Wang, 2002) instantiates these mechanisms at the level of biophysically-plausible pools of neurons with excitatory and inhibitory connections and has numerous model parameters tuned by physiological measures. The diffusion model (Ratcliff, 1978) is a cognitive model that can be fitted to a range of behaviors and conditions. We investigated how parameters of the cognitive-level diffusion model relate to the parameters of a neural-level spiking model. In each simulated "experiment," we generated "data" from the spiking neural network by factorially combining a manipulation of choice difficulty (via the input to the spiking model) and a manipulation of one of the core parameters of the spiking model. We then fitted the diffusion model to these simulated data to observe how manipulation of each core spiking model parameter mapped on to fitted drift rate, response threshold, and non-decision time. Manipulations of parameters in the spiking model related to input sensitivity, threshold, and stimulus processing time mapped on to their conceptual analogues in the diffusion model, namely drift rate, threshold, and non-decision time. Manipulations of parameters in the spiking model with no direct analogue to the diffusion model, non-stimulus-specific background input, strength of recurrent excitation, and receptor conductances mapped on to threshold in the diffusion model. We discuss implications of these results for interpretations of fits of the diffusion model to behavioral data.

Keywords Decision-making · Diffusion model · Spiking neural network · Response times · Accumulation of evidence

Introduction

Developing a theoretical explanation of cognition or perception often begins by considering the classic three levels of analysis proposed by Marr (1982). The *computational level* asks what the system does and why, considering jointly the goals of the organism and the structure of the environment, as typified by Bayesian theories of the mind (e.g., Anderson, 1990; Oaksford & Chater, 2007; Tenenbaum et al., 2011). The *algorithmic level* asks how computations are implemented, what representations and processes underlie cognition and perception, without necessarily being concerned with their biological realization in neural substrates, as typified by many cognitive models (e.g., Busemeyer, Townsend, Wang, & Eidels, 2015; Love, 2015; Sun, 2008). The *implementation level* asks how mechanisms are physically realized within the biological substrate, namely neurons and their connections in the brain, as typified by traditional theoretical work in computational neuroscience (e.g., Carnvale & Hines, 2006; Dayan & Abbott, 2005). Traditionally, theorists have worked within only one level of analysis, which reflected some combination of training and experience, a thoughtful decision regarding the most appropriate level of analysis for answering a particular question, or a philosophical conviction that one level of analysis was the right level while others were unconstrained approximations. More than four decades after Marr's insights, discussion, and debate continues about the relative usefulness of different

✉ Thomas J. Palmeri
thomas.j.palmeri@vanderbilt.edu
https://catlab.psy.vanderbilt.edu

1 Neuroscience Institute, Carnegie Mellon University, Pittsburgh, PA, USA

2 Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA, USA

3 Squarespace, New York, NY, USA

4 Psychology Department, Vanderbilt University, Nashville, TN, USA

5 Vanderbilt Vision Research Center, Vanderbilt University, Nashville, TN, USA

levels of analysis (e.g., Carandini, 2012; Frank, 2015; Jones & Love, 2011; Logan, Schall, & Palmeri, 2015; Love, 2015).

Rather than embrace one of Marr's levels as the "right" level, some research attempts to build bridges between levels (e.g., Frank, 2015; Ratcliff & Frank, 2012; Zandbelt et al., 2014). In this article, we specifically examine relations between two widely cited and influential models of decision-making: at the implementation level, a *spiking neural network model of decision-making* developed by Wang and colleagues (Wang, 2002; Wong & Wang, 2006), and at the algorithmic level, the more abstract cognitive *diffusion model*[1] (DDM) of Ratcliff and colleagues (Ratcliff, 1978; Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998). The two models share many conceptual similarities. They both explain behavior from decision-making tasks in which subjects must decide among alternative choices given some potentially noisy evidence,[2] both solve this task by assuming that this evidence is accumulated over time to a response threshold, and both have been used to explain aspects of neural activity underlying these forms of decisions. However, the two models differ in scale, scope, and tractability in ways that confer each modeling framework with distinct advantages and disadvantages.

The diffusion model (Ratcliff & Rouder, 1998), in its most common form, assumes a single accumulator with upper and lower evidence thresholds for each of the two possible choice alternatives. The model has relatively few parameters and requires only seconds (or less) to simulate predictions and fitting parameters to observed data often takes minutes (or less). Despite having relatively few degrees of freedom, the model can capture much of the key variability observed in behavior across different subjects, species, tasks, and conditions (e.g., Ratcliff, 2006; Ratcliff & McKoon, 2008; Ratcliff & Smith, 2004; Ratcliff et al., 2010; see also Nosofsky & Palmeri, 1997, 2015). Moreover, the observation that certain neurons in the brain seem to instantiate an accumulation of evidence like that assumed by the diffusion model (e.g., Hanes & Schall, 1996) has spawned growing theoretical research equating accumulator models with different measures of neural activity (Forstmann et al., 2016; Logan, Schall, & Palmeri, 2015; Palmeri, Logan & Schall, 2015; see also Palmeri et al., 2017; Turner et al., 2017). In some cases, the evidence accumulation process has been directly identified with the activity of individual neurons (e.g., Boucher et al., 2007; Mazurek et al., 2003; O'Connell, Shadlen, Wong-Lin, & Kelly, 2018; Purcell et al., 2010; Purcell et al., 2012; Purcell & Palmeri, 2017; Ratcliff et al., 2003; Ratcliff et al., 2007; Smith & Ratcliff, 2004). In other cases, the cognitive processes associated with diffusion model parameters have been linked to more global measures of neural activity including EEG (e.g., O'Connell et al., 2018; O'Connell et al., 2012; Philiastides et al., 2006; Schurger et al., 2012) and functional MRI (e.g., Forstmann et al., 2008; Turner et al., 2013, 2015; White et al., 2012). The tractability and sufficient neural plausibility of abstract cognitive models like the diffusion model make them extremely powerful frameworks to understand normal behavior as well as illness, injury, and disease (e.g., Montague et al., 2012; Logan, Schall, & Palmeri, 2015; Wiecki et al., 2015).

The spiking neural network model (Wang, 2002) explains how decisions can be implemented in realistic neural circuits. In this model, parameters define the cellular and synaptic mechanisms that govern the interactions among pools of neurons within a circuit. These models allow us to ask questions about how changes in network structure, synaptic mechanisms, and neural connectivity can ultimately impact neural responses and behavior. In doing so, the models provide a framework to understand how failures at the cellular level can be linked to behavioral deficits in healthy and clinical populations (Anticevic, Murray, & Barch, 2015; Rolls & Deco, 2011). However, the additional complexity of these models precludes direct fits to behavioral data. For example, the spiking neural network model that we used has 2000 simulated neurons, with each neuron defined by 10 differential equations, with 4 million simulated connections between neurons, and a large number of parameters. A Python implementation of this model required a good part of a day running on an older laboratory workstation to simulate predictions for one set of experimental conditions with a fixed set of parameters. Without any further optimization, it could well take months or more to fit a fully parameterized model to observed data, putting aside questions of whether different parameters are uniquely identifiable or not. And with the relatively large number of parameters of the spiking neural network model, it is not entirely clear which parameters might be allowed to vary and which might need to be fixed a priori across subjects and conditions.

Here we aim to link between the abstract diffusion model and the low-level spiking neural network model. If we can understand the relationships between the two, we may be able to take advantage of the strengths of both frameworks. For example, if we can identify mappings between diffusion model parameters and purported biological mechanisms instantiated by the spiking neural network model, then it may be possible to develop testable hypotheses about neural circuits based on diffusion model fits to behavior. Even if the mapping is not a perfect one-to-one match, we

[1] Variously called the diffusion decision modeling or drift diffusion model (DDM), the Ratcliff diffusion model, or most simply just the diffusion model (of decision-making).

[2] With noise (variability) originating from sources that are internal, external, or both, with variability arising within a single decision (within-trial variability), across decisions (between-trial variability), or both.

may still be able to identify a diffusion model parameter with a subset of potential biological mechanisms. This raises the possibility that a single cognitive mechanism, captured by a single diffusion model parameter, may have several potential neural implementations that are not necessarily consistent across experimental manipulations and subject populations.

To make this mapping, using an approach also used by Frank (2015; see also Ratcliff & Frank, 2012), we first simulated behavioral data using the spiking neural network model, treating the model like a single subject in a decision-making task. We systematically varied different biological mechanisms by systematically varying different parameters in the spiking neural network. Next, we fitted the diffusion model to the simulated data from simulated experimental conditions in much the same way that we would fit the diffusion model to actual observed human or non-human data in different experimental conditions. We then examined how the best-fitting diffusion model parameters systematically varied in response to the simulated behavioral changes introduced by having varied the underlying biophysical mechanisms in the spiking neural network model. A key advantage of this approach is that unlike fitting actual human or non-human behavior, we know the underlying "ground truth" biophysical mechanisms that were manipulated because we manipulated specific parameters in the spiking neural network model. In a sense, we asked how well we could reverse engineer parameters of the spiking neural network model based on fits of the diffusion model, which gives us insight into the potential of reverse engineering when it comes to fitting the diffusion model to behavioral data from real organisms with unknown neural mechanisms.

As we will see, in some cases, we found a direct relationship between the spiking neural network model and the diffusion model. For example, manipulations of the neural threshold firing rate parameter of the spiking neural network model largely mapped onto differences in the fitted diffusion model threshold parameter. But in other cases, we found less obvious or anticipated relationships. For example, manipulations of recurrent excitation in the spiking neural network model — the strength of connections between neurons in a pool that code for the same choice — also mapped onto differences in the fitted threshold parameter of the diffusion model. While there were some clear and meaningful relationships between cognitive-level and neural-level models, there were also some complex and less intuitive relationships. These convergent and divergent mappings between models at different levels of analysis might influence both the interpretation of diffusion model parameters for explaining differences between groups and individuals and the understanding of the functional role of particular brain areas during decision-making tasks.

Overview of a Perceptual Decision-Making Task

Using the spiking neural network model (described in detail later) as our simulated "subject," we simulated data from a two-alternative forced-choice perceptual decision-making experiment with six levels of difficulty, with stimulus conditions ranging from those permitting chance performance to near-perfect performance. Differences in stimulus difficulty map on to differences in input to the spiking neural network model (and differences in drift rate for the diffusion model).

We modeled our simulations after the parameterizations of the spiking neural network model originally used by Wang (2002) to account for data from a motion-discrimination task[3] with primates (Wang, 2002). In this task (e.g., Britten et al., 1992; Mazurek et al., 2003), subjects (human or monkey) view an aperture containing a large number of dots. At each video frame, a proportion of dots are offset coherently from their original location to create apparent motion, while the remaining dots are placed randomly within the aperture. The dots move in one of two directions (e.g., leftward or rightward), and the subject's task is to indicate the direction of motion (e.g., by an eye movement or keypress). We (as did Wang, 2002) consider the RT version of this task (e.g., Roitman & Shadlen, 2002), where the subject responds when ready to make a decision following the onset of motion. The percentage of dots that are coherently offset determines the motion strength (expressed as percent coherence), which varies randomly across trials. Trials with higher motion strength are easy, producing high accuracy and fast responses; trials with lower motion strength are difficult, producing low accuracy and slow responses; trials with zero coherence have no correct response (rewarded randomly). Simulated accuracy and response times (RTs) from a random dot motion decision task are illustrated (and further explained) later in Fig. 3. Note that in some of the figures, we illustrate accuracy and mean RTs (e.g., Fig. 3A), though we always fitted full distributions of correct and error RTs (e.g., Fig. 3B).

Both the spiking neural network and the diffusion model explain decisions like this by accumulating noisy evidence over time. The stimulus (in this case, motion strength and direction) and fidelity of the sensory-perceptual

[3] While the spiking neural network model was applied to data from a motion discrimination task, the model does not instantiate motion-detection circuits in extrastriate visual cortex, nor does it take a motion signal as input. Rather, different levels of motion coherence simply map onto different numeric levels of input to the spiking neural network model. We can think of the model as generically simulating choice decision-making across various levels of difficulty rather than specifically simulating performance in a motion discrimination task per se.

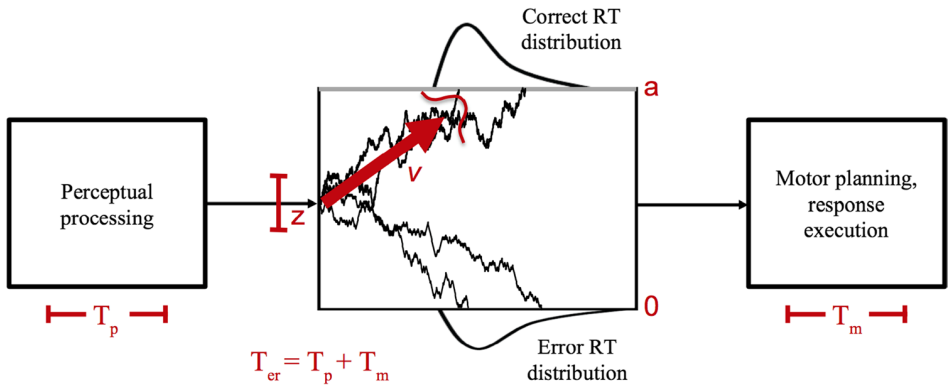

Fig. 1 The diffusion model assumes that RT is partitioned into (1) time for perceptual processing of sensory information (T_p), (2) time for a decision stage assuming accumulation of evidence toward an upper or lower threshold for one of the two possible response alternatives, and (3) time for motor planning and execution (T_m); the times for the first and last stages are summed as a single free parameter (T_{er}). The decision stage assumes evidence starting at a point (z) between the upper (0) and lower (a) response threshold, with a noisy accumulation over time determined by a drift rate (v) with variance s^2. The diffusion model allows for variability across trials in starting point, non-decision time, and drift rate. Decision time is given by the time taken for accumulated evidence to hit one of the two thresholds, with choice given by which threshold is first hit

representations are reflected in the input parameter in the spiking neural network and the drift rate parameter in the diffusion model. For the motion-discrimination task, these inputs are likely computed in brain areas associated with motion perception (e.g., Britten et al., 1992), but neither model specifies the exact computations by which these inputs are generated. By keeping the sensory and perceptual processing generic, the results can apply far more widely than dot motion coherence tasks, to other kinds of decision-making tasks (e.g., Nosofsky & Palmeri, 2015; Ratcliff et al., 2010; Richler & Palmeri, 2014). In addition, neither the spiking neural network nor the diffusion model specify the nature of the motor response, so the models can in principle apply equally to the initiation of saccadic eye movements, limb movements, or finger keypresses (Cisek et al., 2009; de Lafuente et al., 2015).

Dot motion coherence is a bottom-up stimulus manipulation that affects accuracy and RTs. Following Wang (2002; Roitman & Shadlen, 2002), we simulated six levels of coherence (0.0%, 3.2%, 6.4%, 12.8%, 25.6%, 51.2%), and these were instantiated as six levels of the input parameter to the spiking neural network model.

In addition to simulating six coherence (difficulty) levels, in each "experiment," we also simulated the effects of a second independent (or quasi-independent[4]) variable. For sake of simplicity, we assumed selective influence (e.g., Sternberg, 1998), whereby manipulation of this second independent variable affected the value of one and only one parameter in the spiking neural network. For example, different levels of learning might influence the gain or sensitivity on the inputs to network, different levels of speed-accuracy tradeoff might influence the threshold level of pooled activity needed to initiate a choice response, and different levels of some brain disorder might affect the conductance of a particular neurotransmitter channel.

We simulated data from the spiking neural network model by crossing the first motion coherence (difficulty) independent variable that affects the input to the spiking model with the second independent (or quasi-independent) variable that affects one parameter of the spiking model and then fitted the diffusion model to that data, with a particular eye to what differences in diffusion model parameters were required to account for simulated manipulations of the spiking neural network parameter.

Diffusion Model

The diffusion model (Forstmann et al., 2016; Ratcliff, 1978; Ratcliff & Rouder, 1998) is a member of a class of cognitive models that assume that decisions are made by an accumulation of evidence over time to a response threshold (e.g., see Ratcliff & Smith, 2004). The basic elements of the diffusion model are illustrated in Fig. 1. In a common parameterization, evidence begins at a starting point (z) and accumulates toward the upper (a) or lower (0) response threshold. When one of the thresholds is reached, a response is made. In the case of a motion discrimination task, the upper a threshold might be associated with a rightward

[4] Quasi-independent in the sense that they could simulate difference between different subjects (for example, caused by differences in genetics, development, IQ, sex, age, experience, disease, damage, and whole host of other factors not manipulated by the experimenter).

response and the lower *0* threshold associated with a leftward response.

Accumulation in the diffusion model is noisy. The rate of accumulation is called the drift rate (v) and its within-trial variability over accumulation is specified by a diffusion coefficient (s^2), which is often fixed by convention. The drift rate is assumed to take on a different value in different stimulus conditions. In the case of the motion-discrimination task, drift rates have been shown to scale linearly with motion strength (e.g., Palmer et al., 2005; Mazurek, Ditterich, & Shadlen, 2003; Churchland et al., 2008). The specific scaling is dictated by a drift-rate sensitivity parameter (k) such that $v = k \cdot C$, where C is the motion strength (proportion coherence), where k can be thought of as a gain control parameter that is modulated by learning or attention (e.g., Treue & Maunsell, 1999).

Both the time for sensory/perceptual encoding and the time for motor planning/response execution are captured by a non-decision time parameter (T_{er}). The RT on a given trial is determined by adding time to reach a decision threshold and the T_{er}.

In addition to the within-trial variability introduced by the noisy accumulation of evidence, applications of the diffusion model often assume between-trial variability in drift rate (normally distributed with mean 0 standard deviation η), non-decision time variability (uniformly distributed with range s_t), and starting point variability (uniformly distributed with range s_z). In typical applications of the diffusion model, the drift rate (v) varies over stimulus conditions, response threshold (a) varies over speed-accuracy stress conditions, and the relative starting point (whether z differs from $a/2$) varies over bias manipulation conditions; other parameters are often assumed to remain constant across conditions. Related models assume a discrete random walk instead of a continuous diffusion (e.g., Link, 1975; Nosofsky & Palmeri, 1997, 2015) or eliminate within-trial variability and only preserve between-trial variability (e.g., Brown & Heathcote, 2008), but generally account for behavioral data with conceptually similar cognitive mechanisms (e.g., Donkin et al., 2011).

When fitted to response probabilities and distributions of correct and error RTs (e.g., Ratcliff & Tuerlinckx, 2002), the diffusion model captures most of the key properties of observed data, including the shape of correct and error RT distributions; how those distributions change over experimental conditions, fast errors and slow errors; and how behavior changes under different speed-accuracy stress instructions (e.g., Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998), as well as individual differences in decision-making (e.g., Ratcliff et al., 2006, 2010; Shen & Palmeri, 2016).

Spiking Neural Network Model

Unlike cognitive models, neural models are developed at the implementation level of Marr's hierarchy. Wang (2002) proposed a neural model of decision-making that was based on a model originally developed for working memory (Brunel & Wang, 2001) and applied it to the motion-discrimination task (which is why we illustrate using a motion discrimination task here). The model is comprised of a network of thousands of integrate-and-fire neurons that account for the synaptic events and changes in membrane potential that lead to spikes. The influence of spikes on the post-synaptic membrane potential is governed by a system of differential equations that mimic the current dynamics associated with specific ion channels. Synaptic weights determine the strength of connections between different neural pools. Following known biological constraints, the influence of a single spike is relatively short lived (< 25 ms), but recurrent, excitatory connections within the network allow the model to integrate inputs over longer timescales (Wang, 1999), allowing for decision-making (as well as working memory under different parameterizations of the same model).

A number of considerations led us to choose the Wang (2002) spiking neural network model as our candidate neural-level model to compare with the cognitive-level DDM. Other neurally constrained models have articulated the network of different brain areas implicated in various decision-making tasks (e.g., Frank, 2006; Frank & Claus, 006) with neural elements somewhat more abstract than the particular integrate-and-fire neurons used by Wang (2002), and indeed the relationships between such models and the DDM have been examined (Ratcliff & Frank, 2012). The spiking neural model does not simulate all of the detailed processes by which spikes are generated, or any of the morphological properties of real neurons, so it is a more abstract model than how an even lower-level biophysical model might be realized (one that would be even less computationally tractable for our purposes). Key was that we needed a neural-level model that captured neural mechanisms in ways not directly reflected in the cognitive-level DDM, but that also led to behavioral predictions of choice probability and response time. In addition, past work has often considered the spiking neural network model of Wang (2002) to be in essence a neural implementation of an accumulation of evidence process (e.g., Wong & Wang, 2006; Wong et al., 2007), so investigating the mappings between the parameters of this specific neural model and those of the DDM seemed justified.

To simulate a two-alternative choice task, the model assumes that two choice-selective pools of excitatory pyramidal neurons compete to select a response (e.g.,

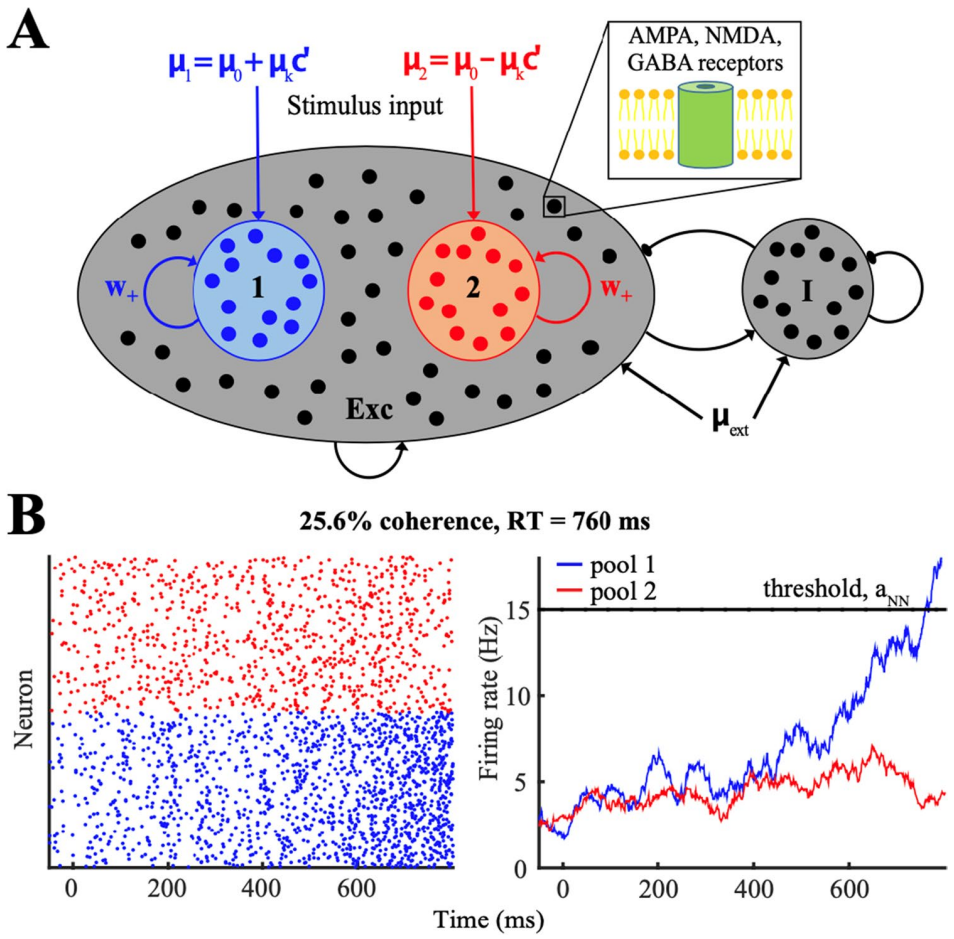

Fig. 2 Spiking neural network model (Wang, 2002). **A** Illustration of the neural network architecture. Neurons are modeled with leaky integrate-and-fire dynamics and have three currents determined by AMPA, NMDA, and GABA receptor conductance values. Two subsets of neurons in the excitatory pool are selective for the two possible choices in the task. These two choice-selective pools are exhibit strong recurrent connections (w_+) to other neurons in the same pool and compete with the other choice-selective pool via an inhibitory pool of neurons. All neurons receive background input (μ_{ext}). Upon stimulus presentation, the two choice-selective pools receive differential input ($\mu_1$and μ_2) based on the stimulus condition (c) and input strength (μ_k). **B** Left: raster plot of spikes from the choice-selective pools, blue represents neurons voting for choice 1 and red represents neurons voting for choice 2. Right: pooled firing rates of the two choice-selective pools. A decision is made and the trial ends when the pooled firing rate of pool 1 or pool 2 reach the neural network threshold (a_{NN} = 15 Hz). On this example trial with coherence c = 25.6%, the pooled firing rate for choice 1 (blue) reached threshold 760 ms after stimulus onset

rightward motion versus leftward motion); a pool of inhibitory interneurons mediates competition between the choice-selective pools and an additional pool of non-selective neurons represents neurons that are not directly associated with the potential choices (Fig. 2). The combination of recurrent excitatory connections and inhibition-driven competition allows the model to generate attractor dynamics that reproduce observed decision-making behavior and neural responses (Wang, 2002).

All neural pools are driven by background inputs that take the form of spike trains with Poisson statistics governed by a mean rate. Following the onset of a stimulus and after a sensory encoding delay, choice-selective pools receive an additional input (increased rate) that depends on the external stimulus. The input rate depends on the difficulty of the discrimination between the alternatives, which in our case is motion strength, as determined by motion coherence. Stronger motion strength corresponds to larger input to the pool associated with a correct choice and weaker input to the pool corresponding to an incorrect choice. As with the diffusion model, input rates scale linearly with motion strength according to an input sensitivity parameter (see Appendix), reflecting the observed linear relationship between motion strength and direction-selective sensory neurons that are a likely input source (e.g., Britten et al., 1992). A choice is made when the pooled firing rates of one choice-selective

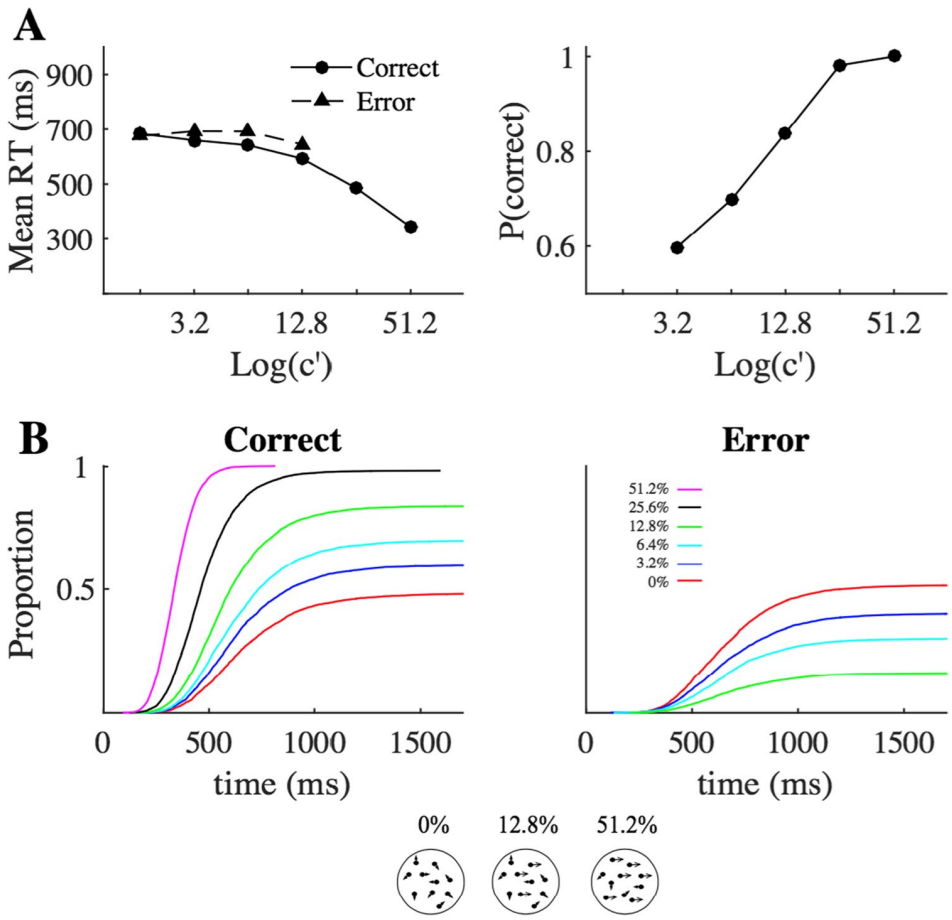

Fig. 3 Simulated "data" from the spiking neural network model (assuming the default parameters from Wang, 2002). **A** shows mean correct and error RT and P(Correct) as a function of motion coherence (x axis). **B** show correct and error (degenerate) cumulative RT distributions as a function of motion coherence (colored lines per the legend). Note that error RTs are not shown for the 25.6% and 51.2% motion coherence conditions because few errors were produced

pool reach a fixed firing rate threshold. The pool that reaches the firing rate threshold first determines the choice, and the time when the pooled activity reaches that threshold determines the RT. Most often, the pool that represents the correct choice and receives the larger input will be the first to reach threshold and determine the choice. But this is not always the case due to noise, especially for more difficult decisions, where the pool that represents the incorrect choice can be the first to reach threshold, which results in the model making errors.

Before delving into the details of the spiking neural network model, we first provide a general overview of its biological mechanisms that we will explore. Because both the spiking neural model and the diffusion models assume a form of evidence accumulation to a threshold, we can use this conceptual connection to establish a priori hypotheses about relationships between the two. We anticipated that mechanisms that have similar effects on the evidence integration process should produce common behavioral predictions and clear parameter mappings between the models. Of course, even though there might be a conceptual similarity — neural thresholds should act like diffusion thresholds, and neural input strength should act like diffusion drift rates — that does not guarantee that there will be a simple one-to-one mapping between models. In particular, the spiking neural network has non-linear dynamics that are in certain ways quite different from the accumulation dynamics of the diffusion model. In other cases, there is no obvious conceptual link between the parameters of the two models — recurrent excitation, synaptic weights, and ionic conductance values in the spiking neural model have no analogue in the diffusion model — so those relationships are not immediately clear a priori.

Roadmap

We first describe the spiking neural network model, including the parameters that we will explore and the procedures for simulating choices and RTs from the spiking model. We then briefly illustrate its predictions of behavior in the standard random dot motion task, replicating results previously presented by Wang (2002). We next describe the diffusion model and the procedures for fitting it to the simulated data from the spiking neural network model. We then describe the simulated experiments in which we systematically manipulated a single spiking neural network

mechanism by varying one of its parameters and generated at each spiking neural model parameter level a new set of behavioral observations for each motion strength condition; for example, we simulated data for five different levels of spiking neural network model threshold and at each threshold simulated six levels of dot motion coherence. We then describe fits of the diffusion model to each simulated 5×6 factorial experiment, examining how manipulations of a particular spiking neural network model parameter mapped on to differences in the drift rate, threshold, and non-decision time parameters in the fitted diffusion model.

Simulation and Model Fitting Methods

Simulating the Spiking Neural Network Model

Network architecture We used a neural architecture designed for perceptual decision-making, with all default model parameters and other aspects of its architecture taken from Wang (2002). This spiking neural network model consists of four neural pools (see Fig. 2). Two pools of 240 choice-selective excitatory pyramidal neurons represent the alternative choices in the task (one pool for a decision that motion is rightward, one pool for a decision that motion is leftward). An additional pool of 1120 background excitatory neurons represents non-choice-selective neurons (i.e., neurons representing responses not relevant for the current task). A pool of 400 inhibitory interneurons implements competition between these pools.

Network Input All simulated neurons receive stochastic external background inputs as spike events sampled from a Poisson process with a mean spike rate of μ_{ext} throughout each simulated trial. Following Wang (2002), we used a default background input value of $\mu_{ext} = 2400$ Hz. In one of our simulated experiments, we systematically manipulated the amount of background input to the two choice-selective pools using $\mu_{ext} = 2390$, 2395, 2400 (default), 2405, and 2410 Hz; below this range, recurrent dynamics were not strong enough to support long-term integration over time, and above this range, the model became unstable in the pre-stimulus period leading to preemptive choices.

On each simulated trial, we allowed a 1-s pre-stimulus period for the network to reach a stable baseline state before stimulus presentation. Then after a stimulus processing time (T_s),[5] additional stimulus-driven inputs were added to the inputs to the choice-selective pools, while the background and inhibitory pools continued to receive only background input (Fig. 2). Stimulus processing time was not explicitly included in the original Wang (2002) model, but we include it here both for biological plausibility and because this time is assumed to contribute to the non-decision time assumed by the diffusion model. To identify the relationship between diffusion model parameters and stimulus processing time in the spiking neural model, in our simulated experiments, we explored a range of T_s values: 0 (default), 50, 100, 150, and 200 ms (Figs. 3 and 4).

After the 1-s pre-stimulus period and stimulus processing time T_s, stimulus inputs were presented to the choice-selective neural pools (added to the background inputs μ_{ext}) until the choice is made and the simulated trial ends. The mean spike rate of the stimulus inputs are governed by the equations:

$$\mu_1 = \mu_0 + \mu_k c$$

$$\mu_2 = \mu_0 - \mu_k c$$

where μ_1 corresponds to the pool representing the one choice (say, leftward for a particular simulated trial), and μ_2 corresponds to the other choice (say, rightward for a particular simulated trial). For simplicity, all simulations assume symmetry between leftward and rightward choices (the decision is not biased and parameters for the two pools are identical); so we will sometimes refer to pool 1 as the one associated with the correct choice and pool 2 with the incorrect choice (in the same way that for the diffusion model, we will sometimes refer to the upper threshold as the one associated with the correct choice and the lower threshold with the incorrect choice). The stimulus inputs have a fixed component, μ_0, and scale linearly with motion strength in percent coherence, c (following Britten et al., 1992). The input sensitivity parameter, μ_k, determines how input scales with motion strength. Following Wang (2002), we used a default value of $\mu_0 = 40$ Hz and $\mu_k = 40$ Hz. To identify the relationship between diffusion model parameters and input sensitivity, we explored the following range of μ_k values: 20, 30, 40 (default), 50, and 60 Hz. These values produced a large range of simulated RTs and choice probabilities (see Fig. 5), while keeping the neural network within a region of network dynamics that allows for competition between the two pools and for a choice decision to emerge (Wong & Wang, 2006).

Neuron and Synaptic Dynamics Mathematical details of the neuronal and synaptic dynamics are provided in Appendix.

[5] Stimulus processing time (T_s) reflects both afferent conductance delays from the retinal stimulation and time needed to process the stimulus input to drive the appropriate choice-selective pools.

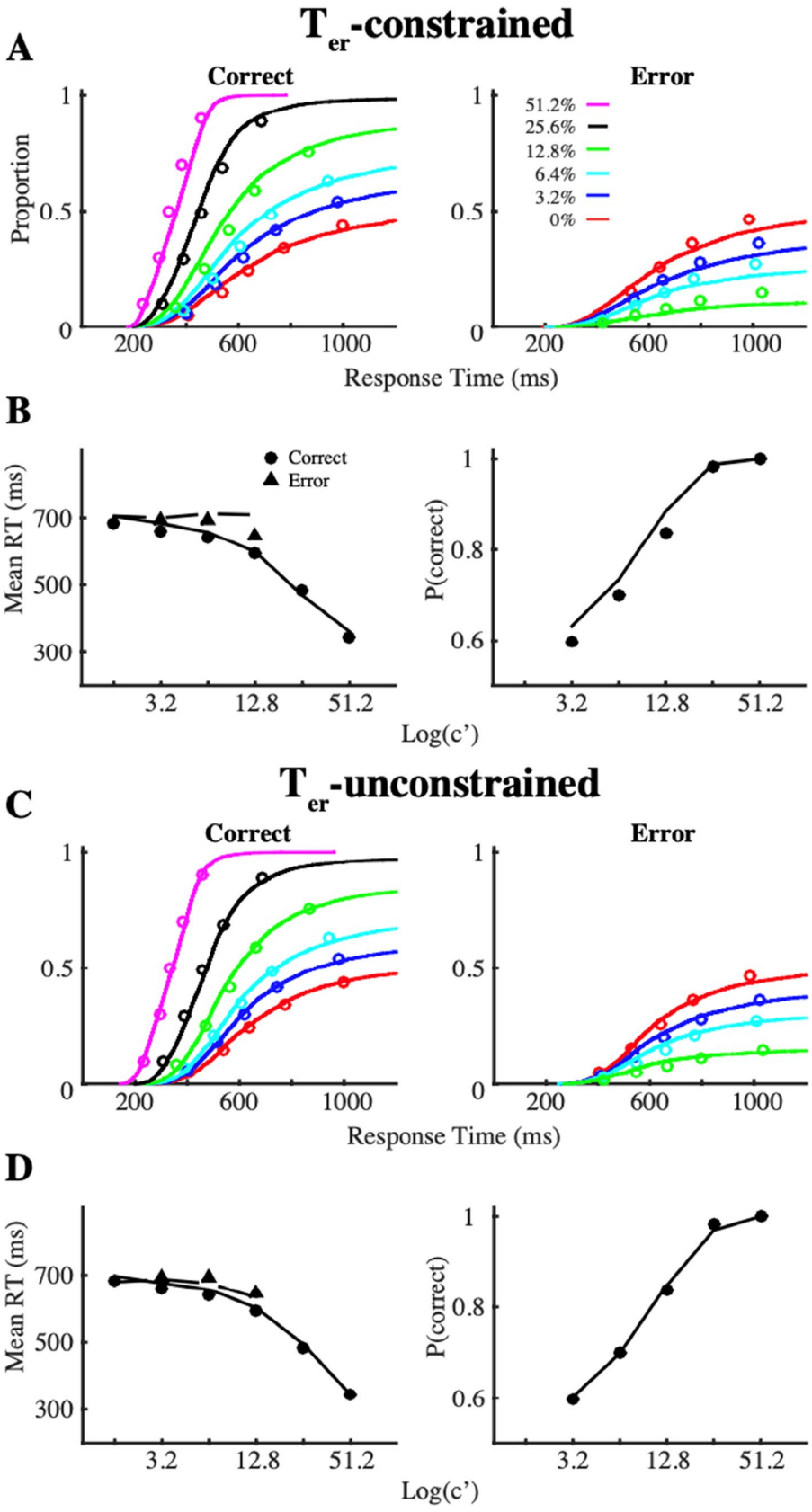

Fig. 4 All panels show simulated "data" from the spiking neural network model (assuming default parameters from Wang, 2002) as symbols and diffusion model predictions are lines (simulated "data" are identical to those shown in Fig. 3). **A** and **C** show correct and error (degenerate) cumulative RT distributions as a function of motion coherence (colored lines per the legend). **B** and **D** show mean correct and error RT and P(Correct) as a function of motion coherence (horizontal axis). Simulated "data" in **A** and **B** are identical to simulated "data" in **C** and **D**. **A** and **B** show diffusion model predictions for T_{er}-constrained version as lines. **C** and **D** show diffusion model predictions for the T_{er}-unconstrained version as lines

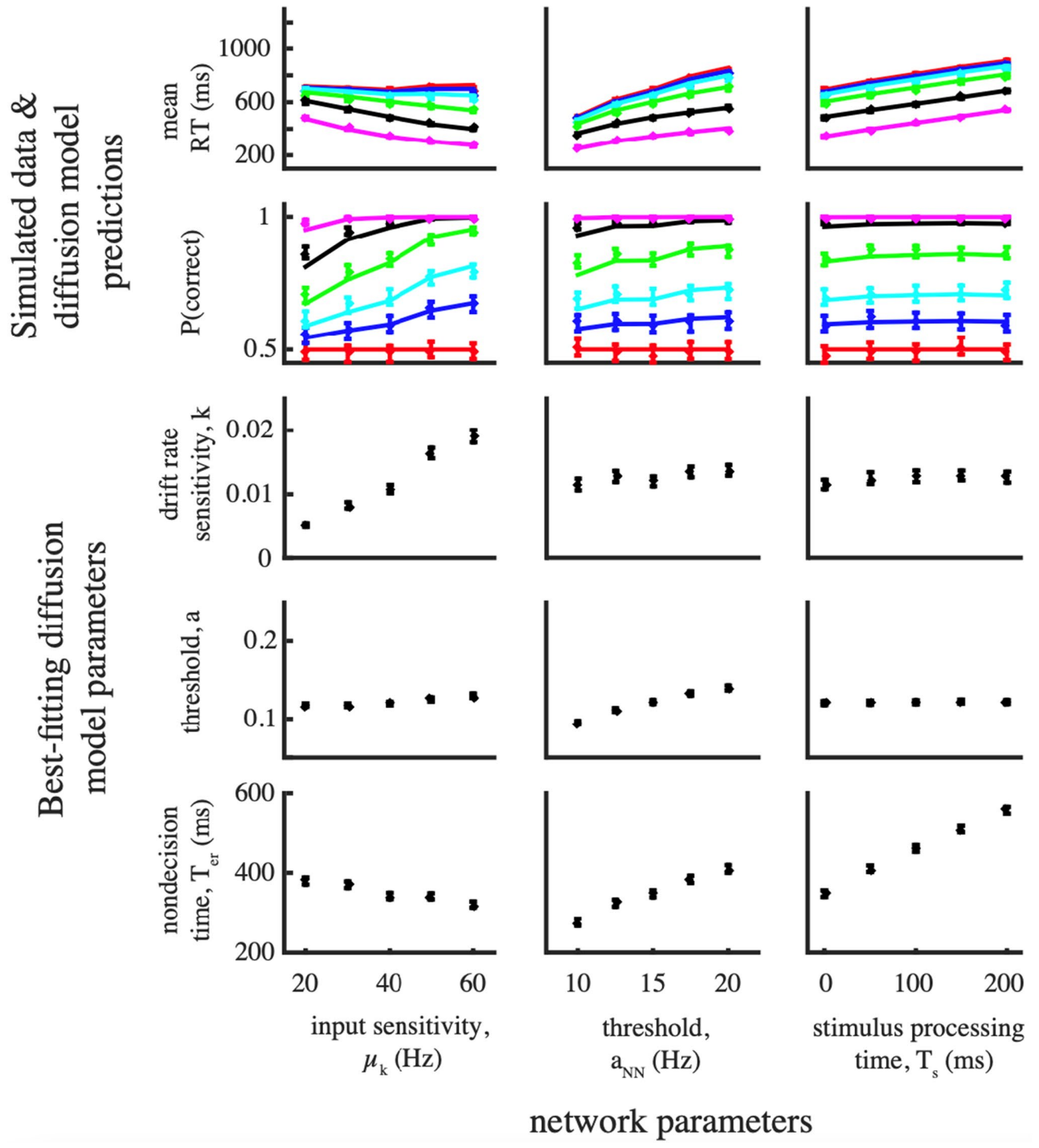

Fig. 5 Mapping between the manipulated spiking neural network parameters and fitted diffusion model parameters. Top two rows show mean RTs and proportion correct, where data points and error bars indicate spiking neural network simulated data and lines indicate fitted diffusion model predictions (color coded by motion coherence following legend as in Figs. 3 and 4). Bottom three rows show the best-fitting diffusion model parameters (error bars indicate 95% bootstrapped confidence intervals); median best-fitting parameters were used to generate diffusion model predictions of mean RTs and accuracy in the top two rows. This figure shows mappings between spiking neural network parameters (input sensitivity, threshold firing rate, stimulus processing time) with a conceptual relationship to diffusion model parameters

Each and every neuron in the network is modeled as a leaky integrate-and-fire neuron (e.g., Abbott, 1999). A differential equation defines the change in membrane potential over time. Once the membrane voltage crosses a threshold level, a discrete spike is generated, the voltage is reset to its resting potential, and it is prevented from spiking again during its refractory period. The spike generated by that neuron serves as a weighted input to other neurons it is connected to.

The change in membrane voltage over time is determined by the synaptic current that serves as the input to the neuron from the other neurons it is connected to. The spiking neural network assumes four types of currents: external AMPA (excitatory), recurrent AMPA (excitatory), recurrent NMDA (excitatory), and recurrent GABA (inhibitory). The precise dynamics of these synaptic currents is given by a series of equations specified in Appendix.

We systematically manipulated two classes of parameters in the spiking neural network model that affect the synaptic current dynamics:

One class of parameters we manipulated was the synaptic weights that control the strength of connections between neural pools. The spiking neural network is connected all-to-all, meaning that any single neuron, j, is connected to every other neuron in the network with weight w_j. The default synaptic weight is $w_j = 1.000$. Neurons representing the same choice (say, leftward) have stronger connections $w_j = w_+$, whereas neurons representing the opposite choice have weaker connections w_- (see Appendix for details). Following Wang (2002), we used the default value $w_+ = 1.700$. To identify the relationship between diffusion model parameters and recurrent synaptic weights, in our simulated experiments, we used $w_+ = 1.650$, 1.675, 1.700 (default), 1.725, 1.750; beyond this range, the model failed to exhibit competitive dynamics necessary to generate a decision (Wong & Wang, 2006).

The other class of parameters we manipulated was the conductance values (g) of three different receptor channels (recurrent AMPA, NMDA, and GABA). We used the following parameter ranges: $g_{rec,AMPA} = 0.0450$, 0.0475, 0.0500 (default), 0.0525, and 0.0550 nS; $g_{NMDA} = 0.163$, 0.164 (default), 0.165, 0.166, and 0.167 nS; $g_{GABA} = 1.290$, 1.295, 1.300 (default), 1.305, and 1.310 nS. As with other choices of manipulated parameter values, these values were chosen because they allowed the model to exhibit appropriate competitive dynamics of the network for it to exhibit decision-making behavior while also producing reasonable differences in RTs and accuracy.

Predicting Choice and RT After stimulus processing time (T_s), the stimulus inputs (μ_1 and μ_2) were presented to the network, and the two pools (rightward vs. leftward) competed until the average firing rate of one pool reached a fixed threshold (a_{NN}). Following Wang (2002), average pooled firing rates were calculated using a sliding window[6] of 50 ms with a step size of 5 ms, and a default threshold of $a_{NN} = 15$ Hz was assumed. Because we assumed complete symmetry between rightward and leftward decisions, we define a correct trial as a trial in which the pool receiving the larger input μ_1 reached threshold first and an error trial as one in which the pool receiving the smaller input μ_2 reached threshold first. To explore the relationship between diffusion model parameters and spiking neural network model threshold, we varied a_{NN} in this range: 10.0, 12.5, 15.0 (default), 17.5, 20.0 Hz.

Simulation Methods For each of the six motion strengths, we simulated 5000 trials of the spiking neural network and recorded each trial's RT and choice (correct vs. error). Numeric simulations of the system of differential equations (Appendix) used timesteps of 0.02 ms. Figure 3 shows the mean correct and error RTs, choice probabilities, and full RT distributions (as degenerate[7] cumulative distribution functions) produced by the neural network using the default parameters assumed in Wang (2002), which are the default parameters we used. The model reproduces key characteristics of behavioral data in perceptual decision-making tasks. Consistent with experimental observations in humans and monkeys (Ratcliff & McKoon, 2008; Roitman & Shadlen, 2002), the proportion of correct choices increases, and the mean RT decreases with increased motion strength (Table 1).

To generate each "simulated experiment," we repeated this process while systematically varying one spiking neural network parameter at a time, keeping all others at their default values. In each simulated experiment, five different values of a given neural model parameter were used (as noted earlier), and we simulated the six motion strengths for each parameter setting. The range of parameter values was chosen to produce reasonable differences in accuracy and RT while also maintaining the competitive and attractor dynamics of the model necessary for it to exhibit decision-making (Wong & Wang, 2006). In total, each simulated experiment contained 5000 trials at each of the six motion strengths for each of five spiking neural network model parameter values, for a total of 30,000 simulated trials.

In summary, simulated experiments were conducted varying the following spiking neural network model parameters: stimulus processing time (T_s), input sensitivity (μ_k), neural response threshold (a_{NN}), background input (μ_{ext}), recurrent synaptic weight (w_+), and conductances for AMPA recurrent ($g_{rec,AMPA}$), NMDA (g_{NMDA}), and GABA (g_{GABA}) receptor channels.

[6] Using a slightly larger or smaller window size did not qualitatively change model simulations.

[7] So-called "degenerate" distributions asymptote at the rate of correct or error proportion in each condition, rather than asymptote at 1.

Table 1 Fits to behavioral equations to quantify the effects of coherence on RTs and accuracy for the simulations using standard parameters from Wang (2002). In the equations, ***coh*** represents coherence and can take values of 0.000, 0.032, 0.064, 0.128, 0.256, 0.512 and ***acc*** represents accuracy. *p* values for all coefficients are smaller than 0.001. Larger coherence values resulted in smaller RTs (negative regression coefficient in the top table) and more correct responses (positive regression coefficient in the middle table)

$$\boldsymbol{RT = B_0 + B_1 coh}$$

Parameter	*Value*	*Standard Error*
B_0	0.689	0.00164
B_1	-0.692	0.00681

$$\boldsymbol{Logit[P(corr)] = B_0 + B_1 coh}$$

Parameter	*Value*	*Standard Error*
B_0	-0.0914	0.0208
B_1	14.62	0.274

$$\boldsymbol{RT = B_0 + B_1 coh + B_2 acc}$$

Parameter	*Value*	*Standard Error*
B_0	0.713	0.00249
B_1	-0.655	0.00738
B_2	-0.0388	0.00306

Fitting the Diffusion Model

Overview We fitted the diffusion model to behavioral data (error and correct RT distributions and accuracy) simulated by the spiking neural network model. As described earlier, the key conceptual parameters of the diffusion model are the drift rate (v), threshold (a), and non-decision time (T_{er}). Like most applications of the diffusion model, we allowed subsets of these parameters to vary freely across different experimental conditions, as described in detail below. Parameters that govern across-trial variability of parameters (s_z, s_t, and η) allow the diffusion model to capture certain aspects of observed data (e.g., Boehm, et al., 2018), and like most applications, these parameters were assumed to be the same across different experimental conditions. Because the spiking neural network model assumed no bias, we fixed the starting point (z) halfway between the upper and lower

Table 2 Bayesian information criterion (BIC) and Chi-squared values for the T_{er}-constrained and T_{er}-unconstrained diffusion model fits to spiking neural network simulated data

	BIC	χ^2
T_{er}-constrained	87,519	1,426
T_{er}-unconstrained	87,188	431

response threshold of the diffusion model. The standard deviation of within-trial variability, s, was fixed at 0.1 by convention (Ratcliff & Rouder, 1998).

Before fitting the diffusion model to simulated experiments and examining fits to manipulations of spiking neural network model parameters, we explored how well different parameterizations of the diffusion model fitted the simulated data from the basic motion discrimination task (before introducing manipulations of the second independent/quasi-independent variable). Drift rate is usually assumed to vary with stimulus difficulty (e.g., Nosofsky & Palmeri, 1997; Palmeri, 1997; Ratcliff & Rouder, 1998), so we expect that differences of drift rate v should fully explain behavior as a function of motion strength. As described earlier, we assumed that $v_C = k \cdot C$, where C is the proportion motion coherence and k is a gain parameter that scales motion strength onto drift rate. We confirmed that a model assuming that only drift rate varied linearly with motion strength, with all other parameters held constant across motion strength, could qualitatively explain key aspects of the simulated spiking neural network model data (Fig. 4A–); we also confirmed with nested model comparisons that fitted diffusion model threshold (a) remains relatively fixed at a constant value with motion strength.[8]

However, our initial model explorations revealed that the quantitative fit could be improved by allowing for some systematic differences in non-decision time (T_{er}) with stimulus strength as well (Table 2). We observed the same improvement when we fitted the same model to existing behavioral data from human subjects performing the motion-discrimination task (Purcell & Kiani, 2016), indicating that this may be a property of real data as well. To ensure that our results were not dependent on whether non-decision time was free to vary with motion strength (the model producing better fit) or whether non-decision time was fixed across motion strength (as commonly assumed in many studies), we fitted both versions of the model to all simulated experiments. For the T_{er}-constrained model, T_{er} was fixed across motion strengths. For the T_{er}-unconstrained model, T_{er} varied linearly with motion strength as follows: $T_{er} = T_0 + T_{coh} \cdot c$, where c is the motion strength (proportion coherence) and T_{coh} is a free parameter. T_0 is an intercept term (i.e., the non-decision time at 0% motion strength). We report the results for both models in Fig. 4 and summarize the T_{er}-constrained model in Table 2. Ultimately, our qualitative results were identical for both sets of model assumptions. While T_{coh} is useful for the best quantitative fitting of the diffusion model to simulated spiking network data at different motion strengths, for the rest of the study when we refer to T_{er}, we are precisely referring to T_0.

In the main body of the manuscript, we fitted the diffusion model to simulated experiments, where we crossed the manipulation of five levels of one specific spiking neural network model parameter with six levels of motion coherence. For these diffusion model fits, we continued to fit full correct and error RT distributions like those displayed in Fig. 4, but we show only mean correct RTs and accuracy in the later figures for sake of clarity. All diffusion model parameters were held fixed across simulated experiment conditions (levels of a spiking neural network model parameter) except for the drift rate gain parameter (k_v), threshold (a), and non-decision time (T_{er}).

Model Fitting Approach We generated diffusion model predictions of cumulative distribution functions for correct and error RTs using a combination of closed-form solutions and numerical methods described in Tuerlinckx (2004). We used a Bayesian information criterion (BIC; Schwarz, 1978) statistic built on quantile maximum likelihood estimation (QMLE; Heathcote et al., 2002) to estimate the best-fitting diffusion model parameters for each simulated experiment.[9] Log likelihood, $ln(L)$, was computed as follows:

$$ln(L) = \sum_{x} \sum_{m} \sum_{r} \sum_{i} N_{i,r,m,x} ln(\pi_{i,r,m,x})$$

where i indexes RTs falling in bins determined by the 0.1, 0.3, 0.5, 0.7, and 0.9 quantiles of the cumulative probabilities

[8] For drift rate, we compared the linear assumption described in the text with a more complex model assuming a separate free drift rate parameter for each motion coherence condition; for threshold, we compared the constant value assumption described in the text with a more complex model assuming a separate free threshold parameter for each motion coherence condition. Nested model comparisons were conducted using the same BIC-based methods described later in this article.

[9] This is a common approach to maximum likelihood parameter estimation for models that predict response time (see Heathcote et al., 2002; Farrell & Lewandowsky, 2018 for details). Observed RT distributions are defined in terms of quantile bins, and model predictions are evaluated based on how well the predicted probabilities account for the observed frequencies with which responses fall into each RT bin. This essentially turns a continuous data space into a discrete data space using a multinomial distribution to link model predicted probabilities to observed frequencies and amounts to evaluating a model based on its fit to the (defective) cumulative RT distribution functions (see Van Zandt, 2000). While the diffusion model does predict probability density functions for RT, fitting the cumulative distribution function helps to mitigate against the overinfluence that outliers (overly fast or overly slow individual RTs) can have on likelihood using a pdf approach; for example, just a single observed RT from one trial out of thousands that falls below the minimum possible RT predicted by a set of diffusion model parameters can send the log likelihood to negative infinity using a pdf approach.

of the RT distributions, r indexes trial outcomes (correct or error), m indexes motion strength, and x indexes the manipulation of the spiking neural network parameter. $N_{i,m,r}$ is the observed number of trials falling within a particular RT bin at a given trial outcome, motion strength, and condition, and $\pi_{i,m,r}$ is the corresponding diffusion model predicted probability of an RT falling within that bin. Note that π is defined as the joint distribution over stimulus conditions and outcomes so it implicitly fits the choice probabilities (accuracy) for each condition as well as RT. BIC was then computed from

$$BIC = -2ln(L) + Mln(N)$$

where M is the number of free parameters (16 for the T_{er}-unconstrained model and 11 for the T_{er}-constrained model) and N is the total number trials across all conditions. We fitted RT quantiles because they are robust to the influence of outliers, while still providing a good summary of the shape of the RT distribution (e.g., Heathcote et al., 2002; Ratcliff, 1979; Van Zandt, 2000). For some simulated conditions (in particular, high motion coherence), the observed error rates were extremely low, and it was necessary to use alternative RT binning procedure to prevent these data from exerting an undue influence on the fitting process; specifically, if the error rate was less than 5%, we binned data using the median (two bins), and if it was less than 2%, we did not fit error RTs for that particular condition at all (zero bins); this is an approach we have used in past work (e.g., Boucher et al., 2007; Servant et al., 2019). For all fits, BIC was minimized using a hill-climbing SIMPLEX method (Nelder & Mead, 1965) implemented in MATLAB (MathWorks). We also used a bootstrapping procedure (Wagenmakers et al., 2004) with 1000 starting points to obtain confidence intervals of the best-fitting diffusion model parameters.

To draw conclusions about the relationship between diffusion model and spiking neural network model parameters, we used a nested model testing approach (e.g., see Farrell & Lewandowsky, 2018). First, we fitted an unrestricted (all-free) version of the diffusion model where particular key parameters of interest (threshold a, non-decision time T_{er}, and drift rate gain k_v) were allowed to vary freely across conditions where a particular spiking neural network parameter was manipulated. Second, we fitted restricted versions of diffusion models where a key parameter (a, T_{er}, or k_v) was fixed across those conditions. The best-fit BIC values of the four models (all free, a-fixed, T_{er}-fixed, and k_v-fixed) were compared to determine which diffusion model parameter needed to be free to adequately fit the simulated neural spiking model data well. If the BIC of a fixed model (a-fixed, T_{er}-fixed, or k_v-fixed) was smaller than or reasonably close to the BIC of the all-free model, then differences in that diffusion model parameter were not necessary to explain the simulated behavioral changes caused by differences in the spiking neural model parameter. If the BIC of the fixed model was somewhat larger than that of the "all free" model, then that diffusion parameter had a weak mapping to the spiking neural network parameter. If it was much larger, then the diffusion parameter had a strong mapping to the spiking neural network parameter.[10]

Modeling Results

We begin by describing manipulations of spiking neural network model parameters that we expected to have an intuitive mapping onto diffusion model parameters — specifically, those arguably most likely to correspond to drift rate, threshold, and non-decision time. We then move to describing manipulations of spiking neural model parameters that do not have any clear a priori mapping onto diffusion model parameters.

Input Sensitivity (μ_k) Input sensitivity (μ_k) in the spiking neural network model can be seen as representing the efficiency of stimulus processing and encoding. Mathematically, in our simulations, μ_k multiplicatively scales the inputs to the spiking model (as a function of motion strength), thereby acting as a gain parameter on the inputs to the network.[11] Because μ_k scales the input to the neural model as a function of motion coherence (difficulty), we expected differences in this neural parameter to map onto differences in drift rate gain (k_v), which scales the drift rate of the diffusion model as a function motion coherence (difficulty).

We simulated the spiking neural network at five levels of μ_k: 20, 30, 40 (default), 50, and 60 Hz. Simulated RTs decreased and accuracy increased as a function of input sensitivity. As expected, we see a direct one-to-one mapping between μ_k in the spiking neural network model and k in the diffusion model (Fig. 5, Table 3). Larger values of μ_k in the spiking model results in shorter RTs and higher accuracy with a corresponding increase in k_v in the fitted diffusion

[10] We chose to compare the fit of the unrestricted (all-free) model with the fit of each restricted (one-fixed) model to mirror a common model comparison approach whereby adding a parameter constraint (in this case fixing the value of one free parameter across conditions) allows the modeler to ask whether flexibility in that parameter is necessary to maintain an adequate fit to the observed data (in other words, does the added parameter constraint "break" the model, producing a significantly worse fit; for example, see Farrell and Lewandowsky, 2018).

[11] The spiking neural network model itself does not provide a theory of the inputs or the gain parameter, in much the same way that the diffusion model does not provide a theory of the drift rates (e.g., Nosofsky & Palmeri, 1997, 2015).

Table 3 Nested model comparisons (T_{er}-constrained). We tested which diffusion model parameters needed to vary to capture changes in a given spiking neural network parameter (rows). Values in the first column are BIC when all diffusion model parameters (threshold a, non-decision time T_{er}, or drift rate gain k_v) were allowed to vary as we manipulated a spiking network parameter (indicated by labels on the left). Subsequent columns show the difference in BIC when either a, T_{er}, *or* k_v were held fixed as network parameters changed versus when they all allowed to vary. A small or negative value in these columns indicates that the corresponding "fixed" diffusion model parameter was not necessary to explain differences in behavior for the corresponding spiking neural parameter. A large positive value in the final three columns indicates that the corresponding diffusion parameter had a strong mapping to the corresponding spiking neural network parameter

	Free model	*Fix a-free*	*Fix T_{er}-free*	*Fix k_v-free*
Input sensitivity	437,590	265	1,075	12,872
Threshold	437,994	2,607	3,910	−150
Background input	441,697	11,080	5,232	−445
Stim. processing time	438,918	42	10,648	90
Recurrent excitation	441,833	11,830	636	−506
AMPA conductance	440,382	12,436	4,841	−469
NMDA conductance	440,862	16,355	5,774	−373
GABA conductance	440,212	10,541	4,938	−380

model. Interestingly, we find that no other parameter that we vary in the spiking neural network maps to diffusion drift rate, suggesting that this result is a true one-to-one mapping (for the set of parameters manipulated).

Threshold Firing Rate (a_{NN}) The spiking neural network makes a decision when the firing rate for one of its two direction-selective neuronal pools crosses its threshold firing rate (a_{NN}). We suspected a direct mapping between a_{NN} and diffusion threshold because the two are analogous — they both determine the threshold on accumulated activity for when a decision is made.

We simulate the spiking neural network at five levels of a_{NN}: 10, 12.5, 15 (default), 17.5, and 20 Hz. Simulated RTs increased and accuracy increased for larger threshold firing rates. We see that there is mapping between threshold firing rate and diffusion model threshold — as the value of a_{NN} increases, the fitted diffusion model threshold increases as well (Fig. 5, Table 3). However, a surprisingly large proportion of the systematic differences in the simulated data is also captured by differences in the non-decision time parameter (T_{er}). In fact, fixing non-decision time resulted in worse fits than did fixing threshold (Table 3), suggesting that non-decision time maps more strongly to a_{NN} than does threshold.

This complex mapping of a_{NN} makes some sense when considering the spiking neural network as an attractor network with two fixed points (one where pool 1 wins, and another when pool 2 wins). Although a_{NN} is the criteria for when the behavioral response is initiated, the neural network is essentially "committed to a choice" at an earlier point in time because of its competitive dynamics. The point at which the mean firing rates of the two pools diverge is when the network diverges to an attractor state and "makes" a choice. After this point, the period of competition is essentially over, and the network is unlikely to leave the attractor state and "change its mind." In many cases, pooled firing rates diverge (i.e., the choice is made) well before a_{NN} is crossed. In this way, further increasing a_{NN} has little effect on the accuracy of responses, but does cause longer RTs. Analogously, increases in diffusion model T_{er} account for longer RTs, without any effect on accuracy. On the other hand, changing diffusion model threshold (a) impacts both RTs and accuracy.

Stimulus Processing Time (T_s) Stimulus processing time (T_s) in the spiking neural network model encapsulates biophysical conductance delays and the time needed for upstream neurons to process incoming stimulus information. This sensory-perceptual processing time in the diffusion model is a component with the non-decision time parameter (T_{er}). As such, we expected T_s to map to diffusion T_{er}.

We simulated the spiking neural network at five levels of T_s: 0 (default), 50, 100, 150, and 200 ms. Figure 5 shows that RTs become longer, accuracy remains constant, and fitted diffusion model non-decision time increases with increasing T_s, as predicted (Table 3). Because T_s and T_{er} are the most closely related parameters between the spiking neural network and diffusion model, this experimental result serves as a simple and intuitive validation of our cross-fit methodology. Note though that even when stimulus processing time in the spiking neural network is set to 0, the diffusion model fits say that there is a significant amount of nondecision time; this difference is likely because the nonlinear dynamics intrinsic to the spiking neural network model cause the two pools of neurons associated with the two choices to remained at baseline, locked in competition with one another for quite a while, which "looks like" additional nondecision time from the perspective of the diffusion model fitting data stimulated by that process.

Background Input (μ_{ext}) While all simulated neurons receive background input (μ_{ext}) in the spiking neural network model, we only manipulated the level of background input for the two pools associated with the correct and incorrect response (both manipulated up or down in the same way).

Physiologically, μ_{ext} represents the afferent inputs from a variety of sources when there is no stimulus present; for example, background input might reflect the internal state of the subject prior to the stimulus onset. An increase in background input (μ_{ext}) makes choice-selective pools have an increased baseline firing rate before stimulus onset, meaning that they are both closer to their response thresholds, and there is greater variability in their firing rates because a larger Poisson spike rate comes with larger spiking variability.

We expected that increased μ_{ext} would present behaviorally in simulations as shorter RTs and decreased response accuracy. Based on this anticipated speed-accuracy tradeoff, we expected that μ_{ext} in the spiking neural network model may map on to diffusion threshold (a). Note that while manipulations of baseline (starting point) in a (linear integrator) model like the diffusion model (more precisely in multi-accumulator variants like the dual diffusion model, e.g., Ratcliff et al., 2007) are identifiable with manipulations of threshold, the inherent nonlinearities in the spiking neural network model do not guarantee that manipulations of background input (μ_{ext}) in the spiking neural network model will map uniquely onto manipulations of threshold in the diffusion model.

We simulated the spiking neural network at five levels of background input μ_{ext}: 2390, 2395, 2400 (default), 2405, and 2410 Hz; here, we only manipulated μ_{ext} in the choice-selective pools, with non-choice neurons and inhibitory neurons continuing to receive μ_{ext} = 2400 Hz. Simulations (Fig. 6) produced decreased RT and accuracy with increased levels of background input. In fits of the diffusion model, manipulated background input in the spiking neural network mapped very strongly on to diffusion threshold and somewhat strongly on to non-decision time (Table 3).

We also conducted simulations in which we varied μ_{ext} to all neurons in the spiking neural network (not shown). In this case, mean RTs were shorter, but, somewhat surprisingly, accuracy was unaffected, even at very large values of μ_{ext} (e.g., 3000 Hz). In fits of the diffusion model, this resulted in a strong mapping on to diffusion non-decision time but no mapping at all on to threshold. Manipulating the balance between excitation and inhibition in the spiking neural network, which can be done by selectively manipulating the background input to choice-relevant pools only, may be necessary for the model to simulate the kind of speed-accuracy tradeoff reflected by differences in the threshold parameter in the diffusion model.

Strength of Recurrent Excitation Within Choice-Selective Pools (w+) The w_+ parameter in the spiking neural network model is the strength of connections between neurons within a pool selective for the same choice. In a real brain, w_+ could increase through long-term potentiation and the formation of new synapses and could decrease with synaptic pruning and certain forms of neural degeneration. For example, Wang et al. (2011) suggested that age-related working memory declines might be associated with molecular mechanisms (e.g., cAMP signaling) that could cause a decrease in such recurrent excitation.

There is no obvious, a priori, intuitive mapping between the w_+ parameter in the spiking neural network model and a parameter in the diffusion model. On one hand, w_+ can be thought of as controlling how quickly firing rates rise as a form of multiplicative self-excitation, perhaps analogous to the rate of accumulation of evidence (i.e., drift rate) in the diffusion model. On the other hand, Wong and Wang (2006) reported that an increase in w_+ in the spiking neural network model could cause faster and less accurate decisions, a characteristic of speed-accuracy tradeoff, captured by threshold (a) in the diffusion model.

We simulated the spiking neural network at five levels of w_+: 1.650, 1.675, 1.700 (default), 1.725, and 1.750. As w_+ increased, both mean RTs and accuracy decreased (Fig. 6). When fitting the diffusion model to this simulated data, a very strong mapping on to diffusion threshold and a strong mapping on to non-decision time were observed (Fig. 6, Table 3). w_+ is another parameter of the spiking neural network that implements a speed-accuracy tradeoff. We confirmed that that baseline firing rate (and firing rate variability) were unaffected by manipulations of w_+. In the spiking neural network model, a change in the balance of excitation and inhibition (in this case, a change in excitation caused by a change in recurrent inputs) can be a neural mechanism of speed-accuracy tradeoff.

Conductance of Receptors on Excitatory and Inhibitory Neurons The currents that mediate recurrent excitatory connections in the network are implemented by AMPA and NMDA receptors. Physiologically, AMPA and NMDA conductance may be manipulated via neuromodulators. For example, dopaminergic neuromodulation through D1 receptors catalyze the potentiation of NMDA receptors (Cepeda et al., 1993). Conductance can also be altered pharmacologically via local or systemic injections of agonists or antagonists. For example, in response time tasks, small doses of the NMDA-antagonist ketamine injected into awake-behaving primates have been shown to increase both accuracy and RTs during visual search (Shen et al., 2010).

While there is no a priori, intuitive mapping between these conductance parameters of the spiking neural network and parameters of the diffusion model, we expected both to have similar behavioral effects and thereby map on to the same diffusion model parameters because they both mediate the strength of recurrent excitatory connections (their main difference is their time constants).

Here, we varied NMDA conductance at five levels (g_{NMDA} = 0.163, 0.164, 0.165 (default), 0.166, and 0.167 nS)

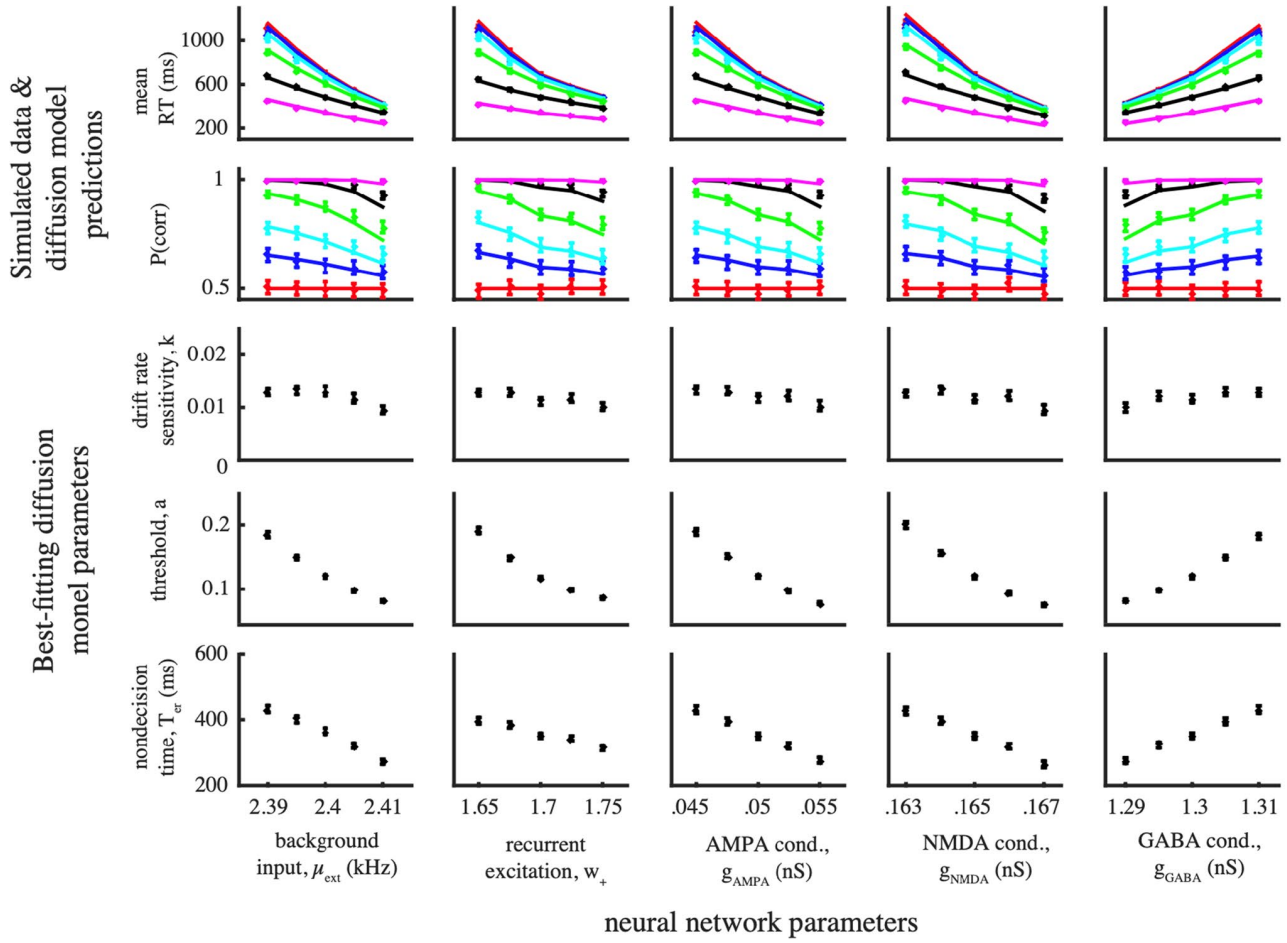

Fig. 6 This figure shows mappings between spiking neural network parameters with no obvious a priori conceptual relationship to diffusion model parameters. Format is the same as Fig. 5. Top two rows show mean RTs and proportion correct, where data points and error bars indicate spiking neural network simulated data and lines indicate fitted diffusion model predictions. Bottom three rows show the best-fitting diffusion model parameters (error bars indicate 95% confidence intervals); median best-fitting parameters were used to generate diffusion model predictions of mean RTs and accuracy in the top two rows

and AMPA conductance at five levels ($g_{rec,AMPA} = 0.0450$, 0.0475, 0.0500 (default), 0.0525, and 0.0550 nS). As we expected, manipulations of both parameters had similar results. Increases in the conductance caused RTs and accuracy to decrease. When fitting the diffusion model to these simulated data, manipulation of these conductance parameters both mapped on to threshold and non-decision time in the diffusion model (Fig. 6, Table 3). This is yet another speed-accuracy tradeoff caused by a change in the balance between excitation and inhibition (in this case, an increase in excitation).

The inhibitory currents (g_{GABA}) in the spiking neural network are implemented by GABA receptors that mediate feedback inhibition. In a real brain, neuromodulation or psychopharmacological manipulations are possible ways that GABA conductance could be manipulated.

We simulated the spiking neuron network model manipulating GABA conductance values ($g_{GABA} = 1.29$, 1.295, 1.3 (default), 1.305, and 1.31 nS). Increasing GABA conductance results in slower RTs and more accurate decisions. GABA conductance implements a speed-accuracy tradeoff. When fitting the diffusion, manipulations of GABA conductance mapped on to diffusion model threshold and non-decision time parameters (Fig. 6, Table 3). In this case, manipulations of GABA conductance caused changes in baseline firing rate, which we observed earlier as another neural manifestation of a speed-accuracy tradeoff (see also Brunel & Wang, 2001, for effects of decreasing NMDA:GABA ratio on activity in the spiking neural network).

In other simulations (not shown), we manipulated the conductance values of inhibitory neurons and found the converse effects. Increasing AMPA and NMDA conductance

on inhibitory neurons caused slower RTs and more accurate decisions, and increasing GABA conductance on inhibitory neurons (i.e., inhibiting the inhibitory neurons) caused faster RTs and less accurate decisions. When fitting the diffusion model, manipulations of these spiking neural network inhibitory conductance parameters mapped to diffusion model a and T_{er} parameters, but in the opposite direction of the mapping of their excitatory counterparts. Again, the balance between excitation and inhibition in the spiking neural network model can instantiate speed-accuracy tradeoffs reflected by modulation of threshold and non-decision time in the diffusion model.

Discussion

Numerous successful models of decision-making assume an accumulation of evidence to threshold as a core mechanism for predicting observed response probabilities and distributions of response times. The well-known diffusion model (Ratcliff, 1978; Ratcliff & Rouder, 1998) is a relatively abstract cognitive model with its drift rate parameter representing the mean strength of evidence driving noisy accumulation over time from a starting point to an upper or a lower absorbing boundary associated with each of two choice alternatives, with the distance between the two boundaries a threshold parameter representing some speed-accuracy tradeoff, with response time given by the time at which one of the boundaries is first hit plus a non-decision time parameter representing the time for perception and motor response. A spiking neural network model (Wang, 2002) instantiates accumulation of evidence at the level of biophysically plausible pools of many neurons having various excitatory and inhibitory connections with a large number of model parameters, many of which are informed by neurophysiological measures.

We asked how parameters of the diffusion model at the cognitive level relate to parameters of the spiking neural network model at the biophysical implementation level. To do so, we simulated the spiking neural network model to generate data (response probabilities and distributions of response times) that were then fitted by the diffusion model. Certain individual parameters in the spiking neural network model were systematically manipulated, producing simulated data in a form much like one might observe in a behavior experiment with manipulated independent (or quasi-independent) variables. Having the known levels of parameters of the spiking neural network model and associated best-fitting parameters of the diffusion model allowed us to relate model mechanisms across levels of abstraction. We focused on asking whether parametric differences across simulated conditions produced by the spiking neural network model were accounted for by differences in the drift rate, threshold, or non-decision time parameters in the diffusion model.

In summary, manipulations of parameters in the spiking neural network model related to input sensitivity, threshold, and stimulus processing time mapped on to their conceptual analogues in the diffusion model, namely differences in fitted drift rate, threshold, and non-decision time parameters, respectively. Manipulations of parameters in the spiking neural network model with no direct analogue to the diffusion model, including non-stimulus-specific background input, strength of recurrent excitation, and various receptor conductances, all mapped on to differences in fitted threshold in the diffusion model. Manipulations these parameters likely affect the nonlinear growth of neural activity to threshold in the spiking neural network model in ways that lead to behaviors analogous to that produced by manipulating the level of threshold; this is in some ways analogous to the relationship between non-homogenous drift rate and collapsing bounds in an accumulator model (e.g., see Cisek et al., 2009; Hawkins et al., 2015).

Our work complements work investigating speed-accuracy tradeoffs in the spiking neural network model. Lo, Wang, and Wang (2015) also showed that increasing non-stimulus-specific input to pools of neurons representing the two choice alternatives could trade speed for accuracy and investigated how that non-specific signal might impact the dynamics of the system. We show that other manipulations of spiking neural network parameters (in addition to threshold and baseline input) can produce a speed-accuracy tradeoff. Moreover, we showed that these manipulations map onto differences in fitted diffusion model threshold.

Under the assumption that diffusion model mechanisms are an appropriate model of cognitive mechanisms, differences in fitted threshold in diffusion model fits are often interpreted in terms of potential neural processes that modulate some combination of baseline firing rates or threshold firing rates (e.g., Bogacz et al., 2009; Forstmann et al., 2016; Heitz, 2014; Servant et al., 2019; but see Heitz & Schall, 2012). Differences in fitted diffusion model threshold have been observed under different speed-accuracy instructions given to participants (e.g., Ratcliff, 2006; Ratcliff & Rouder, 1998; Wagenmakers et al., 2008) as well as between different age groups of participants (e.g., Ratcliff et al., 2006, 2010). While these differences in diffusion model threshold could indeed be caused by modulation of the same neural threshold, like the threshold in the spiking neural network model, they could also be caused by different underlying neural processes. For example, perhaps task modulation by speed-accuracy instruction could be caused by modulation of a neural threshold whereas age-related differences could be caused by differences in the strength of recurrent excitation or level of neuromodulation. Both are revealed as

diffusion model threshold in model fits but are quite different neural processes. Of course, any lack of identifiability in underlying neural processes based on differences in diffusion model threshold is founded on an assumption that the spiking neural network model is an appropriate model of the underlying neural dynamics of brain mechanisms involved in decision-making.

It is interesting to note that when fitting data simulated by manipulating parameters in the spiking neural network, we observed that fitted diffusion model threshold and non-decision time tended to change together. Indeed, fits of the diffusion model to performance by different age groups of participants often find that higher diffusion model thresholds are accompanied by longer diffusion model non-decision times (e.g., Ratcliff et al., 2006, 2010).

Cross-fitting studies have been informative in relating alternative models at the same level of abstraction. For example, Donkin et al. (2011) simulated data from the diffusion model and fitted that data with the linear ballistic accumulator (LBA) model (Brown & Heathcote, 2008) and simulated data from the LBA and fitted that data with the diffusion model; both are accumulation of evidence models, but differ in certain assumptions about the nature of variability within and across trials. Overall, they observed a close correspondence between the associated drift rate, threshold, and non-decision time parameters that are core to both models. Data simulated by explicit manipulations of threshold (or drift rate, or non-decision time) in the diffusion model were accounted for by differences in threshold (or drift rate, or non-decision time) in the LBA, and vice versa. The identifiability of parameters was not entirely unique, however. For example, manipulations of threshold to produce simulated data with the LBA were accounted for by differences not only in threshold in the diffusion model, but also by (monotonic) differences in drift rate and (non-monotonic) differences in non-decision time in the LBA. Manipulations of threshold to produce simulated data with the diffusion model were accounted for by differences not only in the threshold in the LBA, but also by (monotonic) differences in non-decision time. While it would be ideal to do cross-fitting in both directions as in Donkin et al. (2011), because of the computational complexity of the spiking neural network model, we could only generate simulated data from the spiking neural network model and fit using the diffusion model, not the other way around.

Our work is certainly not the first to try to relate processes at the level of spiking neurons involved in decision-making with cognitive-level accumulation of evidence models. Perhaps the first was the seminal work of Hanes and Schall (1996) who showed that the measured dynamics of movement-related neurons in Frontal Eye Field (FEF) during saccade decision-making can be explained as variable accumulation of evidence to a fixed threshold, much like the dynamics assumed by the diffusion model. Years of research since involving both neurophysiological work and computational modeling supported and significantly expanded this view (e.g., see Gold & Shadlen, 2007; Logan et al., 2015; Palmeri et al., 2015; Schall, 2001, 2004; Shadlen & Newsome, 2001; Smith & Ratcliff, 2004).

While we took the approach simulating data from the neural model and fitting using the cognitive model, other work has aimed to mathematically simplify the neural model into a form more akin to a cognitive model. Wong and Wang (2006) used a mean-field approach to progressively simply mathematically the spiking neural network model with thousands of integrate-and-fire neurons and dozens of parameters to a far more simplified model with two choice-state variables and far fewer parameters. The resulting model was similar in form (though different in details concerning its dynamics and parameterization) to the leaky competing accumulator (LCA) model of Usher and McClelland (2001), which itself can be related to a diffusion model under certain parameter regimes (Bogacz et al., 2006). Our work complements that of Wong and Wang (2006). They showed a mathematical relationship between the spiking neural network model and an abstract cognitive cousin. We outlined relationships between manipulations of parameters of the spiking neural network and those of the more abstract diffusion model.

For both its tractability and (sufficient) neural plausibility, to understand normal behavior as well as illness, injury, and disease, abstract cognitive models like the diffusion model may be the "just right" level for making significant theoretical progress (e.g., Montague et al., 2012; Logan, Schall, & Palmeri, 2015; Wiecki et al., 2015). Indeed, some have even suggested that any link between low-level neural circuit models and behavior may be "a bridge too far" (Carandini, 2012). When it comes to understanding how differences in cognitive-level mechanisms, processes, and parameters relate to the underlying neural mechanisms, especially those at the level of individual neurons, it can be useful to investigate how theoretical cognitive-level mechanisms map on to measured neural-level dynamics (e.g., Cox et al., in press; Purcell et al., 2010, 2012; Servant et al., 2019). It can also be useful to investigate how properties of abstract cognitive models generalize to models that instantiate cognitive mechanisms within large ensembles akin to those at work within the brain (e.g., Cox, Lilburn, Logan, Schall, & Palmeri, under revision; Zandbelt et al., 2014). And as we have demonstrated here, it can also be productive to investigate these relationships by mapping from cognitive-level mechanisms, processes, and parameters to those instantiated by neural-level models like the spiking neural network model.

Appendix Spiking Neural Network Model Dynamics

Neuron and synapse models. Each neuron in the spiking network is modeled as a leaky integrate-and-fire neuron (e.g., Abbott, 1999) with a membrane potential *V(t)* defined by a differential equation

$$C_m \frac{dV(t)}{dt} = -g_L\left(V(t) - V_L\right) - I_{syn}(t)$$

where C_m is the membrane capacitance, g_L is the leakage conductance, V_L is the resting potential, and $I_{syn}(t)$ is the synaptic current. When the membrane potential $V(t)$ of a neuron reaches a threshold potential V_{th} = -50 mV, the neuron generates a spike and is reset to $V_r = -55$ mV for a refractory period of T_{ref}. For excitatory neurons, $C_m = 0.5$ nF, $g_L = 25$ nS, V_L = -70 mV, and $T_{ref} = 2$ ms; for inhibitory neurons, $C_m = 0.2$ nF, $g_L = 20$ nS, V_L = -70 mV, and $T_{ref} = 1$ ms. We fixed these parameters at values that replicate known biophysical properties of cortical neurons (as per Wang, 2002).

The total synaptic current $I_{syn}(t)$ is the sum of external input currents and excitatory and inhibitory currents from recurrent network connections. There are four types of currents at the synaptic connections: external AMPA, recurrent AMPA, recurrent NMDA, and recurrent GABA. Therefore, the equation for total synaptic current $I_{syn}(t)$ is:

$$I_{syn}(t) = I_{ext,AMPA}(t) + I_{rec,AMPA}(t) + I_{NMDA}(t) + I_{GABA}(t)$$

The individual receptor currents are given by:

$$I_{ext,AMPA}(t) = g_{ext,AMPA}\left(V(t) - V_E\right)s_{ext,AMPA}(t)$$

$$I_{rec,AMPA}(t) = g_{rec,AMPA}\left(V(t) - V_E\right)\sum_{j=1}^{Ce} w_j s_{j,AMPA}(t)$$

$$I_{rec,NMDA}(t) = \frac{g_{NMDA}\left(V(t) - V_E\right)}{1 + [Mg^{2+}]e^{-0.062V(t)/3.57}}\sum_{j=1}^{Ce} w_j s_{j,NMDA}(t)$$

$$I_{rec,GABA}(t) = g_{GABA}\left(V(t) - V_I\right)\sum_{j=1}^{Ci} s_{j,GABA}(t)$$

where $V_E = 0$ and $V_I = -70$ mV are the reversal potentials and $[Mg^{2+}] = 1$ mM is the extracellular magnesium concentration. The sums are over all excitatory connections C_e or all inhibitory connections C_i. We fixed these parameters at values that replicate known biophysical properties of cortical neurons (as per Wang, 2002).

The g variables are conductance values of the receptor channels. All default conductance values were chosen to match known biophysical measurements (Wang, 2002). For excitatory neurons $g_{ext,AMPA} = 2.100$ nS, $g_{rec,AMPA} = 0.050$ nS, $g_{NMDA} = 0.165$ nS, and $g_{GABA} = 1.300$ nS. For inhibitory neurons $g_{ext,AMPA} = 1.620$ nS, $g_{rec,AMPA} = 0.040$ nS, $g_{NMDA} = 0.130$, and $g_{GABA} = 1.000$ nS. To explore the relationship between diffusion model parameters and conductance values, in simulated experiments, we used the following ranges: $g_{rec,AMPA} = 0.045$, 0.0475, 0.050 (default), 0.0525, and 0.055 nS; $g_{NMDA} = 0.163$, 0.164, 0.165 (default), 0.166, and 0.167 nS; and $g_{GABA} = 1.290$, 1.295, 1.300 (default), 1.305, and 1.310 nS. As with other choices of manipulated parameter values, these values were chosen because they allowed the model to exhibit appropriate competitive dynamics while also producing reasonable differences in predicted behavior.

The w variables are synaptic weights that control the strength of connections between neural pools. The network is connected all-to-all, meaning that a single neuron, j, is connected to every other neuron in the network with weight w_j. The synaptic weights were structured with a "Hebbian rule" with stronger connections between neurons representing the same choice ($w_j = w_+$) and weaker connections between neurons representing opposite choices ($w_j = w_-$). All other connections assume $w_j = 1$. Following Wang (2002), we used the default value $w_+ = 1.7$ and w_- was determined by the Eq. $1 - f(w_+ - 1)/(1 - f)$. f is the proportion of excitatory neurons in the spiking neural network that belong to a choice-selective subpool (in our case, $f = 240/1600 = 0.15$). To identify the relationship between diffusion model parameters and synaptic weights, we explored the following ranges: $w_+ = 1.650$, 1.675, 1.700 (default), 1.725, and 1.750. Beyond these ranges, the model failed to exhibit competitive dynamics (Wong & Wang, 2006).

The s variables are gating variables and represent the fraction of receptor channels that are open at a given time. These parameters control the receptor dynamics and are presumed to be a fixed biophysical property of the cells. They are governed by the following equations:

$$\frac{ds_{j,AMPA}(t)}{dt} = \sum_k \delta(t - t_j^k) - \frac{s_{j,AMPA}(t)}{\tau_{AMPA}}$$

where $\tau_{AMPA} = 2$ ms is the decay time and the sum is over the presynaptic spikes from neuron j. For external AMPA currents, the sum is over Poisson spike trains generated with means specified by the parameters μ_{ext}, μ_1, and μ_2 that are independent for each cell. For NMDA receptors:

$$\frac{ds_{j,NMDA}(t)}{dt} = \alpha x_j(t)(1 - s_{j,NMDA}(t)) - \frac{s_{j,NMDA}(t)}{\tau_{NMDA,decay}}$$

$$\frac{dx_j(t)}{dt} = \sum_k \delta(t - t_j^k) - \frac{x_j(t)}{\tau_{NMDA,rise}}$$

where $\tau_{NMDA,decay} = 100$ ms and $\alpha = 0.5$/ms. The rise time dynamics of NMDA receptors is modeled by x_j, where $\tau_{NMDA,rise} = 2$ ms. Finally, for GABA:

$$\frac{ds_{j,GABA}(t)}{dt} = \sum_k \delta(t - t_j^k) - \frac{s_{j,GABA}(t)}{\tau_{GABA}}$$

where $\tau_{GABA} = 100$ ms.

Acknowledgements We thank Xiao-Jing Wang for providing a Python-based implementation of his spiking neural network model that was adapted for use in our simulations. Portions of this article represented an undergraduate honors thesis by the first author (AU) when he was at Vanderbilt University. Much of the original modeling work and writing was done when the second author (BAP) was a graduate student at Vanderbilt University and then a postdoctoral fellow at New York University.

Author Contribution All authors contributed to the study conception and design. Simulation and model fits were performed by Akash Umakantha, with significant input from Braden Purcell and Thomas Palmeri. The first draft of the manuscript was written by Akash Umakantha and Braden Purcell, the final version edited by Thomas Palmeri. All authors read and approved the final manuscript.

Funding This work was supported by NEI grant R01 EY021833, the Temporal Dynamics of Learning Center (NSF grant SMA 1041755), and the Vanderbilt Vision Research Center (P30 EY008126).

Data Availability Simulation code is available from the authors upon request.

Declarations

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Competing interests The authors declare no competing interests.

References

Abbott, L. F. (1999). Lapicque's introduction of the integrate-and-fire model neuron (1907). *Brain Research Bulletin, 50*(5–6), 303–304.

Boehm, U., Annis, J., Frank, M.J., Hawkins, G.E., Heathcote, A., Kellen, D., ... & Wagenmakers, E.J. (2018). Estimating across-trial variability parameters of the diffusion decision model: Expert advice and recommendations. Journal of Mathematical Psychology, 87, 46-75

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2009). The neural basis of the speed–accuracy tradeoff. *Trends in Neuroscience, 33*, 10–16.

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.

Boucher, L., Palmeri, T. J., Logan, G. D., & Schall, J. D. (2007). Inhibitory control in mind and brain: An interactive race model of countermanding saccades. *Psychological Review, 114*, 376–397.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12*, 4745–4765.

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

Brunel, N., & Wang, X.-J. (2001). Effects of neuromodulation in a cortical network model. *Journal of Computational Neuroscience, 11*, 63–85.

Busemeyer, J.R., Townsend, J.T., Wang, Z.J. , & Eidels, A. (Eds.) (2015). *Mathematical and computational models of cognition*. Oxford University Press.

Carandini, M. (2012). From circuits to behavior: A bridge too far? *Nature Neuroscience, 15*, 507–509.

Cepeda, C., Buchwald, N. A., & Levine, M. S. (1993). Neuromodulatory actions of dopamine in the neostriatum are dependent upon the excitatory amino acid receptor subtypes activated. *Proceedings of the National Academy of Sciences, 90*(20), 9576–9580.

Cisek, P., Puskas, G. A., & El-Murr, S. (2009). Decisions in changing conditions: The urgency-gating model. *Journal of Neuroscience, 29*(37), 11560–11571.

Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience, 11*, 693–702.

Cox, G.E., Palmeri, T.J. Logan, G.D., Smith, P.L., Schall, J.D. (in press). Salience by competitive and recurrent interactions: Bridging neural spiking and computation in visual attention. *Psychological Review*. PsyArXiv: https://psyarxiv.com/rkh8g/

Cox, G.E., Lilburn, S., Logan, G.D., Schall, J.D., Zandbelt, B., & Palmeri, T.J. (under revision). Decision making by ensembles of accumulators. *Manuscript under revision*. PsyArXiv: https://psyarxiv.com/qdk7b/

Dayan, P., & Abbott, L. F. (2005). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. MIT Press.

de Lafuente, V., Jazayeri, M., & Shadlen, M. N. (2015). Representation of accumulating evidence for a decision in two parietal areas. *Journal of Neuroscience, 35*(10), 4306–4318.

Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E.-J. (2011). Diffusion versus linear ballistic accumulation: Different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review, 18*, 61–69.

Farrell, S., & Lewandowsky, S. (2018). *Computational modeling of cognition and behavior*. Cambridge University Press.

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., Von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E. J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences, 105*, 17538–17542.

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E. J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology, 67*, 641–666.

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*(6), 272–279.

Frank, M. J. (2015). Linking across levels of computation in model-based cognitive neuroscience. In B. Forstmann & E. J. Wagenmakers (Eds.), *An Introduction to Model-Based Cognitive Neuroscience* (pp. 159–177). Springer Neuroscience.

Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological Review, 113*(2), 300–326.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30*, 535–560.

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science, 274*, 427–430.

Hawkins, G. E., Forstmann, B. U., Wagenmakers, E. J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience, 35*(6), 2476–2484.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9*(2), 394–401.

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*, 150.

Heitz, R. P., & Schall, J. D. (2012). Neural mechanisms of speed-accuracy tradeoff. *Neuron, 76*, 616–628.

Jones, M., & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences, 34*, 169–231.

Link, S. W. (1975). The relative judgment theory of two choice response time. *Journal of Mathematical Psychology, 12*(1), 114–135.

Lo, C.-C., Wang, C. T., & Wang, X. J. (2015). Speed-accuracy tradeoff by a control signal with balanced excitation and inhibition. *Journal of Neurophysiology, 114*(1), 650–661.

Logan, G.D., Schall, J.D., & Palmeri, T.J. (2015). Inhibitory control in mind and brain: The mathematics and neurophysiology of the underlying computation. In B. Forstmann & E.J. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience*, Springer Neuroscience.

Love, B. C. (2015). The algorithmic level is the bridge between computation and brain. *Topics in Cognitive Science, 7*(2), 230–242.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman and Company, New York.

Mazurek, M. E., Roitman, J. D., Ditterich, J., & Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cerebral Cortex, 13*, 1257–1269.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences, 16*, 72–80.

Nelder, J. A., & Mead, R. (1965). A simplex method for function optimization. *The Computer Journal, 7*(4), 308–313.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104*, 266–300.

Nosofsky, R.M., & Palmeri, T.J. (2015). Exemplar-based random walk model. To appear in J.R. Busemeyer, J. Townsend, Z.J. Wang, & A. Eidels (Eds.), *Mathematical and computational models of cognition*, Oxford University Press.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.

O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience, 15*, 1729–1735.

O'Connell, R. G., Shadlen, M. N., Wong-Lin, K., & Kelly, S. P. (2018). Bridging neural and computatio nal viewpoints on perceptual decision-making. *Trends in Neurosciences, 41*(11), 838–852. Chicago.

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision, 5*(5), 376–404.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 324–354.

Palmeri, T. J., Love, B. C., & Turner, B. M. (2017). Model-based cognitive neuroscience. *Journal of Mathematical Psychology, 76*, 59–64.

Palmeri, T.J., Schall, J.D. & Logan, G.D. (2015). Neurocognitive modelling of perceptual decisions. To appear in J.R. Busemeyer, J. Townsend, Z.J. Wang, & A. Eidels (Eds.), *Oxford handbook of computational and mathematical psychology*, Oxford University Press.

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*(35), 8965–8975.

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2010). Neurally-constrained modeling of perceptual decision making. *Psychological Review, 117*, 1113–1143.

Purcell, B. A., & Kiani, R. (2016). Neural mechanisms of post-error adjustments of decision policy in parietal cortex. *Neuron, 89*, 658–671.

Purcell, B. A., & Palmeri, T. J. (2017). Relating accumulator model parameters and neural dynamics. *Journal of Mathematical Psychology, 76*, 156–171.

Purcell, B. A., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2012). From salience to saccades: Multiple-alternative gated stochastic accumulator model of visual search. *Journal of Neuroscience, 32*(10), 3433–3446.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*, 446–461.

Ratcliff, R. (2006). Modeling response signal and response time data. *Cognitive Psychology, 53*, 195–237.

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.

Ratcliff, R., & Frank, M. J. (2012). Reinforcement-based decision making in corticostriatal circuits: Mutual constraints by neurocomputational and diffusion models. *Neural Computation, 24*, 1186–1229.

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.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873–922.

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

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*, 333–367.

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin and Review, 13*, 626–635.

Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology, 60*, 127–157.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review, 9*, 438–481.

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

Rolls, E. T., & Deco, G. (2011). A computational neuroscience approach to schizophrenia and its onset. *Neuroscience & Biobehavioral Reviews, 35*, 1644–1653.

Schall, J. D. (2001). Neural basis of deciding, choosing and acting. *Nature Reviews Neuroscience, 2*, 33–42.

Schall, J. D. (2004). On building a bridge between brain and behavior. *Annual Review of Psychology, 55*, 23–50.

Schall, J. D., Purcell, B. A., Heitz, R. P., Logan, G. D., & Palmeri, T. J. (2011). Neural mechanisms of saccade target selection: Gated accumulator model of visual-motor cascade. *European Journal of Neuroscience, 33*, 1991–2002.

Schurger, A., Sitt, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences, 109*, 2904–2913.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464.

Servant, M., Tillman, G., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2019). Neurally constrained modeling of speed-accuracy tradeoff during visual search: Gated accumulation of modulated evidence. *Journal of Neurophysiology, 121*, 1300–1314.

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology, 86*(4), 1916–1936.

Shen, J., & Palmeri, T. J. (2016). Modeling individual differences in visual categorization. *Visual Cognition, 24*, 260–283.

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.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neuroscience, 27*, 161–168.

Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. Methods, models, and conceptual issuesIn D. Scarborough & S. Sternberg (Eds.), *Invitation to cognitive science* (Vol. 4, pp. 703–863). MIT Press.

Sun, R. (Ed.). (2008). *The Cambridge handbook of computational psychology*. Cambridge University Press.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science, 331*(6022), 1279–1285.

Treue, S., & Maunsell, J. H. R. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *Journal of Neuroscience, 19*, 7591–7602.

Tuerlinckx, F. (2004). The efficient computation of the cumulative distribution and probability density functions in the diffusion model. *Behavior Research Methods, Instruments & Computers, 36*(4), 702–716.

Turner, B. M., Forstmann, B. U., Love, B. C., Palmeri, T. J., & Van Maanen, L. (2017). Approaches to analysis in model-based cognitive neuroscience. *Journal of Mathematical Psychology, 76*, 65–79.

Turner, B. M., Forstmann, B. U., Wagenmakers, E. J., Brown, S. D., Sederberg, P. B., & Steyvers, M. (2013). A Bayesian framework for simultaneously modeling neural and behavioral data. *NeuroImage, 72*, 193–206.

Turner, B. M., van Maanen, L., & Forstmann, B. U. (2015). Informing cognitive abstractions through neuroimaging: The neural drift diffusion model. *Psychological Review, 122*, 312–336.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*, 550–592.

Wagenmakers, E. J., Ratcliff, R., Gomez, P., & Iverson, G. J. (2004). Assessing model mimicry using the parametric bootstrap. *Journal of Mathematical Psychology, 48*(1), 28–50.

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review, 7*(3), 424–465.

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58*, 140–159.

Wang, X.-J. (1999). Synaptic basis of cortical persistent activity: The importance of NMDA receptors to working Mmmory. *Journal of Neuroscience, 19*(21), 9587–9603.

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron, 36*, 955–968.

Wang, X.-J. (2008). Decision making in recurrent neuronal circuits. *Neuron, 60*, 215–234.

Wiecki, T. V., Poland, J., & Frank, M. J. (2015). Model-based cognitive neuroscience approaches to computational psychiatry: Clustering and classification. *Clinical Psychological Science, 3*(3), 378–379.

White, C. N., Mumford, J. A., & Poldrack, R. A. (2012). Perceptual criteria in the human brain. *Journal of Neuroscience, 32*(47), 16716–16724.

Wong, K. F., Huk, A. C., Shadlen, M. N., & Wang, X.-J. (2007). Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. *Frontiers in Computational Neuroscience, 1*, 1–11.

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

Zandbelt, B. B., Purcell, B. A., Palmeri, T. J., Logan, G. D., & Schall, J. D. (2014). Response times from ensembles of accumulators. *Proceedings of the National Academy of Sciences, 111*(7), 2848–2853.