Published in final edited form as:

Annu Rev Psychol. 2016 ; 67: 641–666. doi:10.1146/annurev-psych-122414-033645.

Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions

B.U. Forstmann[1], **R. Ratcliff**[2], and **E.-J. Wagenmakers**[3]

B.U. Forstmann: buforstmann@gmail.com

[1]Amsterdam Brain and Cognition Center, University of Amsterdam, 1018 WS Amsterdam, The Netherlands [2]Department of Psychology, Ohio State University, Columbus, Ohio 43210 [3]Department of Methodology, University of Amsterdam, 1018 WV Amsterdam, The Netherlands

Abstract

Sequential sampling models assume that people make speeded decisions by gradually accumulating noisy information until a threshold of evidence is reached. In cognitive science, one such model—the diffusion decision model—is now regularly used to decompose task performance into underlying processes such as the quality of information processing, response caution, and a priori bias. In the cognitive neurosciences, the diffusion decision model has recently been adopted as a quantitative tool to study the neural basis of decision making under time pressure. We present a selective overview of several recent applications and extensions of the diffusion decision model in the cognitive neurosciences.

Keywords

diffusion decision model; information accumulation; decision making; response time; speed-accuracy trade-off; drift rate

INTRODUCTION

Every day, people make thousands of small decisions. Many of these decisions are trivial (e.g., what pair of socks to wear or what TV series to watch), many are to some degree automatic (e.g., how to greet your colleague in the morning or what word to type next in an email), but all of them are made under time pressure. One simply cannot take hours to ponder over what pair of socks to wear or how to greet a colleague: After some deliberation, a decision needs to be made based on the data at hand. Consequently, most real-life decisions are composed of two separate decisions: first the decision to stop deliberating and act, and then the decision or act itself.

The decision to stop deliberating and act is not straightforward, because it involves a balance between two opposing forces. On the one hand, the quality of decision making improves

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

when it is based on more information; on the other hand, decisions are only acceptable when they are timely. In psychology, this balance is known as the speed-accuracy trade-off, a trade-off that affects basketball players, honeybees, and even acellular organisms such as slime molds (Latty & Beekman 2011).

Several models have been developed to account for the speed-accuracy trade-off and explain how people and animals make decisions under time pressure. The most popular class of models assumes that the decision maker accumulates noisy samples of information from the environment until a threshold of evidence is reached. Such accumulation-to-threshold models are known as sequential sampling models.

Sequential sampling models have been developed in mathematical psychology ever since the 1960s (e.g., Stone 1960). Over the course of several decades, researchers began to understand the benchmark phenomena that underlie decision making under speed stress, and the models became increasingly sophisticated to account for these findings (Luce 1986, Townsend & Ashby 1983). In the early 1990s, it became clear that one particular sequential sampling model—the diffusion decision model (DDM)—stood out as the effective standard model (see sidebar Fitting the Diffusion Decision Model to Data) in the field.

FITTING THE DIFFUSION DECISION MODEL TO DATA

In recent years, three dedicated software packages for fitting the full DDM have become publicly available: DMAT (Vandekerckhove & Tuerlinckx 2007, 2008), fast-dm (Voss & Voss 2007, 2008), and the nonhierarchical HDDM (Wiecki et al. 2013). A reduced diffusion model, EZ, can be fit using code provided by Wagenmakers et al. (2007). Wabersich & Vandekerckhove (2014) added a DDM distribution routine to the Bayesian Markov chain Monte Carlo software JAGS. These packages have been implemented in different systems, namely DMAT in MATLAB, fast-dm as a stand-alone C program (precompiled for Windows but running on Linux), and HDDM in Python. Ratcliff & Childers (2015) performed an extensive comparison of the methods. The public availability of software to fit the DDM to data has greatly contributed to the model's popularity and use in practical research settings.

Even though the DDM had been successful as a mathematical process model that accounted for the speed and accuracy of decision making under a wide variety of circumstances, initially its domain of application remained relatively limited. Around the turn of the century, this state of affairs changed radically when it became apparent that the DDM not only accounted for observed behavior but also provided an explanation for some of the general dynamics of single-cell firing rates in monkeys. In the following years, neuroscientists have applied and extended the DDM, and presently it provides a point of departure for many modeling attempts in both low-level and high-level cognitive neuroscience.

As we intend to demonstrate, the ever-increasing interest in applying and extending the DDM in the domain of speeded decision making is motivated by the growing realization that

a quantitative approach can greatly help guide empirical work and deepen our understanding of cognition (Forstmann & Wagenmakers 2015).

The outline of this article is as follows. The first section provides historical context and outlines the current standard form of the DDM; the second section lists the advantages of using the DDM for both experimental psychology and cognitive neuroscience; the third section provides an overview of DDM applications in cognitive neuroscience, focusing on low-level neural firing rates in monkeys and on high-level brain imaging techniques in humans; and the fourth section outlines recent extensions and exciting new developments. A brief summary of the most important points concludes the article.

SEQUENTIAL SAMPLING MODELS

People often need to make decisions based on information that unfolds over time. An example of this is the idealized work process of a police detective solving a homicide: Following a state of confusion and uncertainty, informative cues become available over time that allow the detective to reduce the uncertainty and hopefully solve the case. However, the decision-making process can be sequential even when all of the information is immediately available. For instance, a chess player contemplating a particular move has all the information available, in the sense that the environment will not offer any more cues as time progresses: All the information is contained in the configuration of the pieces on the board, which can be perceptually encoded at a glance. The problem for the chess player is that mental capacity is limited, and the relevant information can only be extracted and processed piecemeal. Hence, the sequential nature of decision making is a fundamental property of the human nervous system, reflecting its inability to process information instantaneously.

To understand the dynamics of decision making, most studies focus on simple, repeatable choice problems with just two alternatives. For instance, participants in lexical decision are confronted with letter strings that have to be classified as a word (e.g., mango) or a nonword (e.g., drapa); participants in the moving dots task are confronted with a random dot kinematogram (RDK) and have to judge whether a subset of dots move to the left or to the right. The elementary nature of these tasks makes it possible to collect thousands of decisions for a single participant in a single session, providing rich data for modeling. Traditionally, the measures of interest are the response times (RTs) for correct responses and for error responses, the distributions of RTs, and the proportion of correct responses. Note that the simplicity of the tasks does not preclude errors; when participants are instructed to respond quickly, errors inevitably arise, and a participant may well classify drapa as a word.

The data from these elementary decision-making tasks reveal several law-like patterns that any model of decision making should try to account for. Some of these law-like patterns are trivial (e.g., mean RT is shorter for easy stimuli than it is for hard stimuli; increasing speed stress shortens mean RT but increases the proportion of errors), but others are not. For instance, (*a*) mean RT is proportional to RT standard deviation (Wagenmakers & Brown 2007); (*b*) manipulations that increase the speed of correct responses also increase the speed of error responses; (*c*) RT distributions are right-skewed, and this skew increases with task difficulty; and (*d*) for difficult tasks, mean error RT is often slower than mean correct RT, but

this pattern can be reversed by speed stress (e.g., Wagenmakers et al. 2008; for a description of these and other law-like patterns see Brown & Heathcote 2008, Carpenter 2004, Luce 1986, Mulder et al. 2012, Ratcliff 2002, Ratcliff & McKoon 2008, Van Ravenzwaaij et al. 2011).

Sequential sampling models come in various forms. The general idea is that evidence is gradually accumulated and each response (e.g., word/nonword, left/right) is represented by a separate decision boundary. However, the models differ according to whether there are one or two counters and whether the counters are independent; whether they are assumed to be leaky; or whether they exert a top-down influence on the accumulation process. Responses can be determined by an absolute evidence rule (i.e., two fixed thresholds, one for each counter) or a relative evidence rule (one threshold on the difference in accumulated activation; e.g., Bogacz et al. 2006, Ratcliff & Smith 2004, Teodorescu & Usher 2013).

One famous class of sequential sampling models consists of accumulator models (e.g., Van Zandt et al. 2000, Vickers & Lee 1998). A prototypical accumulator model has independent counters and an absolute evidence response rule. Here we focus on a different class of sequential sampling models, which assumes a relative evidence rule: A response is initiated as soon as the difference in evidence accumulated exceeds a prespecified criterion. For discrete evidence accumulation, this account is known as a random walk model; for continuous evidence accumulation, the process is known as a diffusion process.

Interdisciplinary History

The history of random walk models dates back to the early days of probability theory, when much effort was devoted to problems related to gambling. In a famous problem known as the gambler's ruin, two gamblers, A and B, play a sequence of independent games against each other. Each gambler has a (possibly different) starting capital, and every game has a fixed chance p of being won by gambler A. Every time one of the gamblers wins a game, the winner obtains one unit of the other player's capital; the process continues until one of the gamblers is bankrupt (Carazza 1977, Feller 1968). Provided the two starting capitals and the chance p, what is the probability that A goes bankrupt? And what is the expected number of games until this happens? (For answers, readers are referred to Feller 1968.)

The temporal flow of capital in the gambler's ruin process can be represented as a random walk with two absorbing boundaries; whenever a boundary is reached, this signifies that one of the gamblers is bankrupt. Parameter p represents the drift of the process; when $p > 1/2$, the noisy process will tend to drift toward the bound associated with gambler B's bankruptcy.

In the continuous limit of small stakes and with p close to 1/2, the process is known as Brownian motion or Wiener diffusion process. This process was proposed to explain the movement of physical particles influenced by many molecular collisions (e.g., Einstein 1905; for a visual demonstration, see **http://en.wikipedia.org/wiki/Brownian_motion**). The experimental verification of this explanation helped confirm the existence of molecules and atoms, and it earned Jean Perrin the 1926 Nobel Prize in Physics. In mathematics, the

Wiener diffusion process is a prototypical example of a stochastic differential equation (e.g., Smith 2000) with applications in finance, heat flow, and fluid dynamics.

The random walk process is also of considerable interest to statisticians, partly because it is related to sequential analysis; data come in over time and the statistician has to determine when to stop collecting data and make a decision. The sequential analyses initiated by Alan Turing famously helped break the German enigma code, expediting the end of World War II (Good 1979). At about the same time, Abraham Wald proposed the sequential probability ratio test (SPRT; e.g., Wald & Wolfowitz 1948). In the SPRT, each incoming datum is transformed into a log likelihood ratio that quantifies the relative evidence for one hypothesis versus another; the likelihood ratios are added as the data flow in, and the process halts as soon as a predetermined level of evidence is reached. The appeal of this procedure is that it is optimal in the sense that it achieves the fastest mean decision time for a given accuracy (Bogacz et al. 2006). For applications in neuroscience, the optimality of the SPRT is an attractive property, because evolutionary pressures and reinforcement learning mechanisms may have shaped neurons to process information near-optimally, given their innate limitations (Ma et al. 2006).

In psychology, the interest in random walk models started with Stone (1960) and was followed by major contributions from Laming (1968) and Link & Heath (1975). The early models, however, did not account for the relative speed of error RTs across all experimental scenarios. To account for all of the data, the early models needed to be expanded, and this resulted in the model that is the current standard: the DDM.

Current Standard Form: The Diffusion Decision Model

The DDM (e.g., Ratcliff 1978, Ratcliff & McKoon 2008, Voss et al. 2013) assumes that dichotomous decisions are based on the accumulation of noisy evidence, commencing at the starting point and terminating at a decision threshold that is associated with a particular decision or choice. **Figure 1** represents an application of the model to the RDK task. The figure demonstrates that the diffusion process is inherently noisy, causing the choices to be error prone and the RTs to be variable.

The model structure shown in **Figure 1** provides a unified account of the psychological mechanisms that underlie both RTs and the probabilities with which responses are chosen. The model has four key parameters. First, drift rate represents the average amount of evidence accumulated per unit time, and it is an index of task difficulty or subject ability. Second, boundary separation represents the level of caution; increasing boundary separation results in fewer errors (because of the reduced impact of the within-trial diffusion noise) but at the cost of slower responding. Hence, boundary separation implements the speed-accuracy trade-off. Third, the starting point reflects the a priori bias or preference for one or the other choice alternative. Fourth, nondecision time is an additive lag parameter that measures the time for peripheral processes such as encoding a stimulus, transforming the stimulus representation into a decision-related representation, and executing a response. Consequently, the total time for a response is the time to diffuse from starting point to boundary, plus the nondecision time.

In addition to the four key parameters, the DDM features three across-trial variabilities in drift rate, starting point, and nondecision time. Without across-trial variability in any of these, and with boundaries equidistant from the starting point, the model would predict the distributions of RTs for correct and error responses to be identical. However, with across-trial variability in drift rate, the model predicts errors to be slower than correct responses because of probability mixtures of processes with different times and accuracies. For example, for a larger drift rate in the mixture, accuracy will be higher and RT slower for both correct and error responses; for a lower drift rate, accuracy will be lower and RTs longer. Therefore, there will be a smaller number of fast errors and a higher number of slow errors from this mixture (relative to correct responses), leading to slower errors relative to correct responses (for further explanation, see Ratcliff 1978; Ratcliff & McKoon 2008, figure 4). With across-trial variability in the starting point, the model predicts errors to be faster than correct responses (e.g., Ratcliff & Rouder 1998, figure 2; Wagenmakers et al. 2008). When a process starts near the correct response boundary there will be few errors and they will be slow, because the process has to travel a long distance to reach the error boundary. When a process starts near the error boundary there will be more errors and they will be fast, because the process only has to travel a short distance to reach the error boundary. Both patterns are found in the data (with a crossover sometimes so that fast errors occur in easy conditions and slow errors in difficult conditions), and the model explains why these patterns occur. A mixture of starting points gives fast errors overall.

Equipped with these parameters, the DDM provides an excellent account of the law-like patterns observed across virtually all speeded RT tasks. For instance, the DDM accounts for the relationship between mean RTs and the probabilities of the choices (errors and correct responses), including how these covary with stimulus difficulty. In addition, the DDM accounts for the shapes of RT distributions and for how these change as a function of experimental conditions.

Based on a superficial analysis, a skeptic may argue that the excellent fit of the DDM is achieved partly because its parameters make it highly flexible; in other words, the model may be overparameterized and immune to empirical falsification. A deeper analysis, however, shows this concern to be without merit (Wagenmakers 2009). First, experimental designs usually feature multiple conditions, and this allows the model parameters to be constrained in meaningful ways; for instance, only the drift rate parameter is allowed to vary with stimulus difficulty, and only the boundary separation parameter is allowed to vary across conditions with different levels of speed stress. Such constraints severely limit the model's flexibility. Second, Ratcliff (2002) has demonstrated by simulations how the DDM fails to account for fake but plausible data patterns. Third, typical experimental designs allow the model parameters to be adequately recovered. Finally, tests of specific influence show that manipulations of psychological processes affect the associated parameters; for instance, a manipulation of task difficulty mostly affects the drift rate parameter, and a manipulation of speed stress mostly affects the boundary separation parameter (e.g., Voss et al. 2004).

ADVANTAGES OF THE DIFFUSION DECISION MODEL

The DDM can be considered a dynamic version of signal-detection theory (SDT; Gold & Shadlen 2007, Ratcliff 1978, Wagenmakers et al. 2007). In SDT, the decision maker is assumed to assess the diagnosticity of a single sample of information; in the DDM, the decision maker draws a sequence of samples, adding their diagnostic values until a threshold amount of evidence is reached. Thus, like SDT, the DDM allows one to disentangle estimates of ability (i.e., that which is not under strategic control: *d'* in SDT versus drift rate in DDM) from estimates of criteria settings (i.e., that which is under strategic control: *c* in SDT versus boundary separation and starting point in DDM). Unlike SDT, however, the DDM considers not only response proportion but also the shapes of RT distributions, both for errors and correct responses. This has the advantage of finding invariance in evidence used in a decision when speed or accuracy settings are manipulated; in contrast, SDT analyses find that evidence changes when accuracy is stressed.

Below we discuss the advantages that the DDM has to offer, both for the analysis of choice behavior and for cognitive neuroscience.

Advantages for the Analysis of Choice Behavior

At its core, a DDM analysis allows researchers to decompose observed choice behavior into its constituent cognitive processes. By simultaneously taking into account both response accuracy and response latency, the model addresses the speed-accuracy trade-off and allows an assessment of individual ability that is not contaminated by differences in threshold settings or in the speed of peripheral processes unrelated to the decision-making process itself (Wagenmakers et al. 2007). This decomposition makes it possible to evaluate the adequacy of verbal theories such as the global slowing hypothesis of aging, which effectively states that the effect of aging is to decrease drift rate. The decomposition also facilitates the use of the DDM as a cognitive psychometric tool (Riefer et al. 2002, Vandekerckhove et al. 2011) to pinpoint the cognitive processes that are dysfunctional in clinical populations, for instance, in patients with aphasia, hypoglycemia, dysphoria, attention deficit hyperactivity disorder (ADHD), dyslexia, and anxiety disorders.

One criterion for a model's usefulness is whether it does more than simply reiterate what can be obtained from traditional analyses. Below we describe a number of recent DDM applications and highlight how some of these have provided new insights into cognition, individual differences, and differences among subject groups. In other cases, however, the model only provides obvious results. But even in this case, the model still integrates the three dependent variables—accuracy and the shapes of correct and error RT distributions—into a common theoretical framework that provides a mechanistic explanation for the observed data. This stands in contrast to the commonly used hypothesis-testing approaches that are mute on the psychological processes that produce behavior, usually focusing only on accuracy or mean correct RT as the dependent variable. In some cases, separate statistical analyses of each variable tell the same empirical story, but in other cases they are inconsistent. A model-based approach helps resolve such inconsistencies.

Aging—The application of the diffusion model to studies of aging has been especially successful, producing a novel view of the effects of aging on cognition. The general finding in the literature was that older adults are slower than young adults (but not necessarily less accurate) on most tasks, and this has been interpreted as a decline with age in all or almost all cognitive processes. However, application of the DDM showed that this interpretation is generally not correct (Ratcliff et al. 2007b, and references therein). For example, Ratcliff et al. (2010) tested old and young adults on numerosity discrimination, lexical decision, and recognition memory. What they found is that older adults had slower nondecision times and set wider boundaries, but their drift rates were not lower than those of young adults. In contrast, large age-related declines in drift rate have been found in other tasks, such as associative recognition and letter discrimination (Ratcliff et al. 2011, Thapar et al. 2003).

Working memory and IQ—Schmiedek et al. (2007) analyzed data from eight choice RT tasks (including verbal, numerical, and spatial tasks) from Oberauer et al. (2003). They found that drift rates in the diffusion model mapped onto working memory, speed of processing, and reasoning ability measures (all measured by aggregated performance on several tasks).

Similarly, the DDM analyses from Ratcliff et al. (2010, 2011) showed that drift rate varied with IQ (by as much as 2:1 for higher versus lower IQ participants), but boundary separation and nondecision time did not. Note that this is the opposite of the pattern for aging.

Clinical populations—Research on psychopathology and clinical populations commonly uses two-choice tasks to investigate processing differences between patients and healthy controls. For instance, highly anxious individuals show enhanced processing of threat-provoking materials, a pattern that is found reliably only when two or more stimuli are competing for processing resources. White et al. (2010) recently challenged this resource competition account. They conducted a lexical decision experiment with single words (i.e., without resource competition) that included threatening and control words; using a DDM decomposition, they found a consistent processing advantage for threatening words in highly anxious individuals, whereas traditional comparisons showed no significant differences. Because the diffusion model makes use of both RT and accuracy data, it can better detect differences among subject populations than RT or accuracy alone.

In a similar vein, studies of depression have sometimes found mixed patterns of results. In general, depressive symptoms are closely linked with abnormal emotional processing: Whereas nondepressed people have a positive emotional bias, clinical depression is accompanied by a negative emotional bias, and dysphoria is accompanied by evenhandedness (i.e., no emotional bias). However, studies using item recognition and lexical decision tasks often fail to produce significant results. White et al. (2009) used the DDM to examine emotional processing in dysphoric and nondysphoric college students to examine differences in memory and lexical processing of positive and negative emotional words, which were presented among many neutral filler words. They found positive emotional bias in drift rates for nondysphoric subjects and evenhandedness for dysphoric subjects. As before, this pattern was not apparent with comparisons of RTs or accuracy, consistent with previous null findings.

Another study examined the effects of aphasia in a lexical decision task for which the neurolinguistic patients showed exaggerated RTs. A DDM decomposition revealed that both decision and nondecision processes were compromised, but the quality of information processing (i.e., drift rate) did not differ much between patients and controls (Ratcliff et al. 2004a). It is unclear how a traditional statistical analysis could have arrived at a similar insight.

Miscellaneous—Ratcliff & Van Dongen (2009) looked at effects of sleep deprivation using a numerosity discrimination task; Van Ravenzwaaij et al. (2012) looked at the effects of alcohol consumption using a lexical decision task; and Geddes et al. (2010) looked at the effects of reduced blood sugar using a numerosity discrimination task. The main result of all of these studies was a reduced drift rate but with either small or no effect on boundary separation and nondecision time.

By way of contrast, in studies of cognitive development, younger children show larger boundary separation and longer nondecision times than do older children (Ratcliff et al. 2012). Other experiments have found drift rates to be lower for ADHD and dyslexic children relative to normal controls (for ADHD, see Mulder et al. 2010; for dyslexia, see Zeguers et al. 2011). The above applications demonstrate how a comprehensive DDM decomposition of observed choice behavior yields deeper conclusions and insights than traditional methods of analysis.

Advantages for Cognitive Neuroscience

The DDM advantages are particularly acute for the field of model-based cognitive neuroscience, a nascent discipline that combines insights and measurement tools from experimental psychology, mathematical psychology, and neuroscience (Forstmann et al. 2011, Forstmann & Wagenmakers 2015). **Figure 2** shows each discipline's primary concern with the cognitive process and how formal models can act as a hub that connects the contributions of the separate disciplines.

A DDM decomposition allows cognitive neuroscientists to associate brain measurements with specific cognitive processes instead of behavioral data. This comes with a number of advantages. First, the DDM decomposition can confirm that a particular manipulation is selective or processpure; for instance, a manipulation of task difficulty is process-pure when it selectively affects drift rate. When task difficulty is manipulated across blocks, however, other processes such as boundary separation could also be affected. Second, even when a manipulation is not process-pure, a DDM decomposition allows the researcher to isolate and focus on the contribution of the process of interest. In the example above, brain measurements (e.g., fMRI contrasts) may reflect the impact of changes in both drift rate and boundary separation; yet, the DDM parameter estimates can be used to disentangle the joint impact of the two processes and to associate the brain measurements with the process of interest (e.g., drift rate). Finally, the DDM decomposition facilitates an individual differences analysis; for instance, people with relatively large changes in drift rate may show more pronounced activation in frontoparietal network areas, suggesting that these areas are important for stimulus processing.

In addition, as we will outline below in more detail, the DDM has stimulated the development of quantitative models for neural processes. This work suggests that the processes that drive observed choice behavior are qualitatively similar to those that describe the behavior of individual neurons.

THE DIFFUSION DECISION MODEL IN COGNITIVE NEUROSCIENCE: APPLICATION

To bridge the gap between neural process and observed choice behavior, it is helpful to model and estimate the intermediate latent psychological processes. The DDM constitutes an important general framework to understand how neurons process information and how brain activation gives rise to choice and action. Nevertheless, there remains a vast divide between neurons and choice, and one of the main unsolved challenges is to provide a unified account of both low-level and high-level brain processes and of how these determine choice behavior.

Below we first discuss the application of the DDM in low-level neuroscience and single-cell recordings in monkeys, and then turn to the application of the DDM in high-level neuroscience and brain measurements in humans. Due to space limitations, our review is necessarily selective.

Application in Low-Level Cognitive Neuroscience: Neural Firing Rates and Single-Cell Recordings in Monkeys

One of the main reasons for the current popularity of diffusion models in neuroscience is the possibility to observe the behavior of single neurons of monkeys (and occasionally rats) performing simple decision-making tasks such as the RDK. Hanes & Schall (1996) made one of the first connections between theory and single-cell recording data, which was subsequently taken up in work by Shadlen and colleagues (e.g., Gold & Shadlen 2001). As shown in **Figure 3**, the key finding is that the firing rates of single cells in decision-related areas increase to a maximum that is independent of both the speed and the difficulty of the decision. These decision-related areas include the lateral intraparietal cortex (LIP; see Roitman & Shadlen 2002, Shadlen & Newsome 1996), the frontal eye field (FEF; see Ferrera et al. 2009, Hanes & Schall 1996), and other parts of the prefrontal cortex and the superior colliculus (SC; Ding & Gold 2012; Horwitz & Newsome 1999; Ratcliff et al. 2003a, 2007a). These results dovetail nicely with models that assume gradual accumulation of evidence up to a fixed decision criterion.

There is debate about where exactly the accumulation takes place, but it is clear that (at least) LIP, FEF, and SC form part of a circuit that is involved in implementing oculomotor decisions in monkeys performing simple decision-making tasks. The above studies generally support the notion that decision-related information flows from LIP to FEF and then to SC just prior to a decision.

This domain benefits from an abundance of recent high-quality reviews (Ding & Gold 2012; Glimcher 2003; Gold & Shadlen 2001, 2007; Schall 2001, 2013; Shadlen & Kiani 2013) that show a variety of approaches but mainly focus on the accumulation of evidence up to a

decision criterion. In addition, a number of articles present explicit neurobiological models that assume that evidence is gradually accumulated over time; here, evidence is conceptualized as activity in populations of neurons associated with a specific choice alternative (Boucher et al. 2007; Ditterich 2006; Gold & Shadlen 2001, 2007; Hanes & Schall 1996; Platt & Glimcher 1999; Purcell et al. 2010; Ratcliff et al. 2003a; Roitman & Shadlen 2002; Shadlen & Newsome 2001).

The proposed neurobiological models fall into several classes: Some models assume accumulation of a single evidence quantity that can take on positive and negative values (as is consistent with the DDM; Gold & Shadlen 2000, 2001; Ratcliff 1978; Ratcliff et al. 1999, 2003a; Smith 2000); others assume that evidence is accumulated in separate accumulators corresponding to separate choice alternatives (Churchland et al. 2008, Ditterich 2006, Ratcliff et al. 2007a, Usher & McClelland 2001). In this latter class of models, accumulation in separate accumulators can be independent or interactive—so that as evidence grows in one accumulator, it inhibits evidence accumulation in the other accumulator. These two classes of models largely mimic each other at a behavioral level (Ratcliff 2006, Ratcliff & Smith 2004).

In one innovative application, Purcell et al. (2010) used real neural firing rate data as the input for a range of different sequential sampling models of decision making in a visual search task. The models they examined involved independent accumulation models with decay, inhibition, and a gating mechanism (i.e., activity had to be greater than some base level to be involved in the decision). The modeling results revealed that models that included decay or gating provided an excellent account of the observed RT distributions. In the model proposed by Purcell et al. (2010), the stimulus is directly tied to the decision-making mechanism without the involvement of an intermediate short-term memory representation. For highly overtrained monkeys, this is likely appropriate.

E pluribus unum—The modeling efforts for single neurons raise an important question: If single neurons act as noisy evidence accumulators, how does this determine the behavior of a large pool of neurons? In other words, do the properties of an individual neuron scale up to determine the dynamics of the neural population?

This question was recently addressed by Zandbelt et al. (2014), who examined a number of models in which individual neurons act as single redundant accumulators that together constitute a neural ensemble. The decision rule is that when some proportion of neurons from the ensemble have reached their criterion, the decision is made. They found that, under general conditions, the behavior of such a system was relatively insensitive to ensemble size. This suggests that a single diffusion process (used in modeling at the behavioral level) might be implemented in hardware as a combination of multiple accumulators.

Another attempt to bridge the gap from neuron to ensemble comes from modeling efforts that relate diffusion models to models based on spiking neurons (e.g., Deco et al. 2013, Roxin & Ledberg 2008, Wong & Wang 2006). Roxin and Ledberg examined models in which separate populations of spiking neurons are assumed to represent the two choices. They show that such models can be reduced to a one-dimensional model that is similar but

not identical to the standard DDM (their model involves nonlinearity). Wong and Wang present a spiking neuron model and then reduce it to a two-variable model with self-excitation and inhibition. Their approximation is similar to the leaky competing accumulator model (Usher & McClelland 2001). Wang's modeling approach has had a wide range of applications (Wang 2008). Unfortunately, the Wang model is relatively complex, and at this point it is not possible to use it to fit data. However, its strength is that it takes seriously the relationship between neural processes, including synaptic currents, the behavior of neurotransmitters, membrane voltages, etc.

Smith (2010) suggested a different approach. He made an explicit connection between diffusion processes at a macro behavioral level and Poisson shot noise processes at a slightly abstract neural level. The shot noise process describes the cumulative effects of time-varying events (i.e., action potentials) that arrive according to a Poisson process. Smith showed that the time integral of such Poisson shot noise pairs follows an integrated Ornstein-Uhlenbeck process, whose long timescale statistics are very similar to those assumed in the standard DDM.

Single-cell recordings and bias effects—In a two-alternative task, bias toward one or the other alternative can be induced by instructions, by varying the relative proportions of occurrence, or by asymmetric rewards. Such bias can be modeled within the DDM framework in two ways. First, the starting point of the process can be moved nearer the boundary that represents the preferable alternative. Alternatively, the zero point of the drift rate (i.e., the drift criterion) can be altered by increasing drift rates toward the preferred boundary and decreasing drift rates to the nonpreferred boundary (Leite & Ratcliff 2011, Mulder et al. 2012, Ratcliff 1985, Ratcliff & McKoon 2008, Ratcliff & Smith 2004, Starns et al. 2012, Wagenmakers et al. 2007).

In human decision making it is clear that changing the relative proportion of occurrence brings about a change in the starting point of the decision process. This is evident from changes not only in accuracy and mean RT (Mulder et al. 2012), but also in the shape of RT distributions. If the starting point moves nearer one boundary, the RT distribution for that response shifts to lower values. In contrast, if the drift criterion changes, the leading edge of the RT distribution does not change very much (see Ratcliff & McKoon 2008 for a detailed discussion).

Hanks et al. (2011) presented single-cell recording data and human data from the motion discrimination task. They used their conclusions from the human data to support a drift criterion interpretation, but they did not examine RT distributions nor perform the critical test. Their results conflict with the conclusions from Ratcliff & McKoon (2008), who presented data from a bias manipulation in the motion discrimination task and found strong evidence for a change in starting point. For the monkey data, the initial firing rates differed as a function of bias, with an increase in firing rate for neurons corresponding to the more likely decision. Hanks et al. interpreted this as a change in drift rate, but it could also be the way the system changes the starting point of activity in the decision process. Again, explicit modeling of accuracy and RT distributions would make this finding clear.

Single-cell recordings and sequential dependencies—Gold et al. (2008) examined sequential dependencies in a motion discrimination task and found evidence for changes in LIP neuron firing rates, which they interpreted as changes in the drift criterion. In human data, Ratcliff et al. (1999) found that sequential effects were best modeled as changes both in the starting point and in the drift criterion. Again, explicit modeling of the behavioral data would clarify the interpretation of this finding.

Single-cell recordings and trade-off effects—Two recent studies have attempted to manipulate speed-accuracy settings in monkeys (Hanks et al. 2014, Heitz & Schall 2012; see Cassey et al. 2014 for a critique). One immediate problem is that it is extremely difficult to get monkeys to exercise caution and to slow down responding based on rewards in the same way as humans do. This means that monkey must be trained to delay, something that Heitz & Schall accomplished using explicit deadline cutoffs. In the Hanks et al. study, the monkeys naturally produced fast responses; to make the total time of the trial the same for fast and slow responses, the monkeys had to be trained to respond more slowly by using time delays following the response. To move the monkeys back to a speed regime, one of them needed an additional manipulation in which stimulus presentation duration was reduced.

Perhaps the main lesson from this experiment is that it is difficult to get monkeys to adopt different speed-accuracy regimes (see also Cassey et al. 2014). This contrasts with humans: Young adults find it easy to adopt different regimes, whereas older adults require some training, but once reassured that fast responses are acceptable or even desirable, they are able to switch between regimes on a block-by-block basis (Ratcliff et al. 2001, 2003b, 2004).

Surprisingly, both the Hanks et al. (2014) and Heitz & Schall (2012) studies found that the firing rate threshold did not change with speed-accuracy instructions, a finding that seems to contradict the results of human studies that suggest a change in decision boundaries. However, the results for monkeys showed changes in the starting level of activity, with a higher level for the speed regime. This is consistent with a dual racing accumulator model in which changes in starting point are identical to changes in boundary settings (Forstmann et al. 2008, Ratcliff & Smith 2004).

Heitz & Schall (2012) also found a reduction of peak activity in the accuracy regime relative to the speed regime. Hanks et al. (2004) proposed a model in which a boost is added to the drift rate in a diffusion model in the speed regime. However, there is the question of whether monkeys and humans perform the tasks in the same way. This can be examined using RT distributions. In the Heitz & Schall study, the RT distributions for the speed and accuracy regimes hardly overlap. This is inconsistent with most human studies that use instructions (and not time deadlines or signals; e.g., Ratcliff 1988, 2006), in which RT distributions overlap to a great degree (Ratcliff et al. 2001, 2003b, 2004b). The relationship between speed and accuracy manipulations in humans and monkeys is not yet settled (Cassey et al. 2014).

In general, much work in the animal area is limited by the lack of explicit modeling of behavioral data featuring a thorough quantitative analysis of accuracy as well as the shapes

of RT distributions for correct and error responses. If sequential sampling models such as the DDM were fit to the data, this would increase the confidence in the theoretical link between behavior and neural processes (e.g., Purcell et al. 2010, Ratcliff et al. 2007b).

In sum, the application of sequential sampling models in the arena of low-level cognitive neuro-science opens up exciting new prospects. Instead of considering only behavioral data, researchers could test the models on additional findings such as neural firing rates. These data add useful constraints and allow a deeper understanding of the computational mechanisms that ultimately produce overt decisions. The work in this new area can be improved further by rigorous modeling of the behavioral data to confirm the validity of more qualitative conclusions.

Application in High-Level Cognitive Neuroscience: Measuring Human Brain Activity

The application of sequential sampling models in low-level cognitive neuroscience comes with several challenges. For instance, it can be unclear whether monkeys and human carry out an experimental task in the same way (Hawkins et al. 2015), compromising the extent to which neural firing rate results in monkeys generalize to humans. Furthermore, neural firing rates are measured in a select subset of neural structures, making it difficult to assess the network dynamics among larger structures such as frontal cortex, premotor cortex, and the basal ganglia. These challenges can be addressed by using methods from high-level cognitive neuroscience.

Magnetic resonance imaging—In recent years, studies using functional magnetic resonance imaging (fMRI) have started to correlate parameter estimates from sequential sampling models to the blood-oxygen-level dependent signal obtained from fMRI experiments in perceptual decision making. **Figure 4** summarizes the results of these efforts. The summary includes results from seven fMRI studies focusing on evidence accumulation, two studies on decision thresholds, five studies on starting point bias, one study on drift rate bias, and one study on nondecision time (Mulder et al. 2014).

Figure 4 shows the relevance of a large variety of brain areas. Several global patterns emerge. First, individual differences in the accumulation of evidence are mainly associated with regions belonging to the frontoparietal network (top row of **Figure 4**). Second, individual differences in adjusting response thresholds are associated with a frontobasal ganglia network. Third, a more complex pattern arises for choice bias, which is associated with individual differences in both the frontoparietal and the frontobasal ganglia networks. There is only weak evidence for the involvement of brain regions in individual differences in nondecision time.

Electroencephalography—A growing number of studies has started to use sequential sampling models in combination with human neurophysiology measurement techniques such as electroen-cephalography (EEG). The main advantage of these techniques is their high temporal resolution, an advantage that is particularly pronounced for the study of speeded RT tasks.

Philiastides et al. (2006) used multivariate pattern analysis to derive spatiotemporal profiles of activity that could discriminate between relevant stimulus categories (i.e., face versus car) and between different levels of difficulty (i.e., image phase coherence). The results revealed an early (170 ms) and a late (300 ms) event-related potential (ERP) component that were predictive of decision accuracy. In a later study, Ratcliff et al. (2009) showed that within each stimulus coherence level, higher late-component amplitudes were associated with higher DDM drift rates. Hence, this study demonstrated that, for nominally identical stimuli, the amplitude of a single-trial EEG component can be used to measure and predict the quality of information processing.

Other model-based EEG work has tried to elucidate the temporal dynamics of decision making. Van Vugt et al. (2012) employed an EEG experiment to disentangle stimulus- and response-locked processes using an RDK task. They applied a general linear model comparable to event-related fMRI designs including a set of stimulus- and response-locked regressors. Their results revealed spectral changes primarily in the theta band (4–8 Hz), a frequency band associated with cognitive control processes (Cohen 2014). Importantly, changes in the theta band matched the dynamics (i.e., the ramping temporal profiles) of evidence accumulation during the decision process. These results are broadly consistent with recent work showing that weighting discrete stimuli presented in a series and as an input to the accumulation process fluctuated in accordance with delta band (1–3 Hz) oscillations (Wyart et al. 2012).

In another EEG study, Cavanagh et al. (2011) used theta power to quantify trial-to-trial fluctuations in activation of the medial prefrontal cortex (mPFC). They found that an increase in activation of the mPFC—a brain structure thought to be involved in effortful control over behavior—was associated with an increase in the DDM boundary separation parameter. They argued that mPFC signals response conflict and acts in concert with structures in the basal ganglia to increase the response threshold, slowing down response execution and hence creating more time for information accumulation.

Similarly, Boehm et al. (2014) had participants perform an RDK task either under speed stress or under accuracy stress. Trial-by-trial fluctuations in the adjustment of response thresholds under speed stress correlated with single-trial amplitudes of the contingent negative variation (CNV), a slow cortical potential that occurs whenever a stimulus prompts a participant to perform a task. Based on their results, Boehm et al. (2014) concluded that the CNV might reflect adjustments of response caution, which serve to prepare the system for action and facilitate quick decision making.

Taking a different approach, Bode et al. (2012) examined how neural activity preceding the stimulus affects the later decision process. They used a multivariate pattern classification approach to decode choice outcomes in a perceptual decision task from spatially and temporally distributed patterns of brain signals. Interestingly, in addition to decoding choice outcomes based on pre- and poststimulus activity, the authors were able to show that the past history of choices primed the decision process on subsequent trials. More concretely, a DDM decomposition revealed that the starting point of the evidence accumulation process was shifted toward the previous choice, thereby biasing the choice process.

In sum, recent work in high-level cognitive neuroscience has employed the DDM decomposition methodology and related the estimated parameters to brain measurements involving fMRI and EEG; these applications have been varied, concerning individual differences, theta power for specific brain structures, and more generally the construction of spatiotemporal profiles of brain activity. We expect this area of research to continue its ongoing expansion in the near future.

THE DIFFUSION DECISION MODEL IN COGNITIVE NEUROSCIENCE: EXTENSIONS

In this section, we show how the basic framework of the DDM has recently been extended to account for more complicated phenomena in decision making and their neural underpinnings.

Extension to Multi-Alternative Decisions

There is a developing interest in multi-alternative decision-making paradigms such as those concerning visual search (Basso & Wurtz 1998, Purcell et al. 2010), motion discrimination (Ditterich 2010, Niwa & Ditterich 2008), and other more behavioral tasks (Leite & Ratcliff 2010, Ratcliff & Starns 2013). Also, confidence judgments in decision and memory involve multi-alternative decision making (Pleskac & Busemeyer 2010, Ratcliff & Starns 2013, Van Zandt 2002; see also the next section). Many of these approaches compare a variety of competing models, and conclusions about what architectures are most promising are just being reached.

Compared to the present volume of work on two-choice decision making, only a modest amount of research has aimed at modeling both RT and choice proportions in multi-alternative decisions and confidence judgments. It is clear that the two-choice DDM cannot be simply extended to tasks with three or more choice alternatives. However, models with racing accumulators can be naturally extended by adding accumulators for each additional choice. Some models with racing accumulators become standard diffusion models when the number of choices is reduced to two.

Extension to Confidence Judgments

The psychological literature has a long tradition of using confidence judgments to better understand decision making and cognition. Probably the main domain of application of confidence judgments has been memory research (e.g., Egan 1958, Murdock 1974). In this line of research, subjects are often asked to respond on an ordinal many-point scale (e.g., a six-point scale ranging from “very sure” for one choice to “very sure” for the other choice).

In the past there have been several attempts to model the response confidence and response latency jointly (e.g., Murdock & Anderson 1975, Vickers 1979), but recently researchers have proposed more detailed models. Because the confidence choice is an explicit decision, the models have different decision boundaries for each choice.

Sequential sampling models for confidence

In order to model confidence judgments in recognition memory tasks, Ratcliff & Starns (2013) proposed a multiple-choice diffusion decision process with separate accumulators of evidence for the different confidence choices. The accumulator that first reaches its decision boundary determines which choice is made. Five algorithms for accumulating evidence were compared and one of them was successful, in the sense that it produced proportions of responses for each of the choices and full RT distributions for each choice that closely matched empirical data. Within this algorithm, an increase in the evidence in one accumulator is accompanied by a decrease in the others, so that the total amount of evidence in the system is constant. This is one way in which the two-choice DDM can be generalized to multi-alternative decisions (see also Audley & Pike 1965).

Application of the model to experimental data uncovered a relationship between the shapes of z-transformed receiver operating characteristics (z-ROC) and the behavior of RT distributions. For low-proportion choices, the RT distributions were shifted by as much as several hundred milliseconds relative to high-proportion choices. This behavior and the shape of z-ROC functions were both explained in the model by the behavior of the decision boundaries.

Ditterich (2010) argued that behavioral data alone would not be sufficient to discriminate among a number of different multi-alternative models. However, Ratcliff & Starns (2013) applied the decision model to a three-choice motion discrimination task in which one of the alternatives was the correct choice on a low proportion of trials. Like the shifts for the confidence judgment data, the RT distribution for the low-probability alternative was shifted relative to the higher-probability alternatives. The diffusion model with constant evidence accounted for shifts in the RT distribution better than a competing class of models.

Confidence judgments in animals

The animal domain and the human domain have different definitions or measures of confidence. In the animal domain the definition seems to depend on the amount of accumulated evidence: the more the evidence accumulated, the more confident the response. By contrast, in the human domain the measure seems to depend on an explicit choice on a scale, usually among a relatively small number of alternatives. Thus, in the human case, a commitment to a level of confidence is made. Humans find it easy to make decisions on such scales, but it is likely very difficult to train animals to make such judgments. To assess confidence in animals a different kind of task is employed, in which the animals are rewarded for correct choices and are offered smaller rewards if they opt out of the task.

Kepecs et al. (2008) performed an odor discrimination task in which stimuli were mixtures of two odors. Confidence was identified based on distance from the decision boundary and modeling was based only on accuracy. In a delayed version of the task, rats were more likely to move to the next trial without waiting for a reward when the stimulus was more ambiguous. However, this study only reported and modeled accuracy, although there was some discussion of evidence accumulation models.

Kiani et al. (2014) used a more explicit opt-out task in which, on some trials, monkeys could explicitly opt for a lower-value reward rather than risking zero reward. In the experiments, responses were given after a delay so that activity in the LIP neurons was maintained until the signal to respond, at which point a decision criterion was reached. When recordings were made in the area of LIP corresponding to the opt-out responses, there was no strong evidence of accumulation to a criterion for these decisions. This means that decisions can be made both when activity in a neural population reaches a decision criterion and when it does not.

Extension to Value-Based Decision Making

Recent work by Krajbich et al. (2010) set out to understand the role of visual fixation in value-based decision making. In an elegant design, they let people choose between options associated with different subjective values (e.g., a picture of a candy bar versus a picture of an apple) while tracking their eye movements. An extended version of the DDM linking choice preference to eye fixations was fit to the eye fixation data. The results show that this extended DDM in which fixations are involved in the value integration process could provide an excellent fit to the data, providing a new link between fixation and choice data. Interestingly, the duration of fixation was predictive of the choice as well as choice biases.

Extension to Changes of Mind

Resulaj et al. (2009) proposed another interesting extension of the DDM, trying to model how participants change their mind during decision making. Recall that the DDM proposes that a decision is initiated as soon as the accumulated evidence reaches a response boundary. Resulaj et al. challenged this assumption by showing that the simple formalism of the DDM fails to explain what happens when people change their mind. In their study, people had to make a decision about the directionality of a centrally positioned RDK stimulus by moving a handle to a leftward or right-ward response location. Changes in the movement trajectories of the handle revealed that people occasionally changed their mind. The DDM extension of Resulaj et al. allows information that is already in the perceptual processing pipeline to influence and possibly overrule the initial decision.

Extension to Dynamic Thresholds

One of the most popular extensions of the DDM introduces the idea of thresholds or response boundaries that are collapsing instead of fixed (see **Figure 5**). The core idea of this extension is that decisions are based on less and less evidence as time passes; in other words, the decision maker grows increasingly impatient (Bowman et al. 2012, Ditterich 2006, Drugowitsch et al. 2012, Thura et al. 2012; but see Hawkins et al. 2015).

A different implementation of collapsing-bound models is through a rising urgency signal that is parameterized in a so-called gain parameter. This gain parameter increases with the duration of the decision (Churchland et al. 2008, Cisek et al. 2009, Ditterich 2006, Thura et al. 2012).

Importantly, the fixed- and collapsing-bound models make different predictions for the shape of RT distributions of correct and error responses (**Figure 5*b***). Hawkins et al. (2015)

exploited these differential predictions by fitting data from both humans and nonhuman primates using different versions of the DDM. The results showed that whereas there is occasional evidence for a collapsing-bound DDM, this model outperformed the fixed-bound DDM only under certain circumstances, mostly for monkeys and after extensive practice.

CONCLUDING COMMENTS

The last 15 years have witnessed an explosion of interest in sequential sampling models such as the DDM. This interest was initially fueled by the realization that sequential sampling models provide a principled and plausible account of the macrolevel dynamics of the behavior of single cells. Currently this interest has shifted somewhat, and many applications in high-level cognitive neuroscience use models such as the DDM to decompose performance into its constituent psychological processes, such that brain measurements may be connected not to observed behavior, but to specific latent processes of interest.

Another recent shift in interest is evident from research efforts that aim to extend the DDM to novel tasks and new dependent variables, and to probe its adequacy under a set of circumstances that had not been originally considered. New applications and extensions of the DDM now appear on a regular basis, and they constitute one of the most exciting recent trends in the neuroscience of speeded decision making. The work described here is a testament to the symbiosis that is slowly arising between mathematical psychology and cognitive neuroscience; this symbiosis and the accelerated development of quantitative models for brain and behavior hold much promise for the future.

ACKNOWLEDGMENTS

We would like to thank Martijn Mulder, Leendert van Maanen, and Guy Hawkins for valuable discussions. This research line is financially supported by the European Research Council (B.U.F., E.-J.W.), the Dutch Research Foundation NWO (B.U.F.), and the National Institute of Aging (R.R.; grant R01-AG041176).

Glossary

Speed-accuracy trade-off
the universal finding that response time can be shortened at the expense of a higher error rate

Diffusion decision model (DDM)
a model of speeded decision making in which noisy information is accumulated over time until a threshold of evidence is reached

RDK
random dot kinematogram

Response time (RT)
the time between stimulus onset and response execution in a decision-making task performed under substantial time pressure

SPRT
sequential probability ratio test

Drift rate
parameter in the diffusion decision model that quantifies the information used in the accumulation process

Boundary separation
parameter in the diffusion decision model that quantifies response caution and accounts for the speed-accuracy trade-off

Model-based cognitive neuroscience
interdisciplinary field that studies cognition by combining insights and models from mathematical psychology with measurement tools from neuroscience

Single-cell recording
recordings of spiking activity for individual neurons as measured for instance in monkeys and rats

LIP
lateral intraparietal cortex

Functional magnetic resonance imaging (fMRI)
popular brain-imaging technique used to locate brain areas that are relatively active during task performance

Electroencephalography (EEG)
popular method for measuring electrical activity along the scalp, used to study the temporal aspects of information processing in the brain

LITERATURE CITED

Audley RJ, Pike AR. Some alternative stochastic models of choice. Br. J. Math. Stat. Psychol. 1965; 18:207–25.

Basso MA, Wurtz RH. Modulation of neuronal activity in superior colliculus by changes in target probability. J. Neurosci. 1998; 18:7519–34. [PubMed: 9736670]

Bode S, Sewell DK, Lilburn S, Forte JD, Smith PL, Stahl J. Predicting perceptual decision biases from early brain activity. J. Neurosci. 2012; 32:12488–98. [PubMed: 22956839]

Boehm U, Van Maanen L, Forstmann B, Van Rijn H. Trial-by-trial fluctuations in CNV amplitude reflect anticipatory adjustment of response caution. NeuroImage. 2014; 96:95–105. [PubMed: 24699015]

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced choice tasks. Psychol. Rev. 2006; 113:700–65. [PubMed: 17014301]

Boucher L, Palmeri TJ, Logan GD. Inhibitory control in mind and brain: an interactive race model of countermanding saccades. Psychol. Rev. 2007; 114:376–97. [PubMed: 17500631]

Bowman NE, Kording KP, Gottfried JA. Temporal integration of olfactory perceptual evidence in human orbitofrontal cortex. Neuron. 2012; 75:916–27. [PubMed: 22958830]

Brown SD, Heathcote A. The simplest complete model of choice response time: linear ballistic accumulation. Cogn. Psychol. 2008; 57:153–78. Proposed a competing nonstochastic model for decision making. [PubMed: 18243170]

Carazza B. The history of the random-walk problem: considerations on the interdisciplinarity in modern physics. Riv. Nuovo Cimento Ser. 2. 1977; 7:419–27.

Carpenter RHS. Contrast, probability, and saccadic latency: evidence for independence of detection and decision. Curr. Biol. 2004; 14:1576–80. [PubMed: 15341745]

Cassey P, Heathcote A, Brown SD. Brain and behavior in decision-making. PLOS Comput. Biol. 2014; 10:e1003700. [PubMed: 24991810]

Cavanagh JF, Wiecki TV, Cohen MX, Figueroa CM, Samanta J, et al. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat. Neurosci. 2011; 14:1462–67. [PubMed: 21946325]

Churchland AK, Kiani R, Shadlen MN. Decision-making with multiple alternatives. Nat. Neurosci. 2008; 11:693–702. [PubMed: 18488024]

Cisek P, Puskas GA, El-Murr S. Decisions in changing conditions: the urgency-gating model. J. Neurosci. 2009; 29:11560–71. [PubMed: 19759303]

Cohen MX. A neural microcircuit for cognitive conflict detection and signaling. Trends Neurosci. 2014; 37:480–90. [PubMed: 25034536]

Deco G, Rolls ET, Albantakis L, Romo R. Brain mechanisms for perceptual and reward-related decision-making. Prog. Neurobiol. 2013; 103:194–213. [PubMed: 22326926]

Ding L, Gold JI. Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. Cereb. Cortex. 2012; 22:1052–67. [PubMed: 21765183]

Ditterich J. Evidence for time-variant decision making. Eur. J. Neurosci. 2006; 24:3628–41. [PubMed: 17229111]

Ditterich J. A comparison between mechanisms of multi-alternative perceptual decision making: ability to explain human behavior, predictions for neurophysiology, and relationship with decision theory. Front. Neurosci. 2010; 4:184. [PubMed: 21152262]

Drugowitsch J, Moreno-Bote R, Churchland AK, Shadlen MN, Pouget A. The cost of accumulating evidence in perceptual decision making. J. Neurosci. 2012; 32:3612–28. [PubMed: 22423085]

Egan JP. Recognition memory and the operating characteristic. 1958 USAF Oper. Appl. Lab Tech. Note AFCRC-TN-58-51, Hear. Commun. Lab., Indiana Univ., Bloomington, IN.

Einstein A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Ann. Phys. 1905; 17:549–60.

Feller, W. An Introduction to Probability Theory and Its Applications. Wiley; New York: 1968.

Ferrera VP, Yanike M, Cassanello C. Frontal eye field neurons signal changes in decision criteria. Nat. Neurosci. 2009; 12:1458–62. [PubMed: 19855389]

Forstmann BU, Dutilh G, Brown S, Neumann J, von Cramon DY, et al. Striatum and pre-SMA facilitate decision-making under time pressure. PNAS. 2008; 105:17538–42. Showed that increased activity in a corticobasal ganglia network is associated with interindividual differences in response caution. [PubMed: 18981414]

Forstmann, BU.; Wagenmakers, EJ., editors. An Introduction to Model-Based Cognitive Neuroscience. Springer; New York: 2015.

Forstmann BU, Wagenmakers EJ, Eichele T, Brown S, Serences J. Reciprocal relations between cognitive neuroscience and cognitive models: opposites attract? Trends Cogn. Sci. 2011; 6:272–79.

Geddes J, Ratcliff R, Allerhand M, Childers R, Wright RJ, et al. Modeling the effects of hypoglycemia on a two-choice task in adult humans. Neuropsychologia. 2010; 24:652–60.

Glimcher PW. The neurobiology of visual-saccadic decision making. Annu. Rev. Neurosci. 2003; 26:133–79. [PubMed: 14527268]

Good IJ. Studies in the history of probability and statistics. XXXVII. A.M. Turing's statistical work in World War II. Biometrika. 1979; 66:393–96.

Gold IJ, Law C-T, Connolly P, Bennur S. The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. J. Neurophysiol. 2008; 100:2653–68. [PubMed: 18753326]

Gold IJ, Shadlen MN. Representation of a perceptual decision in developing oculomotor commands. Nature. 2000; 404:390–94. [PubMed: 10746726]

Gold IJ, Shadlen MN. Neural computations that underlie decisions about sensory stimuli. Trends Cogn. Sci. 2001; 5:10–16. [PubMed: 11164731]

Gold IJ, Shadlen MN. The neural basis of decision making. Ann. Rev. Neurosci. 2007; 30:535–74. [PubMed: 17600525]

Hanes DP, Schall JD. Neural control of voluntary movement initiation. Science. 1996; 274:427–30. Proposed a connection between diffusion processes and buildup of activity in neurons involved in decision making. [PubMed: 8832893]

Hanks T, Kiani R, Shadlen MN. A neural mechanism of speed-accuracy tradeoff in macaque area LIP. eLife. 2014; 3:e02260.

Hanks TD, Mazurek ME, Kiani R, Hopp E, Shadlen MN. Elapsed decision time affects the weighting of prior probability in a perceptual decision task. J. Neurosci. 2011; 31:6339–52. [PubMed: 21525274]

Hawkins GE, Forstmann BU, Wagenmakers EJ, Ratcliff R, Brown SD. Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision making. J. Neurosci. 2015; 35:2476–84. Tested one of the most popular dynamic-threshold extensions of the DDM in both humans and nonhuman primates. [PubMed: 25673842]

Heitz RP, Schall JD. Neural mechanisms of speed-accuracy tradeoff. Neuron. 2012; 76:616–28. [PubMed: 23141072]

Horwitz GD, Newsome WT. Separate signals for target selection and movement specification in the superiorcolliculus. Science. 1999; 284:1158–61. [PubMed: 10325224]

Kepecs A, Uchida N, Zariwala HA, Mainen ZF. Neural correlates, computation and behavioural impact of decision confidence. Nature. 2008; 455:227–31. [PubMed: 18690210]

Kiani R, Corthell L, Shadlen MN. Choice certainty is informed by both evidence and decision time. Neuron. 2014; 84:1329–42. [PubMed: 25521381]

Krajbich I, Armel C, Rangel A. Visual fixations and the computation and comparison of value in simple choice. Nat. Neurosci. 2010; 13:1292–98. [PubMed: 20835253]

Laming, DRJ. Information Theory of Choice-Reaction Times. Academic; London: 1968.

Latty T, Beekman M. Speed-accuracy trade-offs during foraging decisions in the acellular slime mould Physarum polycephalum. Proc. R. Soc. B. 2011; 278:539–45.

Leite FP, Ratcliff R. Modeling reaction time and accuracy of multiple-alternative decisions. Atten. Percept. Psychophys. 2010; 72:246–73. [PubMed: 20045893]

Leite FP, Ratcliff R. What cognitive processes drive response biases? A diffusion model analysis. Judgm. Decis. Mak. 2011; 6:651–87.

Link SW, Heath RA. A sequential theory of psychological discrimination. Psychometrika. 1975; 40:77–105.

Luce, RD. Response Times. Oxford Univ. Press; New York: 1986.

Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. Nat. Neurosci. 2006; 9:1432–38. [PubMed: 17057707]

Mulder MJ, Bos D, Weusten JMH, van Belle J, van Dijk SC, et al. Basic impairments in regulating the speed-accuracy tradeoff predict symptoms of attention-deficit/hyperactivity disorder. Biol. Psychiatry. 2010; 68:1114–19. Reviewed fMRI studies that used a model-based cognitive neuroscience approach to perceptual decision making. [PubMed: 20926067]

Mulder MJ, van Maanen L, Forstmann BU. Perceptual decision neurosciences—a model-based review. Neuroscience. 2014; 277:872–84. [PubMed: 25080159]

Mulder MJ, Wagenmakers EJ, Ratcliff R, Boekel W, Forstmann BU. Bias in the brain: a diffusion model analysis of prior probability and potential payoff. J. Neurosci. 2012; 32:2335–43. [PubMed: 22396408]

Murdock, BB. Human Memory: Theory and Data. Erlbaum; Potomac, MD: 1974.

Murdock, BB.; Anderson, RE. Encoding, storage, and retrieval of item information. In: Solso, RL., editor. Information Processing and Cognition: The Loyola Symposium. Erlbaum; Hillsdale, NJ: 1975. p. 145-94.

Niwa M, Ditterich J. Perceptual decisions between multiple directions of visual motion. J. Neurosci. 2008; 28:4435–45. [PubMed: 18434522]

Oberauer K, Süß H-M, Wilhelm O, Wittmann WW. The multiple faces of working memory: storage, processing, supervision, and coordination. Intelligence. 2003; 31:167–93.

Philiastides MG, Ratcliff R, Sajda P. Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. J. Neurosci. 2006; 26:8965–75. Used multivariate pattern analysis to derive spatiotemporal activity profiles that could discriminate between relevant stimulus categories and between different levels of difficulty. [PubMed: 16943552]

Platt M, Glimcher PW. Neural correlates of decision variables in parietal cortex. Nature. 1999; 400:233–38. [PubMed: 10421364]

Pleskac TJ, Busemeyer J. Two-stage dynamic signal detection: a theory of confidence, choice, and response time. Psychol. Rev. 2010; 117:864–901. [PubMed: 20658856]

Purcell BA, Heitz RP, Cohen JY, Schall JD, Logan GD, Palmeri TJ. Neurally constrained modeling of perceptual decision making. Psychol. Rev. 2010; 117:1113–43. [PubMed: 20822291]

Ratcliff R. A theory of memory retrieval. Psychol. Rev. 1978; 85:59–108. Introduced the DDM for RT distributions and accuracy, with application to recognition memory.

Ratcliff R. Theoretical interpretations of speed and accuracy of positive and negative responses. Psychol. Rev. 1985; 92:212–25. [PubMed: 3991839]

Ratcliff R. Continuous versus discrete information processing: modeling the accumulation of partial information. Psychol. Rev. 1988; 95:238–55. [PubMed: 3375400]

Ratcliff R. A diffusion model account of response time and accuracy in a brightness discrimination task: fitting real data and failing to fit fake but plausible data. Psychon. Bull. Rev. 2002; 9:278–91. [PubMed: 12120790]

Ratcliff R. Modeling response signal and response time data. Cogn. Psychol. 2006; 53:195–237. [PubMed: 16890214]

Ratcliff R, Cherian A, Segraves M. A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of simple two-choice decisions. J. Neurophysiol. 2003a; 90:1392–407. [PubMed: 12761282]

Ratcliff R, Childers R. Individual differences and fitting methods for the two-choice diffusion model of decision making. Decision. 2015; 2:237–79.

Ratcliff R, Hasegawa YT, Hasegawa YP, Smith PL, Segraves MA. Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. J. Neurophysiol. 2007a; 97:1756–74. [PubMed: 17122324]

Ratcliff R, Love J, Thompson CA, Opfer J. Children are not like older adults: a diffusion model analysis of developmental changes in speeded responses. Child Dev. 2012; 83:367–81. [PubMed: 22188547]

Ratcliff R, McKoon G. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 2008; 20:873–922. Reviewed DDM modeling, data, and applications. [PubMed: 18085991]

Ratcliff R, Perea M, Colangelo A, Buchanan L. A diffusion model account of normal and impaired readers. Brain Cogn. 2004a; 55:374–82. [PubMed: 15177817]

Ratcliff R, Philiastides MG, Sajda P. Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. PNAS. 2009; 106:6539–44. [PubMed: 19342495]

Ratcliff R, Rouder JN. Modeling response times for two-choice decisions. Psychol. Sci. 1998; 9:347–56.

Ratcliff R, Smith PL. A comparison of sequential sampling models for two-choice reaction time. Psychol. Rev. 2004; 111:333–67. [PubMed: 15065913]

Ratcliff R, Starns JJ. Modeling confidence judgments, response times, and multiple choices in decision making: recognition memory and motion discrimination. Psychol. Rev. 2013; 120:697–719. [PubMed: 23915088]

Ratcliff R, Thapar A, McKoon G. The effects of aging on reaction time in a signal detection task. Psychol. Aging. 2001; 16:323–41. [PubMed: 11405319]

Ratcliff R, Thapar A, McKoon G. A diffusion model analysis of the effects of aging on brightness discrimination. Percept. Psychophys. 2003b; 65:523–35. [PubMed: 12812276]

Ratcliff R, Thapar A, McKoon G. A diffusion model analysis of the effects of aging on recognition memory. J. Mem. Lang. 2004b; 50:408–24.

Ratcliff R, Thapar A, McKoon G. Application of the diffusion model to two-choice tasks for adults 75–90 years old. Psychol. Aging. 2007b; 22:56–66. [PubMed: 17385983]

Ratcliff R, Thapar A, McKoon G. Individual differences, aging, and IQ in two-choice tasks. Cogn. Psychol. 2010; 60:127–57. [PubMed: 19962693]

Ratcliff R, Thapar A, McKoon G. Effects of aging and IQ on item and associative memory. J. Exp. Psychol.: Gen. 2011; 140:464–87. [PubMed: 21707207]

Ratcliff R, Van Dongen HPA. Sleep deprivation affects multiple distinct cognitive processes. Psychon. Bull. Rev. 2009; 16:742–51. [PubMed: 19648462]

Ratcliff R, Van Zandt T, McKoon G. Connectionist and diffusion models of reaction time. Psychol. Rev. 1999; 106:261–300. [PubMed: 10378014]

Resulaj A, Kiani R, Wolpert DM, Shadlen MN. Changes of mind in decision-making. Nature. 2009; 461:263–66. Introduced a DDM change-of-mind extension allowing information in the perceptual processing pipeline to influence and possibly overrule the initial decision. [PubMed: 19693010]

Riefer DM, Knapp BR, Batchelder WH, Bamber D, Manifold V. Cognitive psychometrics: assessing storage and retrieval deficits in special populations with multinomial processing tree models. Psychol. Assess. 2002; 14:184–201. [PubMed: 12056081]

Roitman JD, Shadlen MN. Response of neurons in the lateral interparietal area during a combined visual discrimination reaction time task. J. Neurosci. 2002; 22:9475–89. [PubMed: 12417672]

Roxin A, Ledberg A. Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. PLOS Comput. Biol. 2008; 4:e1000046. [PubMed: 18369436]

Schall JD. Neural basis of deciding, choosing and acting. Nat. Rev. Neurosci. 2001; 2:33–42. [PubMed: 11253357]

Schall JD. Macrocircuits: decision networks. Curr. Opin. Neurobiol. 2013; 23:269–74. [PubMed: 23246279]

Schmiedek F, Oberauer K, Wilhelm O, Süß H-M, Wittmann W. Individual differences in components of reaction time distributions and their relations to working memory and intelligence. J. Exp. Psychol.: Gen. 2007; 136:414–29. [PubMed: 17696691]

Shadlen MN, Kiani R. Decision making as a window on cognition. Neuron. 2013; 80:791–806. [PubMed: 24183028]

Shadlen MN, Newsome WT. Motion perception: seeing and deciding. PNAS. 1996; 93:628–33. [PubMed: 8570606]

Shadlen MN, Newsome WT. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. 2001; 86:1916–36. [PubMed: 11600651]

Smith PL. Stochastic dynamic models of response time and accuracy: a foundational primer. J. Math. Psychol. 2000; 44:408–63. [PubMed: 10973778]

Smith PL. From Poisson shot noise to the integrated Ornstein-Uhlenbeck process: neurally principled models of diffusive evidence accumulation in decision-making and response time. J. Math. Psychol. 2010; 54:266–83.

Starns JJ, Ratcliff R, McKoon G. Evaluating the unequal-variability and dual-process explanations of zROC slopes with response time data and the diffusion model. Cogn. Psychol. 2012; 64:1–34. [PubMed: 22079870]

Stone M. Models for choice reaction time. Psychometrika. 1960; 25:251–60.

Teodorescu AR, Usher M. Disentangling decision models: from independence to competition. Psychol. Rev. 2013; 120:1–38. [PubMed: 23356779]

Thapar A, Ratcliff R, McKoon G. A diffusion model analysis of the effects of aging on letter discrimination. Psychol. Aging. 2003; 18:415–29. [PubMed: 14518805]

Thura D, Beauregard-Racine J, Fradet CW, Cisek P. Decision making by urgency gating: theory and experimental support. J. Neurophysiol. 2012; 108:2912–30. [PubMed: 22993260]

Townsend, JT.; Ashby, FG. Stochastic Modeling of Elementary Psychological Processes. Cambridge Univ. Press; London: 1983.

Usher M, McClelland JL. The time course of perceptual choice: the leaky, competing accumulator model. Psychol. Rev. 2001; 108:550–92. Outlined one of the first models to implement principles from the neuroscience of decision making in a formal psychological model. [PubMed: 11488378]

van Ravenzwaaij D, Brown S, Wagenmakers EJ. An integrated perspective on the relation between response speed and intelligence. Cognition. 2011; 119:381–93. [PubMed: 21420077]

van Ravenzwaaij D, Dutilh G, Wagenmakers EJ. A diffusion model decomposition of the effects of alcohol on perceptual decision making. Psychopharmacology. 2012; 219:1017–25. [PubMed: 21842158]

van Vugt MK, Simen P, Nystrom LE, Holmes P, Cohen JD. EEG oscillations reveal neural correlates of evidence accumulation. Front. Neurosci. 2012; 6:106. [PubMed: 22822389]

Van Zandt, T. Analysis of response time distributions. In: Wixted, JT.; Pashler, H., editors. Stevens' Handbook of Experimental Psychology, Vol. 4: Methodology in Experimental Psychology. 3rd. Wiley; New York: 2002. p. 461-516.

Van Zandt T, Colonius H, Proctor RW. A comparison of two response time models applied to perceptual matching. Psychol. Bull. Rev. 2000; 7:208–56.

Vandekerckhove J, Tuerlinckx F. Fitting the Ratcliff diffusion model to experimental data. Psychon. Bull. Rev. 2007; 14:1011–26. [PubMed: 18229471]

Vandekerckhove J, Tuerlinckx F. Diffusion model analysis with MATLAB: a DMAT primer. Behav. Res. Methods. 2008; 40:61–72. [PubMed: 18411528]

Vandekerckhove J, Tuerlinckx F, Lee MD. Hierarchical diffusion models for two-choice response times. Psychol. Methods. 2011; 16:44–62. [PubMed: 21299302]

Vickers, D. Decision Processes in Visual Perception. Academic; New York: 1979.

Vickers D, Lee MD. Dynamic models of simple judgments. I: Properties of a self-regulating accumulator module. Nonlinear Dyn. Psychol. Life Sci. 1998; 2:169–94.

Voss A, Nagler M, Lerche V. Diffusion models in experimental psychology: a practical introduction. Exp. Psychol. 2013; 60:385–402. [PubMed: 23895923]

Voss A, Rothermund K, Voss J. Interpreting the parameters of the diffusion model: an empirical validation. Mem. Cogn. 2004; 32:1206–20.

Voss A, Voss J. Fast-dm: a free program for efficient diffusion model analysis. Behav. Res. Methods. 2007; 39:767–75. [PubMed: 18183889]

Voss A, Voss J. A fast numerical algorithm for the estimation of diffusion-model parameters. J. Math. Psychol. 2008; 52:1–9.

Wabersich D, Vandekerckhove J. Extending JAGS: a tutorial on adding custom distributions to JAGS (with a diffusion model example). Behav. Res. Methods. 2014; 46:15–28. [PubMed: 23959766]

Wagenmakers EJ. Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. Eur. J. Cogn. Psychol. 2009; 21:641–71.

Wagenmakers EJ, Brown S. On the linear relation between the mean and the standard deviation of a response time distribution. Psychol. Rev. 2007; 114:830–41. [PubMed: 17638508]

Wagenmakers EJ, Ratcliff R, Gomez P, McKoon G. A diffusion model account of criterion shifts in the lexical decision task. J. Mem. Lang. 2008; 58:140–59. [PubMed: 19122740]

Wagenmakers EJ, van der Maas HL, Grasman RP. An EZ-diffusion model for response time and accuracy. Psychon. Bull. Rev. 2007; 14:3–22. [PubMed: 17546727]

Wald A, Wolfowitz J. Optimum character of the sequential probability ratio test. Ann. Math. Stat. 1948; 19:326–39.

Wang XJ. Decision making in recurrent neuronal circuits. Neuron. 2008; 60:215–34. [PubMed: 18957215]

White C, Ratcliff R, Vasey M, McKoon G. Dysphoria and memory for emotional material: a diffusion model analysis. Cogn. Emot. 2009; 23:181–205. [PubMed: 19750142]

White CN, Ratcliff R, Vasey MW, McKoon G. Using diffusion models to understand clinical disorders. J. Math. Psychol. 2010; 54:39–52. [PubMed: 20431690]

Wiecki TV, Sofer I, Frank MJ. HDDM: Hierarchical Bayesian estimation of the drift-diffusion model in Python. Front Neuroinform. 2013; 7:14. [PubMed: 23935581]

Wong KF, Wang XJ. A recurrent network mechanism of time integration in perceptual decisions. J. Neurosci. 2006; 26:1314–28. [PubMed: 16436619]

Wyart V, de Gardelle V, Scholl J, Summerfield C. Rhythmic fluctuations in evidence accumulation during decision making in the human brain. Neuron. 2012; 76:847–58. [PubMed: 23177968]

Zandbelt B, Purcell BA, Palmeri TJ, Logan GD, Schall JD. Response times from ensembles of accumulators. PNAS. 2014; 111:2848–53. [PubMed: 24550315]

Zeguers MHT, Snellings P, Tijms J, Weeda WD, Tamboer P, et al. Specifying theories of developmental dyslexia: a diffusion model analysis of word recognition. Dev. Sci. 2011; 14:1340–54. [PubMed: 22010894]

SUMMARY POINTS

1. The DDM provides an excellent account of the law-like patterns observed across virtually all speeded RT tasks.
2. The DDM accounts for the relationship between mean RTs and the probability of choices (errors and correct responses) in both healthy and diseased populations.
3. One of the main reasons for the current popularity of diffusion models in neuroscience is the behavior of single neurons of monkeys (and occasionally rats) performing simple decision-making tasks.
4. In addition, fMRI and EEG data show specific neural patterns related to DDM model parameters, thereby offering the promise of a mechanistic understanding of latent cognitive processes.
5. Extensions of the DDM include multiple-choice behavior, confidence judgments, value-based decision making, and dynamic decision thresholds.
6. A DDM decomposition of choice performance provides numerous benefits, both for a purely behavioral analysis and for a model-based cognitive neuroscience approach.

FUTURE ISSUES

1. How do people set and adjust criteria for response caution? The speed with which people achieve relatively stable criteria suggests that they bring to bear substantial prior knowledge.
2. To what extent can the DDM prove useful in examining deficits in various neuropsychological disorders?
3. How can the DDM be extended to more complex and multistage decision making?
4. What exactly is the relation between latent processes in the DDM and key structures in the human brain (e.g., control structures in the basal ganglia and structures that support working memory processes in the frontal cortex)?
5. How can we build truly integrated models of decision making, that is, models that include knowledge of how motor processes are implemented in motor cortex and the oculomotor system?
6. To what extent can the DDM be applied to more deliberate economic decision making?
7. Will the DDM be able to quantitatively account for choice behavior in groups of animals?
8. To what extent can the DDM be used to jointly model behavioral and neuroscience data pertaining to clinical populations?

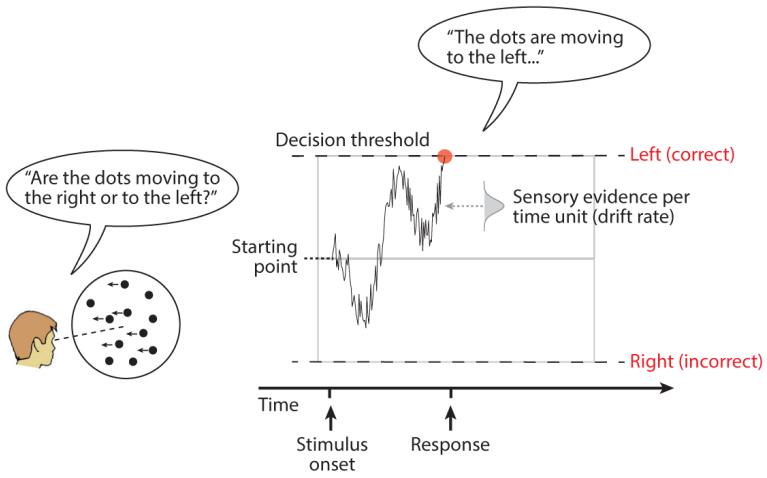

Figure 1.

Schematic representation of the diffusion decision model. Figure adapted with permission from Mulder et al. (2012).

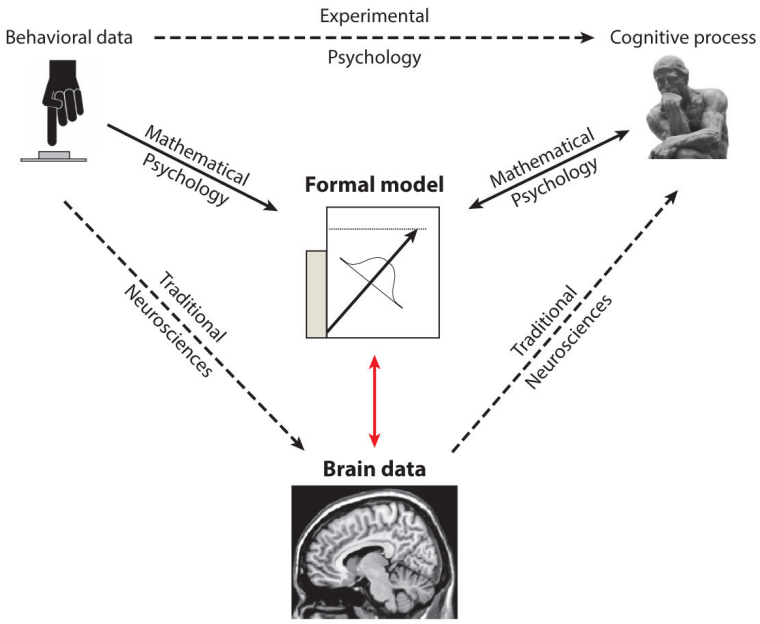

Figure 2.

The model-in-the-middle approach unites the separate disciplines of experimental psychology, mathematical psychology, and cognitive neuroscience in the common goal of understanding human cognition. The red double-headed arrow indicates the reciprocal relation between measuring the brain and modeling behavioral data. Figure adapted with permission from Forstmann et al. (2011).

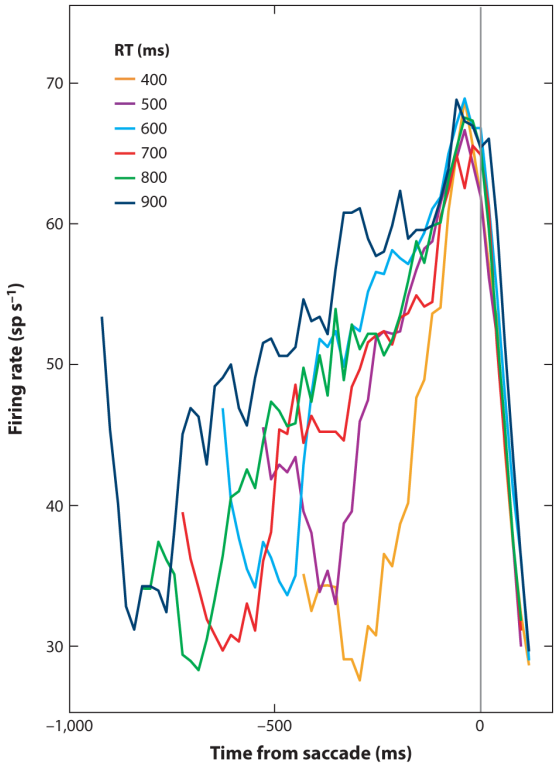

Figure 3.
Immediately before monkeys execute a saccade, all trials reach a stereotyped firing rate. This dynamic is consistent with that of a diffusion model with evidence accumulation to a fixed threshold. Abbreviation: RT, response time. Figure adapted with permission from Gold & Shadlen (2007).

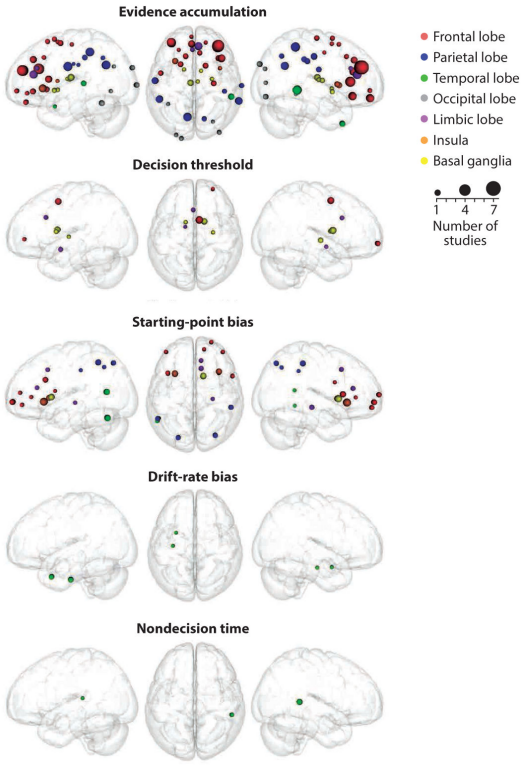

Figure 4.

Summary of peak coordinates reported in functional magnetic resonance imaging (fMRI) studies that correlate blood-oxygen-level-dependent activation with parameter estimates from sequential sampling models. The size of each sphere is proportional to the number of studies that reported a specific region of interest. Only studies reporting whole-brain analyses were included. Figure adapted with permission from Mulder et al. (2014).

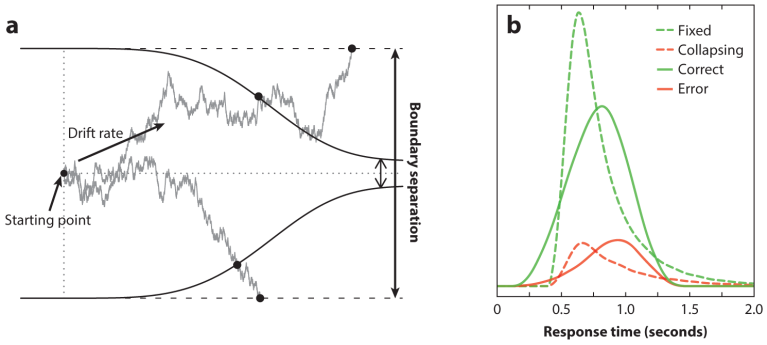

Figure 5.

A standard fixed-bound diffusion decision model (DDM) versus a collapsing-bound DDM. (*a*) The DDM with fixed (*dashed*) or collapsing (*solid*) response boundaries. Models with collapsing boundaries can terminate the evidence accumulation process earlier than models with fixed boundaries, resulting in faster decisions. (*b*) The ways in which the models lead to different predictions for response time distributions, particularly in the tails. Figure adapted with permission from Hawkins et al. (2015).