Trends in Cognitive Sciences

CellPress

Feature Review

Diffusion Decision Model: Current Issues and History

Roger Ratcliff,[1,*] Philip L. Smith,[2] Scott D. Brown,[3] and
Gail McKoon[1]

There is growing interest in diffusion models to represent the cognitive and neural processes of speeded decision making. Sequential-sampling models like the diffusion model have a long history in psychology. They view decision making as a process of noisy accumulation of evidence from a stimulus. The standard model assumes that evidence accumulates at a constant rate during the second or two it takes to make a decision. This process can be linked to the behaviors of populations of neurons and to theories of optimality. Diffusion models have been used successfully in a range of cognitive tasks and as psychometric tools in clinical research to examine individual differences. In this review, we relate the models to both earlier and more recent research in psychology.

Trends

Diffusion models with drift and boundaries constant over time account for accuracy and correct and error response time distributions for many types of two-choice tasks in many populations of participants.

Collapsing decision bounds implement optimal decision making in certain cases, but fits to data show humans use constant boundaries.

Brief stimulus presentation produces time varying input, but data suggest that evidence is integrated to produce constant drift in the decision process. (Other tasks can produce nonstationary evidence.)

Evidence is assumed to vary from trial to trial, as in signal detection theory. This explains why incorrect decisions are often slower than correct decisions.

It is not clear if variability in a sequence of stimulus elements in expanded judgment tasks is equivalent to moment-by-moment internal noise in tasks with a single stationary stimulus.

Modeling Simple Decision Making

Decision making is intimately involved in all of our everyday activities. Many decisions are made rapidly and at a low level cognitively, for example, deciding whether to drive left or right round a car in front. Others, such as deciding which candidate to vote for or which car to buy, are made at a higher level with prolonged deliberation. The diffusion models we discuss are of the former type. In the real world, they involve a rapid matching of a perceptual representation to stored knowledge in memory, which allows us to identify things in our immediate surroundings and determine how we should respond to them. Much of what we have learned about such decisions comes from laboratory tasks in which people are asked to make fast two-choice decisions. The measures of performance are typically response times (RTs) and the probabilities of making the two choices. Researchers are usually interested in how and why RTs and choice probabilities change across experimental conditions, for example, whether a person tries to respond as quickly as possible or as accurately as possible.

There have been a moderate number of models for these tasks and most assume accumulation of noisy evidence to decision criteria representing each of the two choices. The models can include one versus two **accumulators** (see Glossary), decision rules that are relative or absolute, models with drift rate constant or varying over time, discrete or continuous time evidence, stochastic versus deterministic evidence, and models with inhibition and decay. Ratcliff and Smith [1] showed the relationships between the models along with a detailed evaluation of the models (Figure 1, Key Figure).

The standard model that we will discuss was developed by Ratcliff in the 1970s [2] and has only changed in assuming a single diffusion process instead of racing processes [3] and in adding **across trial variability** in starting point [4,5] and **nondecision time** [6]. In this model (Figure 2A), evidence about a stimulus from perception or memory accumulates from a starting point to a boundary or **threshold** (i.e., a criterion), one boundary for each choice. The boundaries represent the amount of evidence that must be accumulated before a response

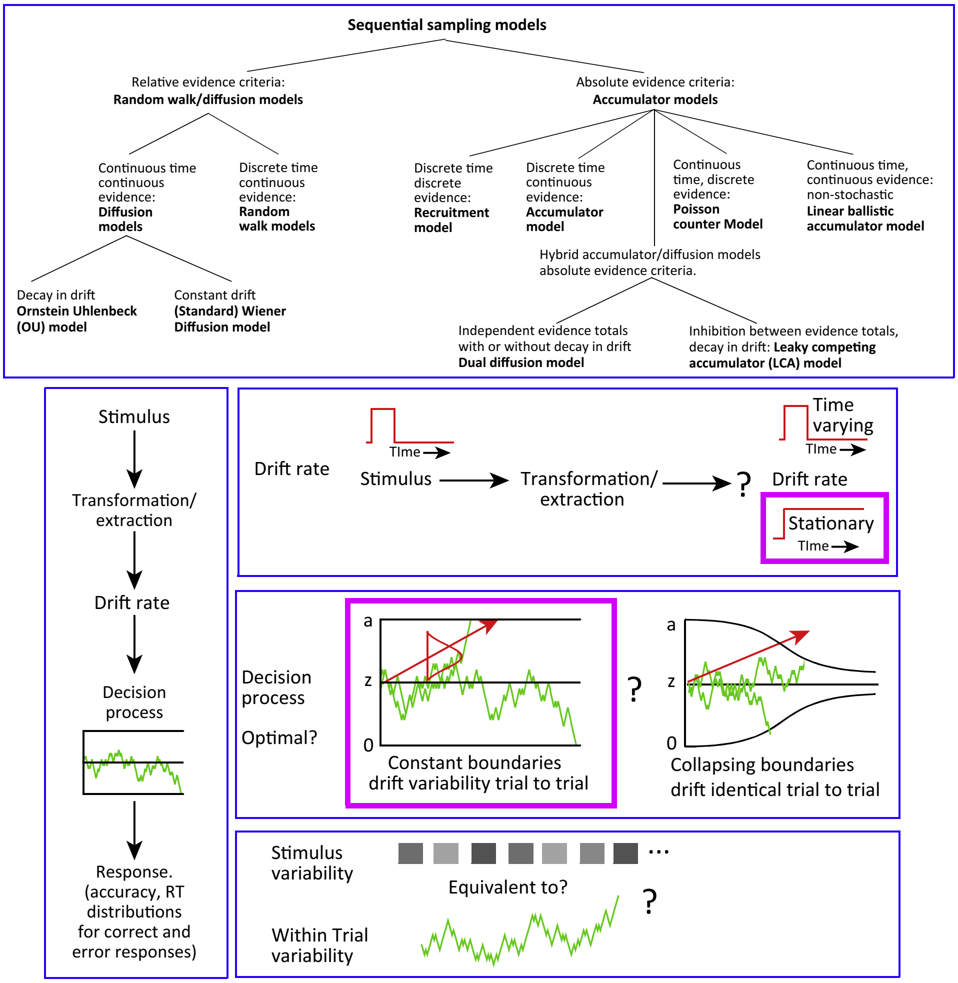
[1]Department of Psychology, The Ohio State University, 1835 Neil Avenue, Columbus, OH, 43210, USA
[2]Melbourne School of Psychological Sciences, Level 12, Redmond Barry Building 115, University of Melbourne, Parkville, VIC 3010, Australia
[3]School of Psychology, University of Newcastle, Australia, Aviation Building, Callaghan, NSW 2308, Australia

*Correspondence: ratcliff.22@osu.edu (R. Ratcliff).

http://dx.doi.org/10.1016/j.tics.2016.01.007

Key Figure

Sequential Sampling Models



Trends in Cognitive Sciences

Figure 1. Relationships between the models, a flowchart of processing, and three issues addressed in the article.

is made. The accumulation process is noisy; at each moment in time, the evidence might point to one or the other of the two boundaries, but more often to the correct than the incorrect one.

The main components of the model for the decision process represent the rate of accumulation and the settings of the boundaries. In Figure 2, the boundaries are set at 0 and *a* with starting point *z*. Evidence accumulates in a noisy manner, and the average rate of accumulation is called the ‘drift rate’. In addition, there are nondecision components: encoding the evidence from a stimulus that will drive the decision process, extracting the dimension(s) of the stimulus that form the basis of the decision from the stimulus or memory, and executing a response. These nondecision components are combined and labeled the ‘nondecision’ component, which has a

Glossary

Accumulator: an assumed structure in an evidence accumulation model that has the purpose of gathering evidence in favor of one response.

Across-trial variability: the assumption that drift rates vary from decision to decision, motivated by the idea that, even if physical stimulus conditions are identical, the internal representation of the decision-relevant information is not.

Attractor model: a network (graph-based) model of interconnected nodes with a dynamic updating process. The updating process causes changes that lead to a stable end state (at the ‘attractor’).

Collapsing boundary: an assumption that the amount of evidence required to trigger a decision (the ‘threshold’) becomes smaller as the time taken to make the decision increases. This contrasts with the standard assumption that the threshold is unchanging.

Confidence: a subjective rating of the likely accuracy of a decision provided by the decision maker.

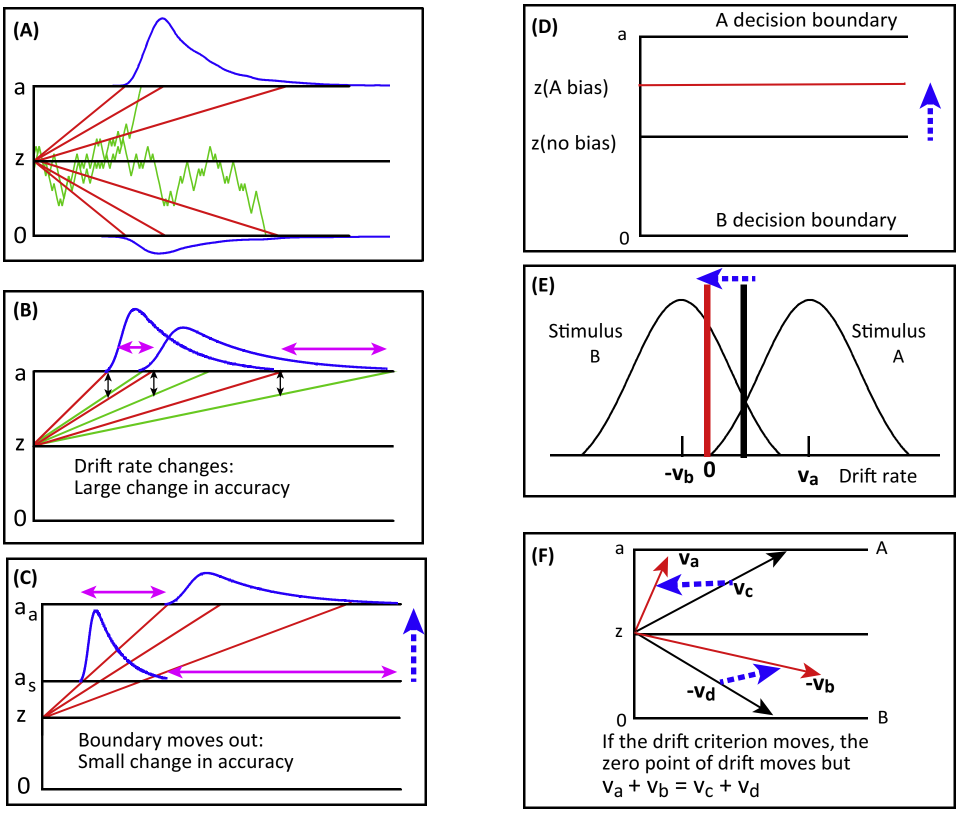
Evidence accumulation: also known as ‘sequential sampling’. The idea that decisions are made by gathering evidence from the environment, continuing until sufficient evidence (a ‘threshold’ amount) is gathered.

Fast errors: an empirical phenomenon in which the mean RT for incorrect responses is longer than that for correct responses. Reliably observed when decision making is easy or decision makers stress speed. It has been important for model development because it is inconsistent with many theories of decision making. See also ‘slow errors’.

Hopfield network: a type of attractor model based on recurrent connections that has been used to model human memory and decision processes, among other things.

Latency–probability (LP) and quantile–probability (QP) plots: parametric plots that show the relationship between the probabilities of different classes of responses and the timing of those responses. Response times can be plotted either as means (LP) or as quantiles (QP).

Nondecision time: the component of RT that is not due to evidence accumulation. Usually modeled as



Trends in Cognitive Sciences

Figure 2. An Illustration of the Diffusion Model. (A) Shows two (irregular) simulated paths in the diffusion model (green). The blue curves represent response time (RT) distributions for correct responses (top) and errors (bottom). The red lines represent the fastest, medium, and slowest responses. (B) Shows the effect of lowering drift rate by a fixed amount. The black double arrows show the effect on fast, medium, and slow average drift rates and the magenta arrows show the effect on the fastest and slowest responses from the blue RT distributions. There is a small change in the leading edge of the distribution and a large change in the tail. (C) Shows the effect of moving a boundary away from the starting point [*a*-speed (a_s) to *a*-accuracy (a_a), the blue broken arrow] to represent a speed–accuracy manipulation (both boundaries would move in a real experiment). The magenta arrows show the effect on the fastest and slowest responses from the blue RT distributions. There is a moderate change in the leading edge of the distribution and a large change in the tail. The difference in effects between (B) and (C) discriminates manipulations that change boundaries from manipulations that change drift rates. (D) Shows how a bias toward the A response can be modeled by a change in the starting point (blue broken arrow with the starting point moving from the black line to the red line). RT distributions change as in (C). (E) Shows how a bias toward the A response can be modeled by a change in the zero point of drift rate (blue broken arrow with the zero point moving from the black line to the red line). (F) Shows the effect of a change in the zero point of drift rate, from (E). Drift rate is first symmetrical (black arrows) and then biased toward A (the red arrows). RT distributions change as in (B). The parameters of the model are boundary separation (a), starting point (z), drift rate (v, one of each condition), nondecision time (T_{er}), which is the duration of encoding and response output processes and the transformation from the stimulus representation to a decision-relevant representation. Parameters of the model are assumed to vary from trial to trial, drift rate is normally distributed with standard deviation η, starting point and nondecision time are assumed to have rectangular distributions with ranges s_z and s_t, respectively.

mean time of T_{er}. To set the measurement scale, one parameter of the model must be fixed (otherwise, e.g., doubling all the rates of **evidence accumulation** while also doubling the boundary separation would not change the predictions of the model). In theory, any model parameter could be fixed; in practice, usually the parameter governing moment-by-moment variability in evidence accumulation is fixed. Concise introductions to this model, which we will refer to as 'the diffusion model' in this review, are available elsewhere [7], as are more general comparative studies of the large class of sequential sampling models (Figure 1).

either a fixed offset or a rectangular additive distribution.

Random walk model: a theory of how decisions are made that posits evidence accumulating semi-randomly over time until enough is gathered in favor of one decision over another.

Response signal task: decision-making paradigm in which an experimenter-controlled signal informs the participant when a response is required. Also called 'time controlled task', 'exogenous RT task', and 'deadline task' (although the latter usually refers to a task in which the same signal duration is used in a block of trials, hence allowing subjects to adopt different criteria for the different deadlines).

Sequential probability ratio test (SPRT): a statistical process for making decisions, due to Wald. The SPRT is identical to the random walk model in certain cases. The SPRT is also optimal in that it minimizes decision time for a given level of accuracy.

Slow errors: an empirical phenomenon in which the mean RT for incorrect responses is longer than that for correct responses. Reliably observed when decision making is difficult and decision makers are careful. It has been important for model development because it is inconsistent with many theories of decision making. See also 'fast errors'.

Speed–accuracy tradeoff: the empirical phenomenon that decision makers can decide to make faster decisions, sacrificing accuracy, or more accurate decisions, sacrificing speed.

Threshold: also known as 'criterion' or 'decision boundary'. The amount of evidence required to trigger a decision response.

Urgency signal: an assumption, related to the 'collapsing boundary', that the accumulated evidence is amplified more and more as the time taken to make a decision grows.

Box 1. Advantages of Diffusion Model Analyses

The model relates speed and accuracy to the same underlying components of processing for fast (less than 1–2 s) two-choice decisions and it explains why speed and accuracy are sometimes correlated and sometimes not (especially across individuals).

An individual can decide to respond as quickly or as accurately as possible. The diffusion model factors out speed–accuracy settings and thus provides better estimates of the quality of the evidence entering the decision process than is available from RT or accuracy data.

The model provides a fit to data that allows us to know whether the model is an adequate description of the data.

When there are limited numbers of materials, fitting the model to fillers and critical items increases the power for the critical items. This is because fillers are weighted heavily in determining model parameters common across conditions [115–118].

When accuracy is at ceiling, it is still possible to estimate drift rates if some of the conditions have lower accuracy. These conditions are weighted most heavily in determining some model parameters and then RTs alone are sufficient to determine drift rates [118,119].

The variability in model parameters is usually smaller than the variability between subjects, which means that the effects of individual difference measures on performance can be measured, for example, IQ and working memory [120–123].

Fitting packages are available (but subject to misuse if not understood): fast-dm [124], DMAT [125], HDDM [126], as well as alternative fitting methods [127]. See evaluations elsewhere [128,129].

Other current sequential sampling models usually offer similar explanations of phenomena in terms of the behavior of model components (parameters) [130,131].

In addition to standard two-choice tasks, the model has been successfully applied to go/no-go, response signal and deadline tasks. Related diffusion models have been applied to multialternative decision making and confidence judgments.

The model can accommodate changes in the rate of evidence accumulation during the time course of a trial [54,56–59]. The distribution of nonterminated processes is well known (derived from the Fokker–Planck forward equation) and this can be used as a new starting point distribution in a second phase of evidence accumulation after a change in drift rate. Changes in other model parameters can be examined using simulation methods.

The diffusion model and others like it have become increasingly influential over the past 10 to 15 years as models of the psychological and neural processes involved in decision making. Box 1 gives a comprehensive list of advantages of the standard model and Box 2 presents a list of the paradigms and areas of research to which it has been applied. There are three main reasons for this success. First, the models account for all the behavioral data, namely accuracy and the shapes and locations of the distributions of RTs for correct responses and for incorrect responses. Second, they have been linked to neural processing for single cells and populations of neurons and they have been linked to aggregate behavior as measured by electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and other imaging methods (Box 3 discusses explicit links between neural models and diffusion models). Third, they have been successful in explaining decision making across wide domains of psychology such as aging, child development, various clinical populations, and animal species, often providing new interpretations of data. For example, as age increases in adults RT increases. In many tasks, fits of the diffusion model show that the quality of the evidence encoded from a stimulus (drift rate) does not decrease; instead, the slowing occurs because the boundaries are set to increase the amount of evidence required for a response and nondecision times are longer.

Current diffusion models are the culmination of 50 years of theoretical and empirical research [2,8–11], which has identified the key features of experimental data that a model must explain and the key properties by which a model can do so. Many recent studies have focused on new phenomena and new areas of application but neglected findings in the older literature. This neglect is potentially detrimental because the older literature contains modeling and experimental work that speaks directly to current issues. Our aims in this review are to redress this neglect

Box 2. Domains of Application of Diffusion Model Analyses

Diffusion and other models have been applied to many basic perceptual and memory tasks such as item and associative recognition [2,15,121,122,132], lexical decision [121,133], perceptual tasks including brightness, letter, motion, visual search, contrast, orientation discrimination tasks [7,14,16,44,103,104,134,135], numeracy judgments [13,121,123], categorization [132,136], and text processing and priming [116,117]. Other tasks that have more interdisciplinary relationships include stop signal tasks [137,138], conflict tasks [70,71], reinforcement learning [69,139], preferential choice and value-based decisions (Box 7), and social decisions [140,141] (Box 7).

Relationships have been established between diffusion model analyses and behavioral measures such as eye tracking [128,137] and pupil dilation [142]. Many studies have established relationships between diffusion models and neurophysiological measures such as single cell recordings in rodents and monkeys [52,88,105,143–148], EEG [42,43,149,150], fMRI [42,43,139,151–155], transcranial magnetic stimulation (TMS) [156], and transcranial alternating current stimulation (tACS) [157]. Diffusion model analyses have also been applied to bees and animal swarms [158,159] and even to slime molds [160].

Diffusion model analyses have been used to study manipulations of state such as sleep deprivation [161], hypoglycemia [162], and alcohol [163]. They have been used to study individual differences in IQ, working memory, and reading measures [120–123], and to examine deficits in populations such as aphasics [164], older adults and children [13,14,121,122], children [165], low literacy adults [166], dyslexics [167], attention deficit hyperactivity disorder (ADHD) [168,169], schizophrenia [170], and in depressed and anxious individuals [171,172].

Box 3. Linking Neural Firing Rates to Diffusion Processes

Qualitative links have been made between neural firing rates and diffusion processes [143,144,173]. More explicit modeling has attempted to link the dynamics of **attractor networks** to decision making by identifying decision making with a neural network entering an attractor state. This uses linear approximations of the network equations and constraints on the network parameters to reduce its dimensionality.

One theoretical challenge is to show that these network models can reproduce the structure found in families of RT distributions for correct responses and errors for real decision-making data.

The attractor model developed by Wang assumes two pools of excitatory neurons coupled via a third pool of inhibitory neurons [174,175]. A two-component linear diffusion approximation to the network dynamics was proposed and the model was reduced to a much simpler representation consistent with current diffusion models. One attractive aspect of this modeling approach is that neurally plausible assumptions are made about inputs based on currents, neurotransmitters, etc.

A different approach assumed two pools with a nonlinear diffusion approximation containing both linear and cubic terms [176]. The cubic term was designed to mimic the attractor dynamics of Wong and Wang [174]. It is an open question whether the cubic diffusion equation can also reproduce the detailed features of accuracy and RT distributions found in human behavioral data.

The attractor network model, the Ising Decision Model [177], is based on a stochastic **Hopfield network**. Stimulus information, represented as the drift of an approximating diffusion process, is identified with (minus) the gradient of a potential field that sets the attractor states of the network. A decision is made when the network first enters one of two neighborhoods surrounding the attractor states. The model successfully reproduces the behavior of RT distributions found in human data and the model makes an explicit theoretical connection between the physics of diffusion and the properties of the potential field that determine the attractor states of that network.

Diffusive noise in a decision process can be derived from a Poisson shot noise model of stimulus representations [178,179]. The Poisson shot noise process represents the variability in the postsynaptic potential across a neural population that is induced by volley of action potentials modeled as a Poisson process. In this model, stimulus information is represented by the difference between excitatory and inhibitory shot noise pairs. The model produced the families of RT distributions predicted by the standard diffusion model.

and to highlight findings in the older literature that present challenges to currently accepted interpretations of data and currently unresolved issues.

The Two-Choice Diffusion Model

Figure 2 shows simulated paths that represent the accumulation of evidence on individual trials (Figure 2A) and it shows the effects of changes in drift rate (Figure 2B) and boundary settings (Figure 2C) on RT distributions. Because there is a minimum on RTs but no maximum, the model automatically produces right-skewed distributions that have the same shape as those found in most simple two-choice tasks. That the model predicts RT distributions that are the same as those found experimentally is one of the most important properties of the model and one of the strongest tests of it [12].

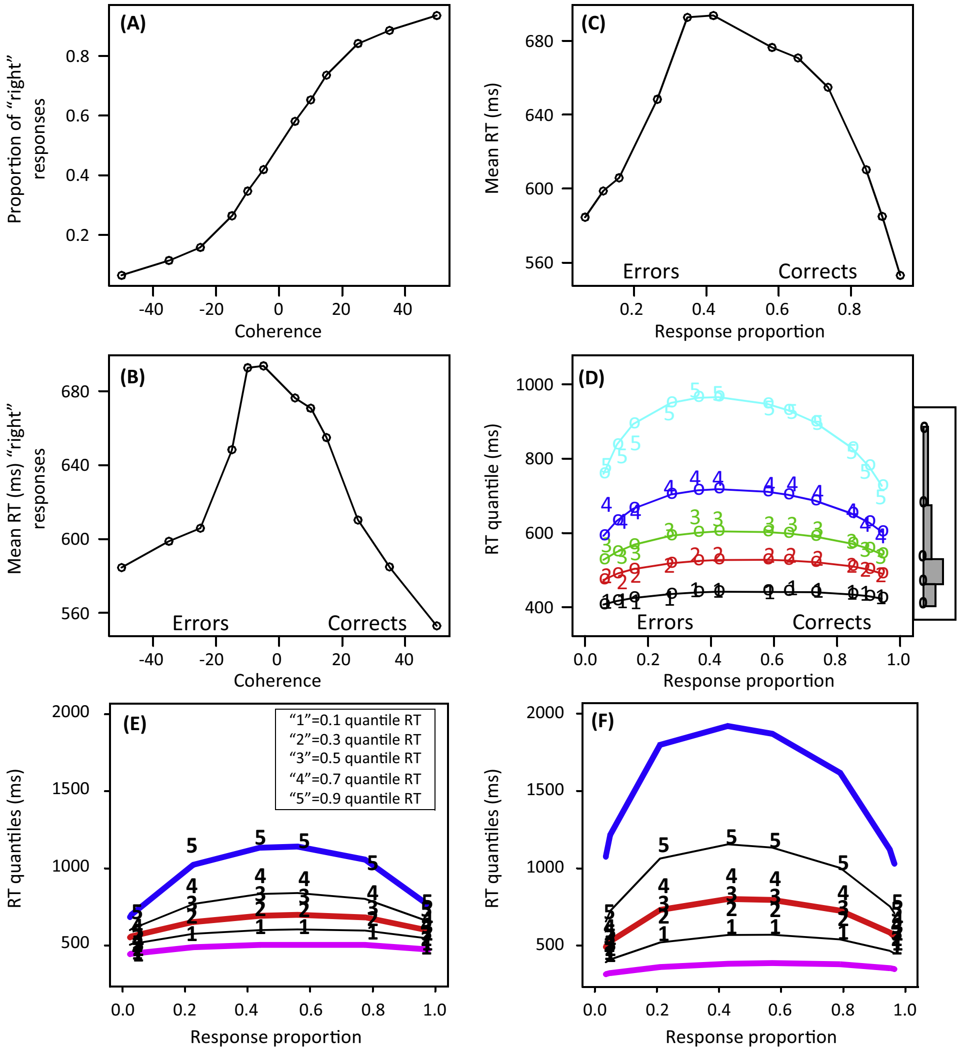
Drift rates are determined by the quality of evidence extracted from the stimulus or memory (often a different value for each condition of the experiment). **Speed–accuracy tradeoff** effects and the effects of bias toward one of the boundaries over the other are ubiquitous in the experimental literature and they are explained naturally by the structure of the model. Usually (Figure 2C), speed–accuracy effects are explained by changes in the boundary settings (e.g., [4,13–16]) and much smaller changes in nondecision time [17,18], although instructions that extremely stress speed can reduce drift rates ([19], see also [20]).

Figure 2D–F illustrates ways in which the model can accommodate bias toward one alternative or the other [21]. If the probability that one alternative is tested is made higher than the other, then the starting point moves toward the higher probability boundary (Figure 2D) [5,22–26]. Some recent investigations have suggested that the effects of this manipulation for both humans and monkeys can be accounted for better by a bias in drift rate, not starting point [27]. However, both accounts make key predictions about the behavior of RT distributions, which were not examined. Bias can also be manipulated by payoffs, for example, paying more for correct responses to one of the alternatives than to the other. In this case, the result is a combination of bias in starting point (Figure 2D) and drift rate (Figure 2E,F) [23–25].

Changes in starting point and drift rates can also explain sequential effects [21,28–33] in which trial-by-trial variations in RT are partly determined by the prior stimulus and the prior response. With a rapid rate of presentation for easy stimuli, sequential effects can extend for several trials, but for slower presentation rates and more difficult stimuli, they are found only for the immediately preceding trial. Adaptive regulatory mechanisms to account for trial-by-trial effects have been proposed by a number of investigators [34–36]. In the diffusion model, sequential effects can be modeled by making the starting point and drift rate functions of prior trials [5].

In many experimental reports, plots of either accuracy or mean RT alone are used to describe data. However, the two dependent variables must be considered simultaneously. The data in Figure 3A–D come from a motion discrimination task in which participants see a display of dots and decide whether a subset of them is moving right or left; the fewer the dots moving coherently, the more difficult the decision (Experiment 1 [7]). The effects are typical – more difficult conditions have lower accuracy (Figure 3A) and slower responses (Figure 3B), but it is the relation between them (the **latency–probability** function in Figure 3C) and the shapes and locations of the full RT distributions (the **quantile–probability** function in Figure 3D) that must be the targets for models. In Figure 3D, the RT distribution is represented by the 0.1, 0.3, 0.5, 0.7, and 0.9 quantile RTs. The numbers are the data and the circles with lines between them are the values predicted by the diffusion model. The model fits the data well, and it does so quantitatively as well as qualitatively. The model is falsifiable in that it must predict the right-skewed shape of RT distributions [37].

Some researchers question the need to fit RT distributions; some simply omit any consideration of them. The importance of distributions is illustrated in Figure 3E,F, which show data from a letter discrimination task with dynamic random pixel noise (Experiment 1 [38]). These data cannot be fit with only drift rate changing (in Figure 3E the magenta line should go through the '1' symbols, the red line through the '3' symbols, and the blue line through the '5' symbols). Figure 3F shows how fitting the median RT alone can give misleading results. The red line goes through the '3' (median) symbols, which indicates a good fit, but the RT distributions that would



Trends in Cognitive Sciences

Figure 3. Relationships between Variables. (A–D) Plots of data from a motion discrimination experiment, Experiment 1 [7]. (A) Response proportion plotted against motion coherence. (B) Mean response time (RT) plotted against motion coherence. (C) Mean RT plotted against response proportion. (D) RT quantiles plotted against response proportion for data (digits) and model predictions (open circles and lines). The quantiles used were the 0.1, 0.3, 0.5, 0.7, and 0.9 quantiles and these represent, respectively, the fastest 10% of responses, the fastest 30%, the median RT, the slowest 30%, and the slowest 10%. The quantiles are stacked vertically and the small inset to the right shows equal area rectangles drawn between quantiles to illustrate what RT distributions derived from the quantiles would look like. (E,F) Fits to a dataset that has large differences in the leading edge of the RT distribution, changes too large for a diffusion model with only drift rate changing over conditions to fit (from Experiment 1 [38]). (E) Data and predictions for fits to RT distribution quantiles. The 0.1 quantile (magenta) and median (red) miss the data. (F) Data and predictions for RT quantiles for fits to accuracy and median RT. The median (red) fits well, but the other quantiles miss badly.

be predicted miss the data badly (the predicted 0.1 and 0.9 quantile RTs miss the data by several hundred milliseconds). Thus, we strongly recommend that predictions for RT distributions be examined in any application of the model to data. Diffusion models for multialternative decision making and **confidence** are discussed in Box 4.

Across-Trial Variability in Model Components

A problem with early **random walk models**, which were discrete time precursors of diffusion models, was that they predicted identical RT distributions for correct and incorrect responses

Box 4. Multialternative Decision Making Including Confidence Judgments

There is no simple generalization of the two-choice diffusion model to multiple alternatives [9,180]. Instead, the usual approach assumes independent racing single boundary diffusion processes. This generalizes easily to any number of alternatives, but differs from the standard model because the racing diffusion model does not include response competition as in the standard model. Response competition can be easily added to the model so that movement toward one boundary entails movement away from the others.

Recent studies have examined models for multialternative decision making that use diffusion processes [24,41,72,79,80,181–189] and a number of algorithms have been used:

(i) Independent racing accumulators with termination when one reaches its decision criterion.
(ii) Independent racing accumulators with a relative stopping rule (termination occurs when one accumulator beats the maximum of the others by some amount).
(iii) Accumulators with dependence between accumulators: inhibition between accumulators that depends on the amount of accumulated evidence.
(iv) Accumulators with dependence: evidence for one alternative is evidence against the others so that the total evidence is constant. When one accumulator is incremented, the others are decremented (termed constant summed evidence or feed-forward inhibition). This can be seen as a generalization of the two-choice model in which evidence for one choice is evidence against the other.
(v) Other architectural choices include whether evidence can fall below zero, whether there is decay, and whether parameters vary across trials. The models that implement these choices are difficult to discriminate (see [186]).

Confidence judgment tasks are also multialternative tasks. In them, subjects are often asked to rate their confidence in a decision about two alternatives. In the majority of applications, the proportions of responses at each level of confidence are the primary data, and signal detection theory has been the dominant model. These tasks have had a long history in psychology [190–192].

Previously, little attention has been paid to confidence RTs [193,194] and even less attention to modeling them. However, recent work in psychology has accounted jointly for response proportions and RT distributions in confidence judgment tasks [186,195,196].

In animal work, it seems impossible to get animals to respond on a scale (e.g., six confidence choices). An opt-out procedure has been used to examine confidence in animals [146,147].

In recent modeling in neuroscience, time has been used as a measure of decision confidence. However, in many studies, RT distributions for different levels of confidence overlap and thus RT cannot uniquely determine the confidence levels.

(when the starting point is equidistant from the boundaries), which is never observed empirically. Several different approaches to this problem have been investigated, including: dynamically changing decision boundaries; nonlinear evidence accumulation processes; and non-normal random walk increments. The most extensively evaluated approach to the problem has been the assumption of trial-to-trial variability in model parameters [2,7,9]. In the standard model, drift rate is normally distributed across trials with standard deviation (SD) η, the starting point is uniformly distributed with range s_z (starting point variability is equivalent to variability in the boundaries), and nondecision time is uniformly distributed with range s_t (Ratcliff examined these parametric forms [39]). Recently, it has been argued that with assumptions of across-trial variability, any pattern of data can be accommodated, but the argument only applies to deterministic models (Box 5).

The assumption that model parameters vary from trial to trial is made by most current models that successfully account for experimental data in psychological applications [1,40]. This assumption has a long history [2,4,9] of extensive testing and it allows models to explain the relative speeds of correct and incorrect responses. When decisions are difficult and decision makers are cautious, incorrect responses are reliably slower than correct responses (see later in relation to **collapsing boundary** models). When decisions are easier and decision makers are hurrying, incorrect responses are reliably faster than correct responses (see [32]).

Across-trial variability in drift rate produces **slow errors** (relative to correct responses) because trials with randomly higher drift rates are associated with fast responses, but very few errors. By

Box 5. Is the Diffusion Model Identifiable or Too Flexible?

If the forms of across-trial variability distributions are unconstrained, then Jones and Dzhafarov argued that the diffusion model and other evidence accumulation models can exactly match any data (response probabilities and RT distributions), rendering the models unfalsifiable [197]. Below are considerations about this argument [198].

(i) The most important point is that the evidence in [197] applies only to completely different deterministic or near-deterministic models, which the diffusion model is not.

(ii) In a deterministic model, if every process travels the same fixed distance (starting point to the boundary), then every RT can be converted into a velocity (drift rate) by velocity = distance/time (drift rate is a constant multiplied by 1/RT), thus producing a one-to-one mapping between drift rates and RTs (Figure IA).

(iii) In such a deterministic model, to account for errors, complex, bimodal distributions of drift, consisting of two unequally sized, asymmetrical lobes, must be assumed. The probability mass in each corresponds to the proportion of responses of each type (Figure IB). Every different RT arises from a different value of drift (because it is a constant times 1/RT) and the drifts for correct responses and errors have opposite signs because errors can only occur when the sign of the drift is wrong.

(iv) Because response probabilities and RT distributions vary from condition to condition, a different bimodal distribution of drift with a different shape is needed for every condition of an experiment. In the standard model, if the distributions of drift rate or starting point are changed modestly [2,39], the model produces similar predictions (within trial noise washes out effects of distribution shape as in the central limit theorem).

(v) The resulting model is complex with highly unintuitive properties. As stimuli become more discriminable and easier, the asymmetry of the distribution of drift and the separation between its positive and negative lobes increases (Figure IB).

(vi) To account for the finding that repeated presentation of the same stimulus can lead to different responses, the model must assume that the sign of the drift can vary from presentation to presentation and that the magnitude of the difference between drifts leading to correct responses and errors increases as the task becomes easier (i.e., error drift rates become more strongly negative).

(vii) In a few cases, neurophysiological data can address the distributions of drift rates. For example, single trial EEG measures are consistent with unimodal distributions rather than bimodal distributions [197] ([42],see Figure 2D,E in main text).

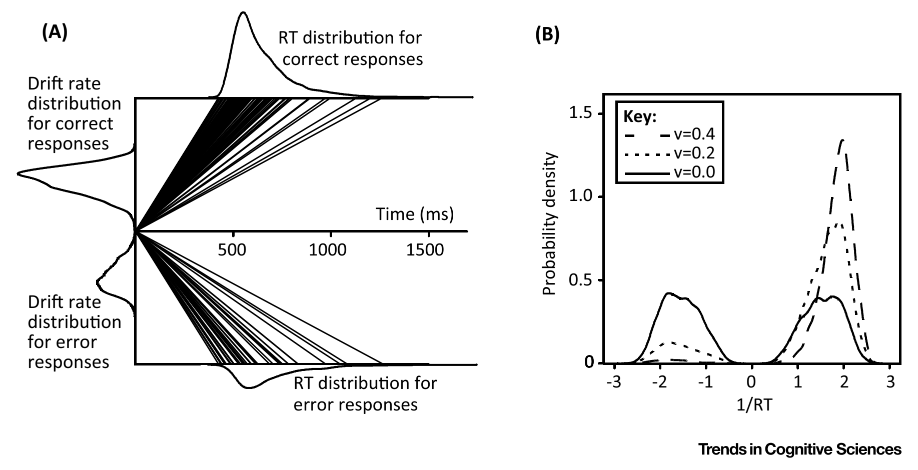
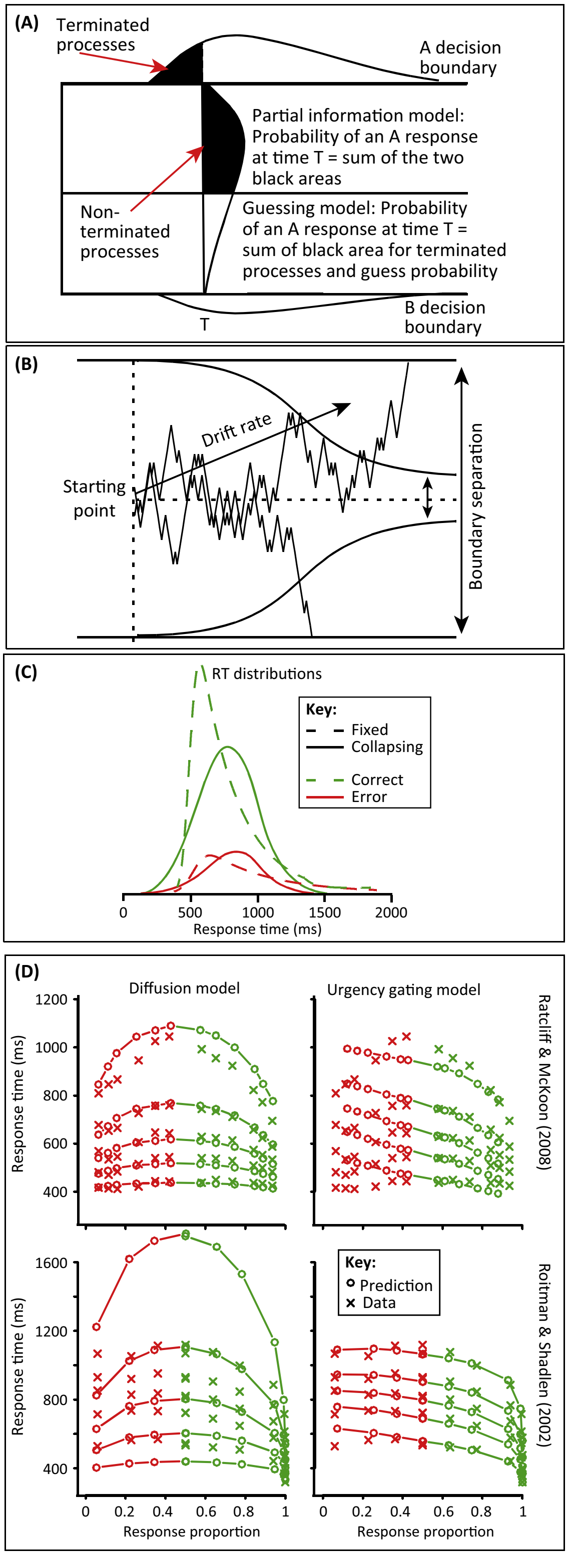


Figure I. Drift Rates and RT Distributions from the Deterministic Model. (A) An example of correct and error RT distributions generated by the standard diffusion model and the distribution of drift rates in the deterministic model used to generate them. (B) Examples of the distributions of drift rate in the deterministic model for different levels of drift rate.

contrast, trials with randomly lower drift rates are associated with slow responses, many of which are also incorrect. From this mixture (**fast errors** with low probability and slow errors with higher probability), error responses are slower than correct responses (see [7], Figure 4). For similar reasons, across-trial variability in starting point gives fast errors. We have seen few, if any, patterns of incorrect RTs versus correct RTs that cannot be accounted for with the across-trial variability assumption, although many such possibilities exist. Note that there are other ways of producing fast or slow errors (relative to correct responses) such as collapsing bounds (discussed later), but few of these have received extensive testing [41].



Trends in Cognitive Sciences

Figure 4. Response Signal and Collapsing Bound Models. (A) Modeling the response signal procedure. (B) An illustration of collapsing boundaries. (C) Sample RT distributions for collapsing and fixed decision boundaries. (D) Sample quantile–probability function for collapsing and fixed boundary models. Data from [7,88].

Another source of support for across-trial variability in drift rate comes from a face/car perceptual discrimination task in which EEG signals are used to sort responses into two groups, those that are more face-like and those that are more car-like ([42], see also [43]). When the diffusion model was applied to the RT and accuracy data for the two groups separately, robust differences in drift rates were produced, showing that the EEG signals indexed the trial-to-trial differences in evidence entering the decision process.

The assumption that drift rate varies from trial to trial [2] has been controversial in some circles [44–46], but across-trial variability in drift rate is no different than variability in signal strength in signal detection theory and, ironically, the latter assumption is almost universally accepted. Alternatives to the across-trial variability assumption have been proposed including the assumption that drift rates ramp up over time (an '**urgency signal**') and the assumption that boundaries collapse over time. Although these can predict slow errors, they cannot account quantitatively for the full range of data and we discuss this shortly.

Response Signal and Go/No-Go Tasks: Implicit Boundaries

In the **response signal task**, a stimulus is presented and then after some amount of time, a signal is given. Participants are asked to choose between two alternatives just as in the usual two-choice procedure except that they are asked to respond as quickly as possible after the signal (in, say, 200–300 ms). The stimulus-to-signal time varies from trial to trial [47–50], which means that processing can be assumed to be the same for all signal times up to each specific signal time [2]. A response signal task is often used in animal studies in which responses are made following a cue (e.g., [51–53]).

In application of the diffusion model to response signal data [17,54,55], there are two response boundaries just as for the usual two-choice task. When a decision is made at some signal lag, responses come from a mixture of processes: those that have terminated at a boundary and those that have not (Figure 4A). As the stimulus-to-signal time increases, a larger and larger proportion of processes will have terminated before the signal. For nonterminated processes, there are two possible hypotheses: that decisions are made on the basis of the partial information that has already been accumulated (which has low accuracy relative to terminated processes [55]) or that they are guesses. In fits of the diffusion model, these hypotheses could not be discriminated [17].

Response signal studies from the 1980s show that drift rate can change from stimulus-to-signal intervals that are about the mean RTs in the usual procedure (600 ms or less) up to intervals of around 2 s [56–59]. Changes in drift rate occur when two sources of information are pitted against each other. For example, early in processing, responses to 'a bird is a robin' are mainly 'true', reflecting the strong association between birds and robins, but later in processing they are 'false'. This differential availability of information over time might be thought similar to data from mouse-tracking paradigms in which, for example, tastiness information becomes available earlier than healthfulness information in a dietary choice task [60]. However, the smooth mouse tracks obtained on single trials in such studies do not match the highly irregular paths of the diffusion model (Figure 2A). The link between smooth mouse tracks or arm-reaching trajectories and the underlying process of evidence accumulation is unlikely to be as simplistic as the one-to-one mapping commonly assumed (e.g., [61]).

In a go/no-go task, participants are to respond to one of the two types of stimuli (e.g., dots moving left) but withhold a response to the other (e.g., dots moving right). In neuropsychological and clinical research, a pervasive view is that the task measures inhibitory control (e.g., [62–64]). However, in the diffusion model, when it is assumed that there are two boundaries, one implicit, the shorter RTs and lower accuracy for 'go' stimuli are explained as a bias of the starting point toward the ‘go’ boundary [65,66]. What has been assumed to be an ability to suppress responses is interpreted simply as a bias in processing (sometimes a combination of biases, as shown in Figure 2D–F).

Another domain that concerns stationarity in processing is how changes in evidence might be detected. Diffusion and related model analyses of tasks in which stimulus information varies from moment to moment provide a theoretical account of change detection [67,68] in which evidence accumulation has to be balanced against changes in the stimulus environment.

Conflict Tasks

Another class of paradigms that appear to require dynamic changes in diffusion model parameters over time are conflict paradigms. In a reinforcement learning paradigm, subjects had to choose one of a pair of letters and feedback indicated which one was ‘correct’ on that trial. Feedback was probabilistic with one letter of the pair being reinforced more often than the other. Later in the session, conflict conditions were created by pairing the letters from different pairs that had low probability of reinforcement. Responses to these conflict pairs had shifts in RT distributions relative to the training pairs and other pairs in which high probability letters were paired with other letters. The conflict conditions were modeled with collapsing decision boundaries that accounted for the shifts in RT distributions [69]. (This pattern of shifts in the RT distributions could also be modeled by shifts in nondecision time.)

A second conflict paradigm is the Eriksen flanker paradigm. In this, ‘>’ symbols are used to indicate the direction of the response (‘>’ right and ‘<’ left). The flanker manipulation involves placing other symbols indicating the same direction or the other direction around the central target that indicates the response. Two diffusion models have been developed, one that assumes dual stages, and one that assumes that evidence driving the process changed continuously over time as a result of attention gradually focusing on the central target [70,71]. The key for testing these models is the behavior of RT distributions: the behavior of error versus correct RT distributions provides the critical tests of the models.

Optimality

Considerations of optimality have played a significant role in the theoretical and experimental analysis of human and animal decision making. Theories of optimality prescribe how the available evidence should be used to produce a best decision, in some specified sense; experimental studies of optimality have investigated whether actual behavior approximates the theoretical ideal. In simple decision making, two different senses of optimality have been promoted (Box 6). One of these is based on Wald's **sequential probability ratio test (SPRT)** from statistics; the other is based on reward rate maximization. Wald showed that a random walk decision process that accumulates the log-likelihood ratios of the observed evidence sequence, given the two decision alternatives, is optimal in the sense of needing the smallest number of evidence samples to reach a prescribed level of accuracy. Optimality in Wald's sense was influential in the development of early random walk models of human decision making in psychology [8,9,22]. A pure (Wiener) diffusion process, with no across-trial variability, can be viewed as a continuous-time log-likelihood ratio accumulator, and is optimal in the same sense [72,73].

A significant limitation of the Wald SPRT test is that it restricts its applicability to real world decision tasks with constant evidence and boundaries and thus it applies only to decisions between pairs of alternatives whose properties are known. It is not applicable, for example, to tasks in which stimuli of varying discriminability are presented in random order in a block of trials unless the boundary settings are set differently for each different difficulty level (and optimally, for each one). In most decision environments, this is not possible because setting the appropriate boundary requires advanced knowledge of the upcoming decision difficulty. However, in the few

Box 6. Optimality: Do Subjects Adopt Optimal Boundary Settings?

There are two senses in common use:

(i) A diffusion process is optimal in that for a single drift rate, the process requires the minimum time on average to produce a given level of accuracy (determined by boundary settings).

(ii) For any experiment with any number of conditions, a value of boundary settings (with bounds constant over time) can be computed that makes the number correct per unit time a maximum (reward rate optimality).

Problems

For (i): If parameters vary from trial to trial and/or there are multiple conditions in an experiment, sense 1 of optimality no longer holds.

It is possible to compute the optimal boundary shape, and this varies as a function of time.

It is difficult to see, especially in the first trials of a task, how enough information could be gathered trial by trial to allow the shape of this optimal bound to be computed. The stimulus condition is almost always not known and the response choice and RT are stochastic and thus reliable information needed to compute the bound is not available. In the derivations of optimality, there is no term in the computations (e.g., based on feedback) representing the duration or effort needed to compute and update the optimal boundary.

For (ii), when experiments have tested whether boundary settings are optimal, generally they are not. Only with feedback does performance of young adults approach reward rate optimal [199]. But older adults rarely moved more than a few percent away from asymptotic accuracy.

In experiments in which blocks of trials are difficult or easy, there is a fixed time for the block, and subjects are instructed to get as many correct regardless of errors. For difficult relative to easy blocks, subjects slow down when it is reward rate optimal for them to speed up [200].

cases where it is possible to set different boundaries for each difficulty level, human decision makers seem fairly efficient at setting those many boundaries in an optimal manner [74,75].

An alternative definition of optimality is reward rate maximization [76], defined as maximizing the number of correct decisions (and the associated reward) per unit time [72,77–81]. This definition of optimality has been promoted, especially in animal studies that use water-deprived animals and liquid rewards, as a biologically principled theory of optimality. In these tasks, it seems likely that animals will be motivated to maximize their reward rate because it also minimizes the time until the next reinforcement. Although reward rate maximization has been promoted as a general definition of optimality with equal applicability to animals and humans, it is not clear that human decision makers are motivated in the same way. Rather than seeking to maximize the returns per unit time, humans seem to be motivated to maximize the returns in the available time. For example, if two students take a 2-h exam and one obtains 60% correct in 1 h while the other obtains 80% correct in 2 h, the second student will perform better on the course. There are some situations in which reward rate is explicitly set as a goal (e.g. ‘speed tests’ in schools) but even there, there is little evidence that people actually attend optimally to this goal. Indeed, when we have investigated this hypothesis, it fails (Box 6).

Collapsing Bounds

The link with optimality theory, on the one hand, and neural studies of decision making, on the other, has led to models in which decision bounds collapse over time. In the collapsing bound model, less evidence is required to trigger a decision as time passes, that is, the boundaries collapse from initially wide spacing toward the center (Figure 4B). Another assumption with much the same effect is that fixed boundaries are maintained, but an ‘urgency signal’ is added to the accumulated evidence [82,83]. This signal is like a gain that magnifies evidence by larger and larger amounts as time passes. Models with collapsing bounds have been identified with urgency gating signals in some recent theoretical accounts of optimal coding in neural populations and in empirical single cell recording studies [41,78,82–85].

Another issue concerns the relative speed of correct and incorrect decisions. In many paradigms, particularly those with difficult decisions and an emphasis on decision caution, incorrect responses are systematically slower than correct responses, on average. Standard evidence accumulation models account for this effect by assuming variability in decision difficulty across trials [2]. Collapsing boundaries or increasing urgency signals provide an alternative way to predict that incorrect responses are slower than correct responses.

Because many of the predictions of fixed bound and collapsing bound models are very similar, a test between them is only likely to be successful if it uses large samples and is carried out at the distribution level. A large-scale investigation of data from hundreds of participants in ten different studies, from three different laboratories addressed these questions empirically [86]. The data of the great majority of participants were better described by the regular, fixed bound diffusion model than by either of the collapsing boundaries or urgency signal variants. What support did exist for the new variants was mostly confined to experiments involving non-human primate participants, or using experimental procedures optimized for non-human primates (but with human participants). Other studies that have identified support for nonstationary models have mostly employed unusual decision-making tasks, for example, with very long decision times or slowly changing stimulus properties [85,87].

Figure 4C,D illustrates, using example experimental data, that the fixed bound model outperformed the collapsing boundaries and urgency signal models. While it is true, in theory, that the collapsing boundaries can help the model to predict slow errors, in practice its predictions did not match data because the model systematically overpredicts the slowing of incorrect responses relative to correct responses. The lower-right panel of Figure 4D shows the best fit of a model with an urgency signal to decision data from monkeys [88] – these data are actually some of the most favorable for the collapsing bounds and urgency signal accounts. Incorrect responses (red crosses) in these data are slower than correct responses (green crosses). However, the urgency signal model overpredicts this effect, with incorrect responses becoming slower from right to left across the plot. This effect is almost never observed in data. Instead, especially in human data, there is almost always a characteristic inverted U shape to these plots, which is accommodated well by the standard fixed bound model, but not by the collapsing bound model (see top row of Figure 4D).

Expanded Judgment Tasks

Most recent applications of diffusion models have been to experimental tasks in which a single stimulus is presented and the noise in the evidence accumulation process arises from moment-by-moment variability in the cognitive representation of the stimulus. However, some of the earliest applications of random walk models [8,22] were to expanded judgment tasks in which a noisy sequence of stimulus elements has to be integrated to make a decision. Studying such tasks was motivated by the Wald SPRT test [89], which provided an optimality theory for decisions about discrete sequences.

There has been a recent resurgence of interest in the application of diffusion and random walk models to expanded judgment tasks, especially in neuroscience, motivated in part by a renewed interest in optimality [76,90,91]. Many researchers have used a coherent motion discrimination task (e.g., the moving dot paradigm mentioned earlier), originally developed in vision science as a pure motion stimulus, as a fast-paced expanded judgment task, assuming that successive states of the motion signal are accumulated directly by the decision process.

At present it is an open question whether the decision process treats variability in a sequence of stimulus elements as equivalent to moment-by-moment internal noise in the cognitive representation of a single stimulus. The hypothesis that external stimulus noise and internal noise are equivalent is an attractive one, but there are enough differences in the perceptual and memory demands between single stimulus decision tasks and expanded judgment tasks to make the equivalence questionable. Only one stimulus representation is required in single stimulus tasks, whereas expanded judgment tasks require a new representation of every element in the sequence, which must be integrated with the memory representation of the elements that precede it.

Currently, little is known about how this memory updating process might take place, how long it might require, or how it might depend on the complexity of the individual stimulus elements. Expanded judgment tasks vary widely in the types of stimuli they use and the way the stimuli are presented. The stimuli have included random dots [44], colored lights [92] and patches [93], line segments [94], sinusoidal gratings [95], clicks and visual pulses [52], and geometric shapes [96], with interelement intervals ranging from a few tens of milliseconds to several hundred milliseconds or longer. Unless memory updating is rapid and effortless, one would expect working memory capacity limitations to have a significant effect on performance.

Consistent with this expectation, the picture of the decision process that emerges from expanded judgment studies is more complex than the one that has come from single stimulus studies, and some of the findings have no obvious counterpart in single stimulus studies. For example, one study [52] reported that the sole source of variability in decision making by human and animal subjects was external noise in the stimulus; they found no effect of internal noise, or leakage, that is, decay of the memory representation of the stimulus sequence with the passage of time.

Other studies [93,95] found there was a strong recency weighting of stimuli: stimulus elements occurring later in the sequence were weighted more heavily in the decision than were earlier elements. Recency weighting has been observed in other studies [97] using longer stimulus sequences, but other studies found both primacy and recency effects [45]. Recency weighting has been explained by various mechanisms, including a gain control process [95] and a working memory capacity limitation [97].

Most studies using expanded judgment tasks have not analyzed RT distributions. When they have been analyzed [94], the decision time distributions were much more variable than those found in single stimulus tasks and could not be predicted by a random walk model with normally distributed increments, which is the discrete time counterpart of the diffusion model with no across-trial variability in drift rates or starting points. Instead, they were better described by a version of the Vickers accumulator model [98], in which only large stimulus elements are accumulated and small ones are ignored.

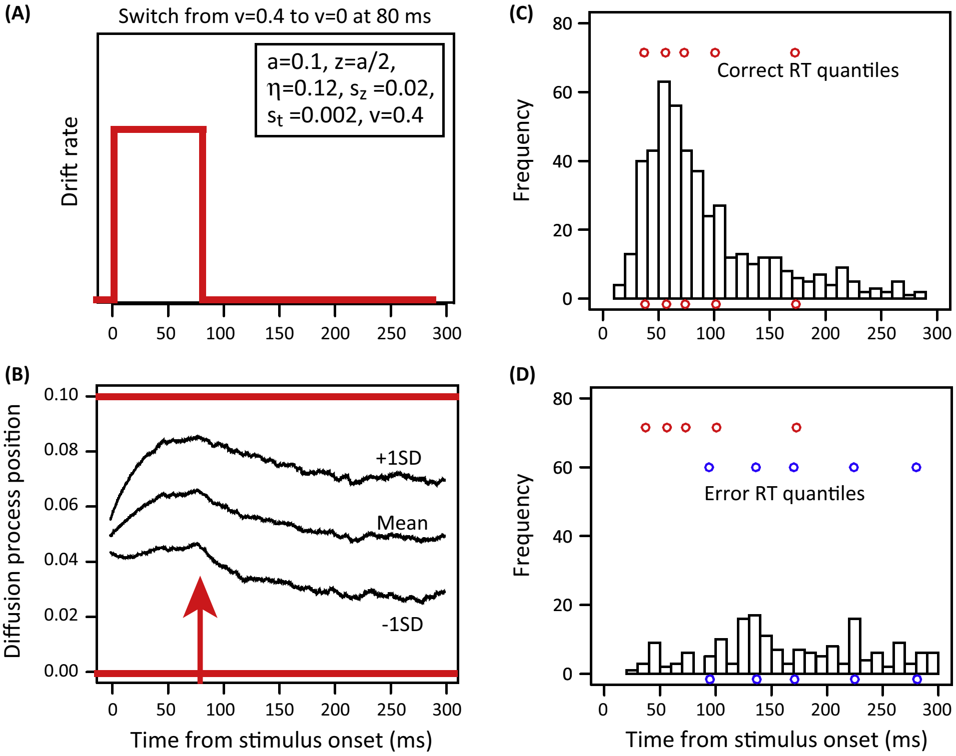
Studies using stimuli that are perturbed by external dynamic noise cast further doubt on the assumption that the decision process treats internal noise and external noise as equivalent. Many studies have assumed that the decision system accumulates the noisy output of a motion discrimination system, in which the noise arises from variability in the stimulus sequence (following [88,99]). These effects are typically modeled using a constant drift diffusion model in which the drift rate is proportional to the mean strength of the motion signal. Other work [38] has found that the effect of dynamic noise on letter discrimination was to shift the entire RT distribution to the right, delaying all responses by a constant amount. They attributed the delay to the time needed to compute a stable stimulus representation, which determines the drift rate of the diffusion process. RT distributions from the dynamic noise task were modeled with a diffusion model with a time-varying drift, in which the growth of the drift rate depended on external noise [100].

These examples highlight the fact that expanded judgment tasks and tasks in which stimuli are presented in external noise differ in significant ways from tasks in which the only source of noise is internal to the decision process. These differences caution against any *a priori* equating of the effects of internal and external noise and suggest that further work is needed to understand the relationship between them.

Brief Stimulus Presentation

There is a growing interest in whether there is nonstationarity in processing that reflects changing stimulus information over time. This is often studied with stimuli flashed for very brief times (e.g., 50 ms) and is also partly motivated by the thought that stimulus quality can change during the time course of an expanded judgment task. In a highly relevant example from about 15 years ago, the question was whether drift rate tracks the stimulus or not. In other words, whether drift rate begins at some value reflecting the stimulus information and then drops to zero when the stimulus turns off, or instead, whether information from the stimulus is integrated over time into a short-term representation that provides a constant drift rate to drive the decision process [101]. Drift rate tracking stimulus availability would be equivalent to the starting point moving nearer the boundary for correct responses, which predicts errors much slower than correct responses because errors would have much further to travel to the incorrect boundary. However, the distributions of RTs for errors and correct responses were similar, supporting a model in which drift rate is driven by a constant representation of the stimulus (see also [14,16,37,102]).

Later work [103] examined manipulations of contrast, stimulus presentation duration, and attention with Gabor patch stimuli and culminated in a model that integrated the processes



Trends in Cognitive Sciences

Figure 5. Brief Stimulus Presentation. Does the Decision Process Track Stimulus Information? (A) Shows drift rate as a function of time for the model in which drift tracks the stimulus (here turning on at zero and off at 80 ms). (B) Shows the mean and plus and minus 1 standard deviation (±1 SD) average paths from simulations (using the random walk method [171]). The red arrow is the point that drift rate turns off and the horizontal red lines are the decision boundaries. (C) Shows response time (RT) distributions for correct responses for the simulations with the red circles showing the 0.1, 0.3, 0.5, 0.7, and 0.9 quantile RTs. (D) Shows RT distributions for error responses for the simulations with the blue circles showing the 0.1, 0.3, 0.5, 0.7, and 0.9 quantile RTs; the red circles are the quantiles for the correct responses from (C).

that construct a representation of stimulus information in short-term memory with the decision process of the diffusion model [104]. Constructing a representation is necessary because stimuli are multidimensional (e.g., size, color, they might correspond to entities in memory) and thus there must be processes to pick out the dimension relevant to a decision. There are neurophysiological data that support this view. For example, in motion discrimination tasks that use a brief pulse of motion [105], there is a long-lasting effect (up to 800 ms) on decision-related firing rates in the lateral intraparietal cortex. Other studies in animals and humans have found that accuracy does not increase after some stimulus duration [37,106–110], suggesting a time by which the relevant stimulus information has been constructed. Integration time, as assessed by a change in accuracy thresholds as a function of stimulus duration, was shown to be the same (around 400 ms) for high and low coherence stimuli [107]. This suggests that the exposure duration effects in this paradigm may reflect perceptual integration processes involved in computation of the drift rate rather than evidence accumulation by a decision process.

Figure 5 illustrates how a nonstationary model mispredicts data when drift rate turns on at a relatively high value for 80 ms and then returns to zero (Figure 5A). Figure 5B shows the average position of the accumulation process rising over the first 80 ms and then falling back toward the starting point. Figure 5C,D shows the difference in the RT distributions for correct and error responses. The red circles are the 0.1, 0.3, 0.5, 0.7, and 0.9 quantile RTs for correct responses and the blue are the quantiles for errors. Overall, error responses are slower than correct responses, not the result found elsewhere [14,16,37,102]. This illustration also demonstrates that RT distributions are the critical test between stationary and nonstationary processes. Some investigations have supported nonstationarity but have not generated the critical predictions for correct and error RT distributions ([52,87,111–113], see also [45,86,114]).

This discussion raises a related question, whether the stimulus evidence driving the decision process turns on abruptly. In a simulation, drift rate was ramped linearly over 50 ms from zero to a constant level [37]. The standard constant drift model fit the simulated data almost perfectly (with different parameters than those used to generate the simulated data – increases in across-trial variability in nondecision time and starting point, and also nondecision time), suggesting that

Box 7. Preferential Choice and Value-Based Decision Making

Diffusion models have been applied to complex preferential choice applications from judgment and decision making that apply to economics and consumer behavior. Wiener and Ornstein–Uhlenbeck diffusion models have been applied to preferential choice, value-based decisions, and economic decisions for more than 25 years. The tasks in these applications are not simple speeded decisions, but rather choices between risky gambles or multiattribute consumer products.

Initial work on preferential choice was carried out using what is called decision field theory (DFT) [201]. In this model, a decision maker's attention switches from one attribute to another over time, and the advantages and disadvantages of each alternative are accumulated into a preference state. When one option reaches a decision threshold, that choice is taken. DFT has been used to fit choice and RT data from choices between risky and uncertain gambles [201]. Later, a multialternative version of DFT was used to account for the context effects (similarity, attraction, compromise effects) on choice found by consumer researchers [181]. In related research, Tsetsos *et al.* [202] used the leaky competing accumulator diffusion model to predict context effects on choice. DFT has also been used to account for puzzling reversals in preference between choice and price of gambles [203] and to explain choice and RTs for intertemporal choice [204].

Recent research on value-based decision making has used the attention drift diffusion model [205,206]. The model assumes that attention to an option changes the drift rate during preference accumulation. Novel eye tracking methods have been used to track attention to options across time and use these measurements to moderate the drift rate across time. This model has achieved impressive success to account for eye fixation data as well as both choice and RT distributions for choices between food items.

This and other diffusion models have been used to account for, among other things, value-based decision, social choice, and purchasing decisions [150,152,153,157,207–210].

ramped drift and constant drift models mimic each other and that constant drift models are a good approximation even if drift rate does ramp up over time.

Concluding Remarks

Current research in modeling decision processes has used diffusion models extensively. They are being applied in clinical and educational domains, economic decision making (Box 7), and the neuroscience of decision making. Here we have examined current issues that have a history in psychology and we have discussed the earlier research and how it complements new research. In some cases, the earlier research provides an answer to new research questions. Although we have separated the issues, many of them are related. Collapsing decision bounds have been argued to be optimal and able to replace assumptions of across-trial variability in drift rate. Likewise, nonstationary drift rates are related to collapsing bounds and are one hypothesis of what would occur with brief stimulus presentation. These basic questions are important as the field uses this modeling approach in clinical and neuroscience domains (see Outstanding Questions).

Acknowledgments

We would like to thank Jerome Busemeyer, Josh Gold, and Tim Hanks for comments on the paper. Preparation of this article was supported by National Institute on Aging grant R01-AG041176, Department of Education/Institute of Educational Sciences grant R305A120189, and Australian Research Council grants DP140102970 and FT120100244.

References

1. Ratcliff, R. and Smith, P.L. (2004) A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* 111, 333–367
2. Ratcliff, R. (1978) A theory of memory retrieval. *Psychol. Rev.* 85, 59–108
3. Ratcliff, R. (1981) A theory of order relations in perceptual matching. *Psychol. Rev.* 88, 552–572
4. Ratcliff, R. and Rouder, J.N. (1998) Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356
5. Ratcliff, R. *et al.* (1999) Connectionist and diffusion models of reaction time. *Psychol. Rev.* 106, 261–300
6. Ratcliff, R. and Tuerlinckx, F. (2002) Estimating parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameter variability. *Psychon. Bull. Rev.* 9, 438–481
7. Ratcliff, R. and McKoon, G. (2008) The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922
8. Stone, M. (1960) Models for choice reaction time. *Psychometrika* 25, 251–260
9. Laming, D.R.J. (1968) *Information Theory of Choice Reaction Time,* Wiley
10. Link, S.W. (1975) The relative judgement theory of two choice response time. *J. Math. Psychol.* 12, 114–135
11. Link, S.W. and Heath, R.A. (1975) A sequential theory of psychological discrimination. *Psychometrika* 40, 77–105
12. Ratcliff, R. *et al.* (2015) Modeling response time and accuracy data. *Curr. Dir. Psychol. Sci.* 24, 458–470
13. Ratcliff, R. *et al.* (2001) The effects of aging on reaction time in a signal detection task. *Psychol. Aging* 16, 323–341
14. Ratcliff, R. *et al.* (2003) A diffusion model analysis of the effects of aging on brightness discrimination. *Percept. Psychophys.* 65, 523–535
15. Ratcliff, R. *et al.* (2004) A diffusion model analysis of the effects of aging on recognition memory. *J. Mem. Lang.* 50, 408–424
16. Thapar, A. *et al.* (2003) A diffusion model analysis of the effects of aging on letter discrimination. *Psychol. Aging* 18, 415–429
17. Ratcliff, R. (2006) Modeling response signal and response time data. *Cogn. Psychol.* 53, 195–237
18. Rinkenauer, G. *et al.* (2004) On the locus of speed–accuracy tradeoff in reaction time: inferences from the lateralized readiness potential. *J. Exp. Psychol. Gen.* 133, 261–282
19. Starns, J.J. *et al.* (2012) Evaluating the unequal-variability and dual-process explanations of zROC slopes with response time data and the diffusion model. *Cogn. Psychol.* 64, 1–34
20. Rae, B. *et al.* (2014) The hare and the tortoise: emphasizing speed can change the evidence used to make decisions. *J. Exp. Psychol. Learn.* 40, 1226–1243
21. Ratcliff, R. (1985) Theoretical interpretations of speed and accuracy of positive and negative responses. *Psychol. Rev.* 92, 212–225
22. Edwards, W. (1965) Optimal strategies for seeking information: models for statistics, choice reaction times, and human information processing. *J. Math. Psychol.* 2, 312–329
23. Diederich, A. and Busemeyer, J.R. (2006) Modeling the effects of payoffs on response bias in a perceptual discrimination task: threshold bound, drift rate change, or two stage processing hypothesis. *Percept. Psychophys.* 97, 51–72
24. Leite, F.P. and Ratcliff, R. (2011) What cognitive processes drive response biases? A diffusion model analysis. *Judgm. Decis. Mak.* 6, 651–687
25. Mulder, M.J. *et al.* (2012) Bias in the brain: a diffusion model analysis of prior probability potential payoff. *J. Neurosci.* 32, 2335–2343
26. Dunovan, K.E. *et al.* (2014) Prior probability and feature predictability interactively bias perceptual decisions. *Neuropsychologia* 61, 210–221
27. Hanks, T.D. *et al.* (2011) Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J. Neurosci.* 31, 6339–6352
28. Falmagne, J.C. (1965) Stochastic models for choice reaction time with applications to experimental results. *J. Math. Psychol.* 12, 77–124
29. Falmagne, J.C. (1968) Note on a simple property of binary mixtures. *Br. J. Math. Stat. Psychol.* 21, 131–132
30. Ollman, R.T. (1966) Fast guesses in choice reaction time. *Psychon. Sci.* 6, 155–156
31. Remington, R.J. (1969) Analysis of sequential effects in choice reaction times. *J. Exp. Psychol.* 82, 250–257
32. Luce, R.D. (1986) *Response Times,* Oxford University Press
33. Kirby, N.H. (1980) Sequential effects in choice reaction time. In *Reaction Times* (Welford, A.T., ed.), pp. 129–172, Academic Press
34. Laming, D.R.J. (1969) Subjective probability in choice-reaction experiments. *J. Math. Psychol.* 6, 81–120

Outstanding Questions

What is the best architecture for multiple choice decisions and confidence judgments? There are a number of models but they are complicated and difficult to test against data and against one another.

What is the signal that initiates the decision process? One assumption is that there is a release from inhibition in the accumulation process, but this needs to be formally modeled. A change detector based on the perceptual signal has been used to model this.

How are decision boundaries set (criterion settings in general such as the criterion in signal detection theory and the drift criterion in the diffusion model)? Optimality theory attempts to do this but no model is completely satisfactory.

The issue of criterion setting is even more of a puzzle because humans can set criteria to verbal instructions in one or two trials. In other words, there is no chance for feedback (in some experiments with human subjects, no feedback is given) to provide information with which to adjust criteria.

Much of the neurophysiological study of decision making uses tasks with responses in different locations in a retinotopic map for eye movement responses or a motor map for motor responses. However, in humans, it is easy to make the alternative responses two arbitrarily chosen words ('one'/'two', 'case'/'plumb', 'cricket'/'football'). A diffusion model could represent the decision process, but the process by which the categorical output is mapped onto the verbal response (such as by using a dynamical speech production model) is unexplored.

The relationship of criterion settings to control structures in basal ganglia and working memory processes in frontal cortex is beginning to be explored.

We are a long way away from truly integrated models of motor processes in decision making in the motor cortex and oculomotor system.

Applications to economic decision-making tasks are beginning to penetrate the field of economics.

Applications that bring this type of cognitive modeling to neuropsychological and educational testing are just beginning.

35. Rabbitt, P. and Rodgers, B. (1977) What does a man do after he makes an error? An analysis of response programming. *Q. J. Exp. Psychol.* 29, 727–743
36. Vickers, D. (1978) An adaptive module of simple judgements. In *Attention and Performance (VII)* (Requin, J., ed.), pp. 599–618, Erlbaum
37. Ratcliff, R. (2002) A diffusion model account of reaction time and accuracy in a two choice brightness discrimination task: fitting real data and failing to fit fake but plausible data. *Psychon. Bull. Rev.* 9, 278–291
38. Ratcliff, R. and Smith, P.L. (2010) Perceptual discrimination in static and dynamic noise: the temporal relation between perceptual encoding and decision making. *J. Exp. Psychol. Gen.* 139, 70–94
39. Ratcliff, R. (2013) Parameter variability and distributional assumptions in the diffusion model. *Psychol. Rev.* 120, 281–292
40. Brown, S.D. and Heathcote, A.J. (2008) The simplest complete model of choice response time: linear ballistic accumulation. *Cogn. Psychol.* 57, 153–178
41. Milosavljevic, M. *et al.* (2010) The Drift Diffusion Model can account for the accuracy and reaction times of value-based choice under high and low time pressure. *Judgm. Dec. Mak.* 5, 437–449
42. Ratcliff, R. *et al.* (2009) Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6539–6544
43. Bode, S. *et al.* (2012) Predicting perceptual decisions from early brain activity. *J. Neurosci.* 32, 12488–12498
44. Palmer, J. *et al.* (2005) The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J. Vis.* 5, 376–404
45. Usher, M. and McClelland, J.L. (2001) The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592
46. Shadlen, M.N. and Kiani, R. (2013) Decision making as a window on cognition. *Neuron* 80, 791–806
47. Reed, A.V. (1973) Speed–accuracy trade-off in recognition memory. *Science* 181, 574–576
48. Schouten, J.F. and Bekker, J.A.M. (1967) Reaction time and accuracy. *Acta Psychol.* 27, 143–153
49. Wickelgren, W.A. (1977) Speed–accuracy tradeoff and information processing dynamics. *Acta Psychol.* 41, 67–85
50. Wickelgren, W.A. *et al.* (1980) Priming and retrieval from short-term memory: a speed accuracy trade-off analysis. *J. Verbal Learn. Verbal Behav.* 19, 387–404
51. Bennur, S. and Gold, J.I. (2011) Distinct representations of a perceptual decision and the associated oculomotor plan in the monkey lateral intraparietal area. *J. Neurosci.* 31, 913–921
52. Brunton, B.W. *et al.* (2013) Rats and humans can optimally accumulate evidence for decision-making. *Science* 340, 95–98
53. Hanks, T.D. *et al.* (2015) Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* 520, 220–223
54. Ratcliff, R. (1980) A note on modelling accumulation of information when the rate of accumulation changes over time. *J. Math. Psychol.* 21, 178–184
55. Ratcliff, R. (1988) Continuous versus discrete information processing: modeling the accumulation of partial information. *Psychol. Rev.* 95, 238–255
56. Dosher, B.A. (1984) Discriminating preexperimental (semantic) from learned (episodic) associations: a speed–accuracy study. *Cogn. Psychol.* 16, 519–555
57. Gronlund, S.D. and Ratcliff, R. (1989) The time-course of item and associative information: implications for global memory models. *J. Exp. Psychol. Learn.* 15, 846–858
58. Ratcliff, R. and McKoon, G. (1982) Speed and accuracy in the processing of false statements about semantic information. *J. Exp. Psychol. Learn.* 8, 16–36
59. Ratcliff, R. and McKoon, G. (1989) Similarity information versus relational information: differences in the time course of retrieval. *Cogn. Psychol.* 21, 139–155
60. Sullivan, N. *et al.* (2015) Dietary self-control is related to the speed with which health and taste attributes are processed. *Psychol. Sci.* 26, 122–134
61. Friedman, J. *et al.* (2013) Linking cognitive and reaching trajectories via intermittent movement control. *J. Math. Psychol.* 57, 140–151
62. Hasher, L. and Zacks, R.T. (1988) Working memory, comprehension and aging: a review and a new view. In *The Psychology of Learning and Motivation* (Vol. 22) (Bower, G.H., ed.), In pp. 193–225, Academic Press
63. Rush, B.K. *et al.* (2006) Accounting for cognitive aging: context processing, inhibition or processing speed? *Neuropsychol. Dev. Cogn. B Aging* 13, 588–610
64. Simmonds, D.J. *et al.* (2008) Meta-analysis of go/no-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia* 46, 224–232
65. Gomez, P. *et al.* (2007) A model of the go/no-go task. *J. Exp. Psychol. Gen.* 136, 347–369
66. Ratcliff, R. *et al.* Modeling the go/no-go task. *Decision* (in press)
67. Ossmy, O. *et al.* (2013) The timescale of perceptual evidence integration can be adapted to the environment. *Curr. Biol.* 23, 981–986
68. Glaze, C.M. *et al.* (2015) Normative evidence accumulation in unpredictable environments. *Elife* 4, e08825
69. Ratcliff, R. and Frank, M.J. (2012) Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neurocomputational and diffusion models. *Neural Comput.* 24, 1186–1229
70. Hubner, R. *et al.* (2010) A dual-stage two-phase model of selective attention. *Psychol. Rev.* 117, 759–784
71. White, C.N. *et al.* (2011) Diffusion models of the Flanker task: discrete versus gradual attentional selection. *Cogn. Psychol.* 63, 210–238
72. Bogacz, R. *et al.* (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced choice tasks. *Psychol. Rev.* 113, 700–765
73. Moran, R. (2015) Optimal decision making in heterogeneous and biased environments. *Psychon. Bull. Rev.* 22, 38–53
74. Deneve, S. (2012) Making decisions with unknown sensory reliability. *Front. Neurosci.* 6, 75
75. Drugowitsch, J. *et al.* (2012) The cost of accumulating evidence in perceptual decision making. *J. Neurosci.* 32, 3612–3628
76. Gold, J.I. and Shadlen, M.N. (2002) Banburismus and the brain decoding the relationship between sensory stimuli decisions reward. *Neuron* 36, 299–308
77. Frazier, P. and Yu, A.J. (2008) Sequential hypothesis testing under stochastic deadlines. *Adv. Neural Info. Proc. Sys.* 20, 465–472
78. Zhang, S. *et al.* (2014) Time-varying boundaries for diffusion models of decision making and response time. *Front. Psychol.* 5, 1364
79. Hawkins, G. *et al.* (2012) An optimal adjustment procedure to minimize experiment time in decisions with multiple alternatives. *Psychon. Bull. Rev.* 19, 339–348
80. Hawkins, G. *et al.* (2012) Context effects in multi-alternative decision making: empirical data and a Bayesian model. *Cogn. Sci.* 36, 498–516
81. Simen, P. *et al.* (2009) Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *J. Exp. Psychol. Hum.* 35, 1865–1897
82. Ditterich, J. (2006) Computational approaches to visual decision making. In *Percept, Decision, Action: Bridging the Gaps* (Chadwick, D.J. *et al.*, eds), p. 114, Wiley
83. Ditterich, J. (2006) Stochastic models of decisions about motion direction: behavior and physiology. *Neural Netw.* 19, 981–1012
84. Bowman, N.E. *et al.* (2012) Temporal integration of olfactory perceptual evidence in human orbitofrontal cortex. *Neuron* 15, 916–927
85. Thura, D. *et al.* (2012) Decision-making by urgency-gating: theory and experimental support. *J. Neurophysiol.* 108, 2912–2930
86. Hawkins, G.E. *et al.* (2015) Discriminating evidence accumulation from urgency signals in speeded decision making. *J. Neurophysiol.* 114, 40–47
87. Cisek, P. *et al.* (2009) Decisions in changing conditions: the urgency-gating model. *J. Neurosci.* 29, 11560–11571

88. Roitman, J.D. and Shadlen, M.N. (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489
89. Wald, A. (1947) Note on the consistency of the maximum likelihood estimate. *Ann. Math. Stat.* 20, 595–601
90. Holmes, P. and Cohen, J.D. (2014) Optimality and some of its discontents: successes and shortcomings of existing models for binary decisions. *Top. Cogn. Sci.* 6, 258–278
91. Summerfield, C. and Tsetsos, K. (2015) Do humans make good decisions? *Trends Cogn. Sci.* 19, 27–34
92. Vickers, D. *et al.* (1971) Discriminating between the frequency of occurrence of two alternative events. *Acta Psychol.* 35, 151–172
93. de Gardelle, V. and Summerfield, C. (2011) Robust averaging during perceptual judgment. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13341–13346
94. Smith, P.L. and Vickers, D. (1989) Modeling evidence accumulation with partial loss in expanded judgment. *J. Exp. Psychol. Hum.* 15, 797–815
95. Cheadle, S. *et al.* (2014) Adaptive gain control during human perceptual choice. *Neuron* 81, 1429–1441
96. Yang, T. and Shadlen, M.N. (2007) Probabilistic reasoning by neurons. *Nature* 447, 1075–1080
97. Pietsch, A. and Vickers, D. (1997) Memory capacity and intelligence: novel techniques for evaluating rival models of a fundamental information processing mechanism. *J. Gen. Psychol.* 124, 229–339
98. Smith, P.L. and Vickers, D. (1988) The accumulator model of two-choice discrimination. *J. Math. Psychol.* 32, 135–168
99. Churchland, A.K. *et al.* (2011) Variance as a signature of neural computations during decision making. *Neuron* 69, 818–831
100. Smith, P.L. *et al.* (2014) Modeling perceptual discrimination in dynamic noise: time-changed diffusion and release from inhibition. *J. Math. Psychol.* 59, 95–113
101. Ratcliff, R. and Rouder, J.N. (2000) A diffusion model account of masking in letter identification. *J. Exp. Psychol. Hum.* 26, 127–140
102. Ratcliff, R. *et al.* (2007) Application of the diffusion model to two-choice tasks for adults 75-90 years old. *Psychol. Aging* 22, 56–66
103. Smith, P.L. *et al.* (2004) Attention orienting and the time course of perceptual decisions: response time distributions with masked and unmasked displays. *Vis. Res.* 44, 1297–1320
104. Smith, P.L. and Ratcliff, R. (2009) An integrated theory of attention and decision making in visual signal detection. *Psychol. Rev.* 116, 283–317
105. Huk, A.C. and Shadlen, M.N. (2005) Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* 25, 10420–10436
106. Barlow, H.B. (1958) Temporal and spatial summation in human vision at different background intensities. *J. Physiol.* 141, 337–350
107. Watamaniuk, S.N.J. and Sekuler, R. (1992) Temporal and spatial integration in dynamic random-dot stimuli. *Vis. Res.* 32, 2341–2347
108. Watson, A.B. (1979) Probability summation over time. *Vis. Res.* 19, 515–522
109. Zariwala, H.A. *et al.* (2013) The limits of deliberation in a perceptual decision task. *Neuron* 78, 339–351
110. Gold, J.I. and Shadlen, M.N. (2003) The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J. Neurosci.* 23, 632–651
111. Kelly, S.P. and O'Connell, R.G. (2013) Internal and external influences on the rate of sensory evidence accumulation in the human brain. *J. Neurosci.* 33, 19434–19441
112. O'Connell, R.G. *et al.* (2012) A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat. Neurosci.* 15, 1729–1730
113. Thura, D. and Cisek, P. (2014) Deliberation and commitment in the premotor and primary motor cortex during dynamic decision-making. *Neuron* 81, 1401–1416
114. Winkel, J. *et al.* (2014) Early evidence affects later decisions: why evidence accumulation is required to explain response time data. *Psychon. Bull. Rev.* 21, 777–784
115. White, C.N. *et al.* (2010) Using diffusion models to understand clinical disorders. *J. Math. Psychol.* 54, 39–52
116. McKoon, G. and Ratcliff, R. (2012) Aging and IQ effects on associative recognition and priming in item recognition. *J. Mem. Lang.* 66, 416–437
117. McKoon, G. and Ratcliff, R. (2013) Aging and predicting inferences: a diffusion model analysis. *J. Mem. Lang.* 68, 240–254
118. Ratcliff, R. (2008) The EZ diffusion method: Too EZ? *Psychon. Bull. Rev.* 15, 1218–1228
119. Ratcliff, R. (2014) Measuring psychometric functions with the diffusion model. *J. Exp. Psychol. Hum.* 40, 870–888
120. Schmiedek, F. *et al.* (2007) Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *J. Exp. Psychol. Gen.* 136, 414–429
121. Ratcliff, R. *et al.* (2010) Individual differences, aging, and IQ in two-choice tasks. *Cogn. Psychol.* 60, 127–157
122. Ratcliff, R. *et al.* (2011) Effects of aging and IQ on item and associative memory. *J. Exp. Psychol. Gen.* 140, 464–487
123. Ratcliff, R. *et al.* (2015) Modeling individual differences in response time and accuracy in numeracy. *Cognition* 137, 115–136
124. Voss, A. and Voss, J. (2007) Fast-dm: a free program for efficient diffusion model analysis. *Behav. Res. Meth.* 39, 767–775
125. Vandekerckhove, J. and Tuerlinckx, F. (2007) Fitting the Ratcliff diffusion model to experimental data. *Psychon. Bull. Rev.* 14, 1011–1026
126. Wiecki, T.V. *et al.* (2013) HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Front. Neuroinform.* 7, 1–10
127. Diederich, A. and Busemeyer, J.R. (2003) Simple matrix methods for analyzing diffusion models of choice probability, choice response time and simple response time. *J. Math. Psychol.* 47, 304–322
128. Ratcliff, R. and Childers, R. (2015) Individual differences and fitting methods for the two-choice diffusion model. *Decision* 2, 237–279
129. Voss, A. *et al.* (2013) Diffusion models in experimental psychology. *Exp. Psychol.* 60, 385–402
130. Ratcliff, R. *et al.* (2005) Aging and response times: a comparison of sequential sampling models. In *Measuring the Mind: Speed, Control, and Age* (Duncan, J. *et al.*, eds), pp. 3–32, Oxford University Press
131. Donkin, C. *et al.* (2011) Diffusion versus linear ballistic accumulation: different models for response time, same conclusions about psychological mechanisms? *Psychon. Bull. Rev.* 55, 140–151
132. Nosofsky, R.M. *et al.* (2011) Short-term memory scanning viewed as exemplar-based categorization. *Psychol. Rev.* 118, 280–315
133. Ratcliff, R. *et al.* (2004) A diffusion model account of the lexical decision task. *Psychol. Rev.* 111, 159–182
134. Thornton, T.L. and Gilden, D.L. (2007) Parallel and serial processes in visual search. *Psychol. Rev.* 114, 71–103
135. Purcell, B.A. *et al.* (2012) From salience to saccades: multiple-alternative gated stochastic accumulator model of visual search. *J. Neurosci.* 32, 3433–3446
136. Nosofsky, R.M. and Palmeri, T.J. (1997) An exemplar based random walk model of speeded classification. *Psychol. Rev.* 104, 266–300
137. Boucher, L. *et al.* (2007) Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol. Rev.* 114, 376–397
138. Logan, G.D. *et al.* (2014) On the ability to inhibit thought and action: general and special theories of an act of control. *Psychol. Rev.* 121, 66–95
139. Frank, M.J. *et al.* (2015) fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. *J. Neurosci.* 35, 484–494

140. Klauer, K.C. *et al.* (2007) Process components of the Implicit Association Test: a diffusion-model analysis. *J. Pers. Soc. Psychol.* 93, 353–368
141. van Ravenzwaaij, D. *et al.* (2011) Does the Name–Race Implicit Association Test measure racial prejudice? *Exp. Psychol.* 58, 271–277
142. Cavanagh, J.F. *et al.* (2014) Eye tracking and pupillometry are indicators of dissociable latent decision processes. *J. Exp. Psychol. Gen.* 143, 1476–1488
143. Hanes, D.P. and Schall, J.D. (1996) Neural control of voluntary movement initiation. *Science* 274, 427–430
144. Gold, J.I. and Shadlen, M.N. (2001) Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16
145. Ratcliff, R. *et al.* (2003) A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of simple two-choice decisions. *J. Neurophysiol.* 90, 1392–1407
146. Kepecs, A. *et al.* (2008) Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455, 227–231
147. Kiani, R. and Shadlen, M. (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324, 759–764
148. Kiani, R. *et al.* (2008) Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017–3029
149. Philiastides, M.G. *et al.* (2006) Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *J. Neurosci.* 26, 8965–8975
150. Polania, R. *et al.* (2014) Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron* 82, 709–720
151. Mulder, M. *et al.* (2014) Perceptual decision neurosciences – a model-based review. *Neuroscience* 277, 872–884
152. Lim, S.L. *et al.* (2011) The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J. Neurosci.* 31, 13214–13223
153. Hare, T.A. *et al.* (2011) Transformation of stimulus value signals into motor commands during simple choice. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18120–18125
154. Heekeren, H.R. *et al.* (2008) The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* 9, 467–479
155. Basten, U. *et al.* (2010) How the brain integrates costs and benefits during decision making. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21767–21772
156. Philiastides, M.G. *et al.* (2011) Causal role of dorsolateral prefrontal cortex in human perceptual decision making. *Curr. Biol.* 21, 980–983
157. Polania, R. *et al.* (2015) The precision of value-based choices depends causally on fronto-parietal phase coupling. *Nat. Commun.* 6, 8090
158. Passino, K.M. *et al.* (2008) Swarm cognition in honey bees. *Behav. Ecol. Sociobiol.* 62, 401–414
159. Marshall, J.A.R. *et al.* (2009) On optimal decision making in brains and social insect colonies. *J. Roy. Soc. Interf.* 6, 1065–1074
160. Latty, T. and Beekman, M. (2011) Irrational decision-making in an amoeboid organism: transitivity and context-dependent preferences. *Proc. Roy. Soc. B* 278, 307–312
161. Ratcliff, R. and Van Dongen, H.P.A. (2009) Sleep deprivation affects multiple distinct cognitive processes. *Psychon. Bull. Rev.* 16, 742–751
162. Geddes, J. *et al.* (2010) Modeling the effects of hypoglycemia on a two-choice task in adult humans. *Neuropsychology* 24, 652–660
163. van Ravenzwaaij, D. *et al.* (2012) A diffusion model decomposition of the effects of alcohol on perceptual decision making. *Psychopharmacology* 219, 1017–1025
164. Ratcliff, R. *et al.* (2004) A diffusion model account of normal and impaired readers. *Brain Cogn.* 55, 374–382
165. Ratcliff, R. *et al.* (2012) Children are not like older adults: a diffusion model analysis of developmental changes in speeded responses. *Child Dev.* 83, 367–381
166. McKoon, G. and Ratcliff, R. (2016) Adults with poor reading skills: how lexical knowledge interacts with scores on standardized reading comprehension tests. *Cognition* 146, 453–469
167. Zeguers, M.H.T. *et al.* (2011) Specifying theories of developmental dyslexia: a diffusion model analysis of word recognition. *Dev. Sci.* 14, 1340–1354
168. Mulder, M.J. *et al.* (2010) Basic impairments in regulating the speed–accuracy tradeoff predict symptoms of attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 68, 1114–1119
169. Weigard, A. and Huang-Pollock, C. (2014) A diffusion modeling approach to understanding contextual cueing effects in children with ADHD. *J. Child Psychol. Psychiatry* 55, 1336–1344
170. Moustafa, A.A. *et al.* (2015) Drift diffusion model of reward and punishment learning in schizophrenia: modeling and experimental data. *Behav. Brain Res.* 291, 147–154
171. Tuerlinckx, F. *et al.* (2001) A comparison of four methods for simulating the diffusion process. *Behav. Res. Instr.* 33, 443–456
172. White, C.N. *et al.* (2009) Dysphoria and memory for emotional material: a diffusion model analysis. *Cogn. Emot.* 23, 181–205
173. Smith, P.L. and Ratcliff, R. (2004) The psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168
174. Wong, K-F. and Wang, X-J. (2006) A recurrent network mechanism for time integration in perceptual decisions. *J. Neurosci.* 26, 1314–1328
175. Rustichini, A. and Padoa-Schioppa, C. (2015) A neuro-computational model of economic decisions. *J. Neurophysiol.* 14, 1382–1398
176. Roxin, A. and Ledberg, A. (2008) Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comput. Biol.* 4, e1000046
177. Verdonck, S. and Tuerlinckx, F. (2014) The Ising Decision Maker: a binary stochastic network for choice response time. *Psychol. Rev.* 121, 422–462
178. Smith, P.L. (2010) 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.* 54, 266–283
179. Smith, P.L. and McKenzie, C.R.L. (2011) Diffusive information accumulation by minimal recurrent neural models of decision-making. *Neural Comput.* 23, 2000–2031
180. Audley, R.J. and Pike, A.R. (1965) Some alternative stochastic models of choice. *Br. J. Math. Stat. Psychol.* 18, 207–225
181. Roe, R.M. *et al.* (2001) Multialternative decision field theory: a dynamic connectionist model of decision-making. *Psychol. Rev.* 108, 370–392
182. McMillen, T. and Holmes, P. (2006) The dynamics of choice among multiple alternatives. *J. Math. Psychol.* 50, 30–57
183. Leite, F.P. and Ratcliff, R. (2010) Modeling reaction time and accuracy of multiple-alternative decisions. *Atten. Percept. Psychophys.* 72, 246–273
184. Niwa, M. and Ditterich, J. (2008) Perceptual decisions between multiple directions of visual motion. *J. Neurosci.* 28, 4435–4445
185. Purcell, B.A. *et al.* (2010) Neurally-constrained modeling of perceptual decision making. *Psychol. Rev.* 117, 1113–1143
186. Ratcliff, R. and Starns, J.J. (2013) Modeling response times, choices, and confidence judgments in decision making: recognition memory and motion discrimination. *Psychol. Rev.* 120, 697–719
187. Usher, M. and McClelland, J.L. (2004) Loss aversion and inhibition in dynamical models of multi-alternative choice. *Psychol. Rev.* 111, 757–769
188. Usher, M. *et al.* (2002) Hick's law in a stochastic race model with speed–accuracy tradeoff. *J. Math. Psychol.* 46, 704–715
189. Zandbelt, B. *et al.* (2014) Response times from ensembles of accumulators. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2848–2853
190. Egan, J.P. (1958) *Recognition Memory and the Operating Characteristic,* Technical note AFCRC-TN-58-51, Hearing and Communication Laboratory, Indiana University
191. Murdock, B.B. (1974) *Human Memory: Theory and Data,* Erlbaum
192. Green, D.M. and Swets, J.A. (1966) *Signal Detection Theory and Psychophysics,* Robert E. Krieger Publishing Company

193. Murdock, B.B. and Anderson, R.E. (1975) Encoding, storage, and retrieval of item information. In *Information Processing and Cognition: The Loyola Symposium* (Solso, R.L., ed.), pp. 145–194, Lawrence Erlbaum Associates
194. Norman, D.A. and Wickelgren, W.A. (1969) Strength theory of decision rules and latency in short-term memory. *J. Math. Psychol.* 6, 192–208
195. Pleskac, T.J. and Busemeyer, J. (2010) Two-stage dynamic signal detection: a theory of confidence, choice, and response time. *Psychol. Rev.* 117, 864–901
196. Ratcliff, R. and Starns, J.J. (2009) Modeling confidence and response time in recognition memory. *Psychol. Rev.* 116, 59–83
197. Jones, M. and Dzhafarov, E.N. (2014) Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. *Psychol. Rev.* 121, 1–32
198. Smith, P.L. *et al.* (2014) The diffusion model is not a deterministic growth model: comment on Jones and Dzhafarov (2014). *Psychol. Rev.* 121, 679–688
199. Starns, J.J. and Ratcliff, R. (2010) The effects of aging on the speed–accuracy compromise: boundary optimality in the diffusion model. *Psychol. Aging* 5, 377–390
200. Starns, J.J. and Ratcliff, R. (2012) Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychon. Bull. Rev.* 19, 139–145
201. Busemeyer, J.R. and Townsend, J.T. (1993) Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychol. Rev.* 100, 432–459
202. Tsetsos, K. *et al.* (2010) Preference reversal in multi-attribute choice. *Psychol. Rev.* 117, 1275–1291
203. Johnson, J.G. and Busemeyer, J.R. (2005) A dynamic, computational model of preference reversal phenomena. *Psychol. Rev.* 112, 841–861
204. Dai, J. and Busemeyer, J.R. (2014) A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *J. Exp. Psychol. Gen.* 143, 1489–1514
205. Krajbich, I. *et al.* (2010) Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13, 1292–1298
206. Krajbich, I. and Rangel, A. (2011) A multi-alternative drift diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13852–13857
207. Krajbich, I. *et al.* (2015) A common mechanism underlying food choice and social decisions. *PLoS Comput. Biol.* 11, e1004371
208. Krajbich, I. *et al.* (2012) The attentional drift-diffusion model extends to simple purchasing decisions. *Front. Psychol.* 3, 1–18
209. Hutcherson, C. *et al.* (2015) A neurocomputational model of altruistic choice and its implications. *Neuron* 87, 451–462
210. Gottlieb, J. *et al.* (2014) Attention, reward and information seeking. *J. Neurosci.* 34, 15497–15504