Analysis of speed-accuracy tradeoffs using the EZ-diffusion model for choice reaction times

Annalise D'Souza

Abstract

Reaction times (RTs) are a ubiquitous variable in Psychology. Most RT experiments primarily test mean speed, with accuracy as a separate secondary analysis. However, speed and accuracy are related in a non-linear way - participants often show tradeoffs between the speed and accuracy at which they perform. Multiple recent studies have shown that diffusion models can be successfully applied to RT distributions, to model parameters such as the speed-accuracy trade-off. An adapted version of the diffusion model for RTs, the EZ-diffusion model, has been found to offer a practical and feasible way to model RT data using the R statistical programming environment (Wagenmakers, van der Maas, & Grasman, 2007).

The current paper focused on modeling speed-accuracy tradeoffs in RT data, using the EZ-diffusion model on data for a two choice RT paradigm. Data were simulated using the RWiener package to model distribution functions of the Wiener diffusion process. The utility of the EZ-diffusion model for speed-accuracy data is discussed.

In human decision-making, the time taken to make a choice is a key dependent variable. Choice reaction time (RT) experiments are widely used in experimental and cognitive psychology, with human performance typically measured by accuracy and RT. Analysis of these two response measures is usually done separately, which implicitly assumes that they are distinct. However, a main insight from the literature on choice RTs is the two measures are linked (Luce, 1986; Van Zandt, 2000; Whelan, 2008). Subjects can choose to focus on either RT or accuracy at the cost of the other measure, termed the speed-accuracy tradeoff (Pachella, & Pew, 1968; Pew, 1969; Wickelgren, 1977; Bogacz, 2014). The present paper considers the speed-accuracy tradeoff in terms of the diffusion model, one of the most successful models of choice RT performance (Ratcliff & Smith, 2004; Wagenmakers, 2009). The paper presents details of the model, focusing on its use in simultaneously analyzing RT and accuracy under a theoretical framework.

Speed-accuracy tradeoffs

The need for a model of RT rises from the many interpretive issues that speed-accuracy tradeoffs pose in choice RT experiments (Pachella & Pew, 1968; Wood & Jennings, 1976). Consider a situation in which a participant is responding in two experimental conditions, A and B (Wagenmakers, Van Der Maas, and Grasman, 2007). The participant responds faster in condition A, but also makes more errors. Using conventional methods of independently analyzing the mean RT and accuracy, it is not possible to say on which condition the subject responded better. More importantly, variations in a participants' criterion for speed versus accuracy will bias RT analyses since they influence RT separately from the independent variable that is being manipulated in the experiment (Jennings, Wood, & Lawrence, 1976). Separating RT and accuracy also leads to researchers focusing on one measure at the expense of ignoring the other (Whelan, 2008).

Analyzing the speed-accuracy tradeoff is usually done in one of two ways (Heitz, 2014).

Speed-accuracy tradeoff functions (SATF) could be calculated and used as dependent variables rather than looking at RT and accuracy individually (Pachella, 1974; Pew, 1969; Swensson, 1972b; Heitz, 2014). SATFs enable testing whether subjects' criterion for speed and accuracy is influenced by different manipulations (Wood & Jennings, e.g., Jennings, Wood, & Lawrence). This process is analogous to calculation of the receiver operating characteristic (ROC) curve in a signal detection experiment (Green & Swets, 1966), and has been termed the LOC or latency operating characteristic (Lappin & Disch, 1972). Obtaining these functions requires multiple condition designs where speed and accuracy are systematically varied. Manipulations include a response-signal paradigm to force the time when a response must be made, or explicit instructions to induce changes in subjects' criteria for speed or accuracy. Another type of function, the conditional accuracy function (CAF), can be obtained from a single condition design by calculating the conditional probability of a correct response for each RT value.

A more comprehensive way of examining the speed-accuracy tradeoff is by using models of the tradeoff process (Vickers et al., 1985; Heitz, 2014). Going back to the example presented earlier, consider a third condition C in which the accuracy is the same as B, but the RT is slower. An explanation using only the speed-accuracy tradeoff will have trouble accounting for such results quantitatively, although it could be use to confirm ordinally that the participant is performing better in condition B than C. Human performance is also influenced by variables other than the speed-accuracy tradeoff, such as stimulus difficulty. Models offer the advantage of simultaneously analyzing multiple predictors of performance and isolating each factor such as the speed-accuracy tradeoff. Second, they offer more meaningful interpretation of data as they go beyond observed variables (RT and accuracy) to make predictions on the unobserved psychological processes that produced the data. Models are also more suited to RT data as they

account for its non-linearity by using the entire RT distributions (instead of just the mean in standard analysis, Whelan, 2008).

Sequential sampling models

This paper focuses on a sequential sampling model and how it applies to the analysis of the speed-accuracy tradeoff. Sequential sampling models were developed based on the statistical literature for the time course of decision-making, termed sequential analysis (Wald, 1947), a key insight from which is that individuals aim to make accurate decisions quickly (Navarro & Fuss, 2009). Out of all the models designed for simple two-choice decisions, sequential sampling models are unique in that they enable the study of both speed and accuracy under a shared theoretical framework (Ratcliff & Smith, 2004). Thus they are ideal for the analysis of a speed-accuracy tradeoff, since they account for the tradeoff and the underlying cognitive processes.

The central tenet of sequential sampling models is that a noisy process accumulates information over time, from a starting point toward one of two response criteria or boundaries (Smith & Ratcliff, 2004). The noisy information accumulation process originates from physics (Einstein, 1905), and can be for one-dimensional discrete data (random walk models) or continuous data (Wiener diffusion or Brownian motion; Ma, Krings, & Millar, 2009). An example of a discrete-time random walk model is flipping a coin until the number of heads exceeds the number of tails by a chosen number, or vice versa, illustrated in Figure 1. The average step size of the random walk or diffusion process is called drift. The process is characterized by the systematic drift component plus normally distributed random noise. The assumption of a noise component accounts for the occurrence of errors, thus offering a way to simultaneously determine both speed of information accumulation and accuracy.

As Figure 1 demonstrates, the diffusion process terminates once the information for one boundary exceeds that of another. This produces two main parameters for the models, drift and

boundary separation. Boundary separation is a property of the decider, and quantifies a system's criterion on the amount of information necessary to make a response. For psychological applications, it characterizes a subject's speed-accuracy tradeoff as low separation prioritizes speed while large separation prioritizes accuracy.

Ratcliff's drift diffusion model (Ratcliff, 1978, 1980, 2002, 2008; Ratcliff & Smith, 2004; Figure 2) is the most influential and widely applied sequential sampling model (Rae et al., 2013). The core of the model uses the Wiener diffusion process. To fit empirical choice RT data, Ratcliff added parameters to assume that the starting point is also a random variable, and that RTs consist of both decision and non-decision times. The model comprehensively explains RT performance by using complete RT distributions for correct and incorrect responses.

The choice of the Wiener diffusion model for this paper was based on the straightforward approach that sequential sampling models offer for analyzing both speed and accuracy (Ratcliff and Smith, 2004), and the diffusion model's extensive success in fitting experimental data from choice RT paradigms (Wagenmakers, 2009; Ratcliff & McKoon, 2008; Vanderckhove & Tuerlinckx, 2007). However, in cases where the speed-accuracy tradeoff is not of primary focus, other RT models are available, including neural based approaches, and can address some of the limitations of sequential sampling models (van Zandt, 2000; Dutilh et al, 2011). Within the family of sequential sampling models itself there are many alternatives for varying experimental situations. Examples are the Ornstein–Uhlenbeck diffusion model for information decay, or accumulator and counter models for two evidence totals (refer Figure 3 for a full taxonomy). Note that although these models have distinctive components, they tend to concur on main insights (Donkin, Brown, Heathcote & Wagenmakers, 2011).

Model assumptions

The aims of the model are to quantify performance in simple two choice decisions. This produces several theoretical assumptions. Since the model is for binary decisions, it doesn't apply to paradigms that have multiple responses (unless these assume a common underlying process and can be recoded as correct or incorrect). The model assumes that paradigms have a single-stage of decision-making, so doesn't apply to multiple-stage processes (ex. reasoning tasks; Ratcliff & McKoon, 2008). Related to this assumption is a need for constant parameter values over time. Based on the previous assumptions, the model applies to fast two-choice decisions with mean RTs smaller than 1500 milliseconds (Ratcliff & Rouder, 1998), as longer responses include unspecified additional stages of processing. Since the imposed RT limit is arbitrary, the model may apply to longer RTs although these must be tested (Voss et al., 2013). It is also possible to include some task violations in the model (ex. post-error slowing, Dutilh et al., 2012).

The assumptions of the general class of sequential sampling models also apply to diffusion models. The first is that a decision is made once activity reaches a certain threshold. The most common approach taken in sequential models is that the decision process is optimal when the decision-maker uses a speed-accuracy tradeoff to determine the termination criterion (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006). The second assumption is that the sampling process is quick but with considerable noise. Third, a continuous sampling process occurs, where information is sampled in sequence, with new information added to old.

Wiener diffusion process

The model assumes a stochastic diffusion process, as demonstrated in the following equation

$$DX(t) = vdt + sdW(t) \quad [1]$$

In the above equation, X is accumulated evidence, dt is a small time interval, $DX(t)$ is the change in X for dt, $W(t)$ is the Wiener noise process (idealized Brownian motion), and $sdW(t)$ are zero-mean random increments with infinite variance[1] s^2dt (Wagenmakers, 2009)[1].

Approximating the diffusion process is challenging since it is a stochastic process with a continuous time and state space. Challenges include an infinite oscillating series when calculating the cumulative distribution function (cdf) and probability density function (pdf), and integrals without closed-forms when parameters are allowed to vary across trials (Tuerlinckx, 2004). To address this, calculations usually result in a discretization of the time or state. In the case of the continuous-time diffusion model, this is done by deriving the model as a limiting version of the discrete random walk model (Ratcliff & McKoon, 2008; Smith, 2000). Several discretization methods for constructing the model were compared using simulations, and their accuracy evaluated using comparison methods with no discretization (Tuerlinckx et al., 2001). The two best methods for simulation speed and accuracy were a rejection-based method and a probability integral transform method.

Wiener first passage time

Computing the diffusion process is numerically difficult, which has lead to a lack of using the model by experimental psychologists (Ratcliff & Tuerlinckx, 2002). One solution is to compute the first passage time (FPT) for the diffusion process. For a stochastic process $X(t)$, the FPT is the time (T) when $X(t)$ first crosses a boundary (Darling & Siegart, 1953). For an initial state of evidence $X(0)$ that falls in the range $0 < X(0) < a$, with boundaries 0 and a, the FPT is specified as the time t for which a boundary is reached, i.e. $X(t) \leq 0$ or $X(t) \geq a$. When applied in the

[1] The chosen variance for the Wiener process variance, s, is arbitrary. The original model fixed it at 0.1, although recent adaptations have since used 1 (Voss, Ruthermord, & Voss, 2004; Lee, Fuss, & Navarro, 2006; Navarro & Fuss).

context of the diffusion model (right image on Figure 1), decisions and the time for a decision are considered as continuous random variables described by the Wiener FPT (Shepp, 1967).

The FPT *problem* refers to finding the pdf of *T* (Siegert, 1951), and is defined as 'a stochastic generalization of a shortest path problem' (Wakuta, 2000). Thus the FPT approach is beneficial for models based on stochastic processes such as random walks or diffusion, as it offers a way to calculate the closed-form solution of the pdf. In most applications it is easier to deal with the pdf than the stochastic process (Ma, Krings, & Millar, 2009). The resulting distribution, denoted *(c, td)* ~ WFPT(*v, a, z),* is called a Wiener first-passage time (WFPT) distribution. Equations and mathematical details of calculating the distributions for the WFPT can be found in Navarro and Fuss (2009) for the pdf, and Blurton, Kesselmeier, and Gondan (2012) for the cdf. Both distribution functions are based on the following formula (Feller, 1968) for calculating the WFPT at the lower boundary, with boundaries set at 0 and *a* such that 0< *z*< *a*,

$$f(t|v,a,w) = \frac{\pi}{a^2} e^{(-vaw-\frac{v^2t}{2})} \sum_{k=1}^{\infty} e^{(-\frac{k^2\pi^2t}{2a^2})} \sin(k\pi w) \quad [2]$$

where *w* = a/2, and t is the time when the boundary is reached. Both distribution calculations have options for small and large time representations. The distinction between performance at each time is crucial for the speed-accuracy tradeoff (ex. more errors at fast RTs).

For the purposes of psychological data (Tuerlinckx et al., 2001), the relevant aspects of the diffusion process are the boundary of absorption, i.e., whether random variable Y takes value of 1 (upper boundary) or 0 (lower boundary), and the time until absorption [*T*<*t*|*Y*=1]. Thus the full sample path is not observed, which makes computing the WFPT suitable for most RT tasks. The accompanying R script uses the WFPT rather than the full diffusion model, based on the reasoning outlined by Navarro and Fuss (2009) on the need for a simple, error-free method to compute the WFPT before estimating the full model. Further, they propose that the WFPT falls

logically within a Bayesian hierarchical characterization of the full model, and can thus be extended using Bayesian techniques (Vanderckhove, Tuerlinckx, & Lee, 2008; Rouder et al., 2003).

Wiener diffusion model parameters

The Wiener diffusion model (Ratcliff, 1978, 1979; Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004; Ratcliff & McKoon, 2008) demonstrated in Figure 2, has four parameters (Table 1) consisting of decision components of drift rate (ξ or v), threshold separation (a), starting-point (β or z), and non-decision time (τ or $T_{er,}$). The rate of accumulation of information (drift rate, v**)**, and is described by the average slope of the diffusion process. Higher drift rates produce faster and more accurate decisions. Drift rate is assumed to be deterministic, which means that it is an intrinsic property of stimulus or participants, and is not under subjective control.

The two subjective parameters are boundary separation (a) and starting point (z). Boundary separation determines the criterion for termination, and the starting point reflects a priori bias for one response boundary over another. This produces a higher speed and accuracy for the biased boundary, and a lower speed and accuracy for the other. The initial bias can be manipulated via payoffs or explicit instructions.

RTs consist of both decision-making and non-decision components

$$RT = DT + T_{er} \quad [3]$$

where DT is decision time. Observed RTs follow the order of encoding time (u), decision time (d), and response output (w) time (Luce, 1986). The nondecision component is calculated by adding u and w and calculating the average (T_{er}**)**. T_{er} doesn't depend on the nature of the stimulus, thus it shifts the entire RT distribution by a constant without affecting response choice.

Since the original four parameters don't capture all empirical data, variability in parameters across trials was added to account for varying responses (Ratcliff & McKoon, 2008). The three additional parameters added to the model are s_t, s_z, and η; for variance in non-decision

time, starting point, and drift rate, respectively. The across-trial variability is assumed to follow a uniform distribution for non-decision time and starting point, s_t, s_z, and a normal distribution for drift rate, η, although the standard model is largely robust to moderate violations in the distributional assumptions of parameter values (Ratcliff, 2013).

Boundary separation and the speed-accuracy tradeoff

Figure 4 shows how the speed-accuracy tradeoff varies as a function of *a.* High boundary separation (large *a*) demonstrates a high level of response caution. By setting the boundaries for response further apart, individuals are less likely to make errors, but at the cost of an increase in RT. Similarly, small boundary separation (small *a*) reflects fast but inaccurate decisions.

By isolating the boundary separation parameter, the diffusion model allows a way of measuring the speed-accuracy criterion and any criterion shifts (Horn, Baven, & Smith, 2011). The diffusion model's predictions for *a* are widely supported by empirical observations in the speed-accuracy tradeoff (Ratcliff & McKoon, 2008; c.f. Rae et al., 2013). Using the diffusion model also allows investigation of the tradeoff in conjunction with the other processes underlying RT and accuracy.

Extending the analogy presented earlier, the diffusion model is conceptually similar to signal detection theory in that it tests an observer's decision criterion under both a signal and noise (Ratcliff & McKoon, 2008). The chosen criterion usually reflects a tradeoff between several measures, and both models use observed data inputs to estimate the unobserved criterion; hit and false alarm rate in signal detection theory, speed and RT in the diffusion model (Wagenmakers, 2007, 2009). Unlike signal detection theory however, observer responses are based on a sequence of noisy samples not a single noisy sample (Macmillan & Creelman, 2005).

Fitting the diffusion model

Parameter estimation

RT data is extremely complex, so before fitting a model, the type of obtained data and its distribution should be considered (refer van Zandt, 2000, for an extensive discussion on fitting RT models). RT data also contains many contaminants, that is data points that are external to the processes under investigation, ex. outliers, random or fast guesses, and delays (Ratcliff, 1993; Ulrich and Miller, 1994). Fast guesses are especially relevant to the speed-accuracy tradeoff, and are an extensive modeling topic on their own (Ollman, 1966; Swensson, 1972). Researchers should develop a strategy to address contaminants before model fitting so the analysis is not biased. Several contaminant solutions have been proposed (reviews by Ratcliff, 1993; Whelan, 2008) such as using data cutoffs or transformations. A modeling solution is adding an 8th parameter to create a mixture model with two components, one for diffusion and one for contaminants (Ratcliff & Tuerlinckx, 2002).

Data from individual participants is modeled separately when estimating parameters. However, if the number of trials per subject is small, data can be combined for similar subjects (Ratcliff, Perea, Colangelo, & Buchanan, 2004). The estimation procedure involves simultaneously searching for estimates for the parameters, in order to obtain an optimal fit of predicted and obtained RT distributions. This can take considerable time since it involves computing predicted distributions for multiple parameters. Two issues that face parameter estimation are model mimicry, and some parameters having only a minimal influence on predictions (Voss et al., 2013). Model mimicry, the ability of the model to account for data from competing models, can occur when predictions from alternating models are similar. To avoid this, large empirical distributions (with large trial numbers) should be used so models can be separated sufficiently (Ratcliff & Tuerlinckx, 2002). Similarly, for parameters that have a minor influence, either a large number of trials should be used, or the parameters should be fixed to

zero (a model matrix method for restricting the number of parameters allowed to vary is outlined in Vanderckhove & Tuerlinckx, 2007).

In fitting the diffusion model, several techniques can be used, each of which vary in speed, precision, and robustness. Ratcliff & Tuerlinckx (2002) compared methods for the multidimensional search for optimal parameters estimates. They found that a maximum likelihood method outperformed chi-square and weighted least squares methods for recovering original parameter estimates, even from small data sets. The maximum likelihood technique is calculated by summing the logarithmic density of predicted RT distributions over all responses, and then maximizing the sum. However, the method was extremely sensitive to contaminant RTs, and had a slow computational speed for large datasets. A multinomial likelihood function is another option, which maximizes the likelihood of observing a certain proportion of responses in a given number of RT bins (Vanderckhove, 2008).

A chi-square approach is much more robust to contaminants and is computational quicker, thus ideal for large datasets. This approach categorizes RT distributions into quantiles, and then calculates the chi-square using observed and expected values for each quantile. Another optimization option is the Kolmogorov-Smirnov (KS) statistic (Voss & Voss, 2007; Voss et al., 2004), which reflects the maximum vertical distance between predicted and empirical RT distributions. Recommendations are to use the maximum likelihood method for small samples due to its efficiency in obtaining true parameters from small datasets, and to use the other methods with larger sample sizes or in the presence of outliers (Voss et al., 2013).

Model Fit

Once the parameters have been fitted, statistical tests can be used to assess model fit by comparing predicted to empirical RT distributions. Typical tests of model fit such as the such as

the chi-square statistic or the KS method are suitable for the diffusion model although they offer a conservative approach since the model predictions are derived from the empirical data which it is being compared with (Voss et al., 2013). One solution is to use Monte-Carlo-simulations to simulate multiple diffusion datasets using estimates from the empirical data, calculate the distribution of the fit for the simulated datasets and use it to assess fit of the empirical model. Another option that avoids the issues of statistical tests of fit tests is to graphically display data. Graphical methods include cdfs of predicted and empirical distributions, or quantile probability plots, for RT quantiles as a function of probability at each condition or stimulus level (refer Heitz, 2014, for a full discussion of plotting techniques). However, graphical displays have no clear interpretations standards of how much distance from predicted data is concerning.

A second issue is that the full model requires a large dataset to be properly fitted, to accurately calculate distributions for both correct and error responses. For the speed-accuracy tradeoff, this would ideally include experimental manipulations to calculate the whole tradeoff (Wagenmakers, 2009). However, focusing on one parameter while keeping others constant, decreases the number of total trials needed. A general issue facing fitting of any statistical model in that the model can never exactly depict reality (Myung, 2000). While a low number of trials might lack the power to detect any misfits, large numbers of trials will produce misfits that are statistically significant but not meaningful. The ideal number of trials depends on the number of parameters allowed to vary, experimental conditions, and the chosen optimization criterion. For example, maximum likelihood optimization is effective even with small samples ($N>40$), while chi-square estimation requires samples over 500 (Ratcliff & Tuerlinckx, 2002).

It should be noted that the goal of the model is to estimate underlying cognitive processes rather than to fit data. Hence when model fit is poor, a researcher must determine if the solution

is dropping few participants who have bad fits, or whether the whole model should be dismissed since its theoretical assumptions are not met by the task in question (Voss et al, 2013).

Model selection

Competing models can be compared using standard model selection techniques such as the likelihood ratio test (LRT; Wilks, 1938) for nested models, or Akaike's information criterion (Akaike, 1973) to point to preferred models based on distance to data. Caution must be exercised when using standard model selection methods however, as most define complexity by the number of free parameters. The absolute number of parameters is more informative when considered along with the amount of data being explained (Myung, Forster, & Brown, 2008), which in the case of the diffusion model is three dependent variables: RT mean, RT variance, and accuracy. Further, overparameterization is a concern when it arises from over-fitting, but in the case of a pre-specified model such as the Wiener diffusion, the parameters are based on underlying theoretical processes rather than to provide a better fit to data (Wagenmakers, 2009). To account for model flexibility, a more suitable parametric bootstrap technique has been proposed to test model adequacy (Wagenmakers et al., 2004). The method accounts for the functional form of parameters by comparing goodness-of-fit from simulated distributions for each competing model, and is also beneficial for speed-accuracy calculations.

In most applications of the diffusion model (Ratcliff & Smith, 2004; Ratcliff, 2008), only a few parameters are allowed to vary, which decreases the issue of overfitting. Model performance for selective parameter variations outperforms that of competing models (Ratcliff & Smith, 2004; Ratcliff, 2009). Thus the model is suited for manipulations of a target parameter alone, such as the boundary separation parameter when the speed-accuracy tradeoff is of interest. However recent results suggest that variations in a target cognitive process may not be

adequately captured by a single parameter, ex. while the speed-accuracy tradeoff is attributed to boundary separation, other parameters such as drift rate can also be affected (Rae et al., 2013).

Software implementations

The main strengths of the model are considerable, such as its ability to simultaneously analyze accuracy and RT, to use the full distributions for correct and error RTs, and to offer a theoretical account of the underlying processes. Despite these advantages, it is rarely implemented (Wagenmakers, 2009). The biggest limitation of the model is its complexity. Even after limiting to the FPT, the model contains many advanced computations. For example, it is difficult to account for the relationship between accuracy and RT since accuracy is binary while RT follows a range. Thus it requires the transformation and scaling of accuracy and RT relative to each other (Ratcliff, 2008). In light of this, several software solutions have been developed to make the model more accessible to researchers.

The simplest of these is the EZ diffusion model (Wagenmakers, van der Maas, & Grasman, 2007). The model contains three equations which transform MRT, VRT, and P_c to drift rate v, boundary separation a, and non-decision time T_{er}. To do so, it assumes that the starting point is unbiased ($z=a/2$), and across-trial variability (s_z, s_t, and eta) is zero. Further, it uses simple equations to calculate parameter estimates with no model fitting procedures. These assumptions have been criticized for being too simple for the targeted data (Ratcliff, 2008), but modifications to the model now allow the starting point to vary and parameter optimization (EZ2, Grasman et al., 2009) and contaminant solutions (robust-EZ, Wagenmakers et al, 2008).

Two more options, DMAT (Vanderckhove & Tuerlinckx, 2007, 2008), and fast-dm (Voss & Voss, 2008), offer more detailed calculations for the diffusion model. Unlike EZ which requires only mean RTs and accuracy as input, fast-dm and DMAT require the full dataset for error and correct RT distributions. Both offer iterative optimization for estimating the model

parameters, and use EZ values as starting estimates when fitting parameters. The optimization criterion used is the chi-square statistic in DMAT, and the Kolmogorov-Smirnov (KS) in fast-dm. DMAT also offers many other options for simulation and model selection techniques.

DMAT is a MATLAB toolbox, while fast-dm is a command-line program in C. EZ-diffusion offers options in Excel, R, Javascript, and MATLAB. Comparisons show that DMAT performs best overall, while EZ-diffusion is good for individual differences (Ravenzwaaij & Oberauer, 2009). No complete implementation exists for the full model in R, although a package has been developed for distribution functions of the Wiener diffusion process (RWiener; Wabersich, 2014; Wabersich & Vanderckhove, 2014). The accompanying script uses the RWiener package to model a simplified version of the Wiener diffusion model for FPT in R (refer Appendix 2 for details of script, assumptions, and steps).

R software implementation

The accompanying script uses a package designed by Wabersich and Vanderckhove (2014), which provides options for calculating diffusion functions using R software. The script uses the package along with other R functions to write a complete script that models the diffusion model for simulated data for a two choice RT paradigm. The functions calculate the pdf and cdf for the WFPT, with distinctions for small and large time representations (Navarro & Fuss, 2009; Blurton et al, 2012).

Note that as with any statistical analysis, the best technique will vary based on the situation under investigation (Voss et al., 2013). For example, although the mean is the ideal technique in many cases of estimating the central tendency of a distribution, there are some situations where other measures such as the median or mode are more suited. Similarly, other methods may be more robust, efficient, or accurate in model fitting. This especially applies to

simulating the diffusion process (Tuerlinckx et al., 2001) and choosing an optimization method (Ratcliff & Tuerlinckx, 2002). Finally, there are many ways of further experimenting with the script to understand the behaviour of the model and test its application under various manipulations.

Task paradigm

The script uses the example of a simple two choice RT paradigm, a color discrimination task, with two conditions, in which speed and accuracy are separately emphasized. It fits the diffusion model for each condition separately, to test if boundary separation (a) varies while assuming that other parameters are constant. The task has been used before with the diffusion model (Voss, Rothermund, & Voss, 2004; Spaniol et al, 2011), and similar perceptual discrimination tasks have also been successfully used many times (Ratcliff and Rouder, 1998; Ratcliff & McKoon, 2008). For new untested paradigms, model assumptions and validity for the paradigm would have to be tested.

Diffusion simulation

Data were simulated using the RWiener's function (rwiener) for obtaining random samples of data with an underlying diffusion process given a set of parameters. The parameter values for the simulations were chosen based on those obtained in empirical data (Ratcliff & Tuerlinckx, 2002; Ratcliff & McKoon, 2008). The algorithm for calculating the diffusion process was based on that the rejection-based method outlined by Tuerlinckx and colleagues (2001). Note that other diffusion algorithms are also outlined in the same paper. Although the rejection-based method is ideal for most situations, other methods may be better suited based on the experimental situation (such as the computational speed, accuracy of simulation, etc.).

The selected number of trials, 200, is reflective of what would be feasible in an experimental situation. It also avoids too many trials since that would produce practice effects on the task (Dutilh, Krypton, & Wagenmakers, 2011; Dutilh, Vanderckhove, & Tuerlinckx, 2009). For cases where the diffusion model is being fit to a task paradigm it has been validated with fewer trials are needed, compared with cases where new paradigms are used.

Model fitting

The optimization procedure chosen is maximum likelihood estimation since it is suitable for small datasets, and outlier removal is first conducted to avoid bias from optimization. Likelihood functions are provided by the RWiener package to calculate the log likelihood and deviance for a set of model parameters. To avoid local minima, the chosen optimization routine follows Vanderckhove (2008) recommendation of using an iterative optimization procedure, in which the final step uses a quasi-Newton method and approximates the Hessian matrix.

As with other model applications (Vanderckhove & Tuerlinckx, 2007; Vanderckhove 2008; Voss & Voss, 2008), the script uses the EZ-diffusion model to obtain starting values before estimating the model parameters. The EZ-diffusion model provides values for drift (v), boundary separation (a), and non-decision time (Ter). It assumes that the starting point (z) is unbiased, hence the value is calculated using the EZ estimate for alpha (z=a/2).

Model selection

Since the simulation assumed that the actual model parameters were unknown, as would be the case in a real experiment, model selection was done by comparing the fit of two competing models for the data, a null model in which the boundary separation parameter does not vary across conditions, and an alternate model in which it does. The package likelihood

functions were used to obtain a test of the likelihood for each model, based on the chi-square goodness of fit (which are both appropriate for testing the diffusion model, Voss et al, 2013).

Since model selection techniques are limited in their suitability to the diffusion model, a graphical presentation of fitted data is also given, plus a Monte-Carlo simulation for the likelihood ratio test in which multiple datasets are generated based on the fitted parameters, and the likelihood ratio assessed for each situation (Voss et al., 2013). To avoid over-fitting (Pitt & Myung, 2002), the Akaka information criterion was used (AIC, Akaike, 1973), which was calculate using the following formula from Myung (2000)

$$AIC_k = 2\ \log(ML_k)+2p_k \qquad [4]$$

Outputs

The script presents tables of performance measures for each condition plus the obtained parameters after fitting the diffusion model (Tables 2 and 3). It also outputs graphs using R's curve function to calculate cdfs of responses in each condition (Figure 5) and RWiener's inbuilt function for plotting RT distributions for correct and error responses (Figure 6).

Summary

This paper outlines how a version of the Wiener diffusion model offers a useful way to examine the speed-accuracy tradeoff under a cognitive framework. The model offers many statistical advantages to the current analysis of RTs and accuracy, which completely ignores the speed-accuracy tradeoff, and the analysis of tradeoffs alone, which ignores the underlying theoretical processes for both the tradeoff and other contributing influences. I demonstrate how the EZ-diffusion model can be applied to data simulated from a choice RT experiment. The success of the model's predictions in matching data on choice RTs and speed-accuracy tradeoffs warrant its consideration for analyses of decision time.

References

Blurton, S. P., Kesselmeier, M., & Gondan, M. (2012). Fast and accurate calculations for cumulative first-passage time distributions in Wiener diffusion models. *Journal of Mathematical Psychology*, *56*(6), 470–475.

Bogacz, R. (2013). Speed-Accuracy Trade-Off. In *Encyclopedia of Computational Neuroscience* (pp. 1–4). Springer. Retrieved from http://link.springer.com/content/pdf/10.1007/978-1-4614-7320-6_319-1.pdf

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision-making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700.

Brown, S. D., Ratcliff, R., & Smith, P. L. (2006). Evaluating methods for approximating stochastic differential equations. *Journal of Mathematical Psychology*, *50*(4), 402–410.

Darling, D. A., & Siegert, A. J. F. (1953). The first passage problem for a continuous Markov process. *The Annals of Mathematical Statistics*, 624–639.

Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E.-J. (2011). Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review*, *18*(1), 61–69.

Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E.-J. (2012). Testing theories of post-error slowing. *Attention, Perception, & Psychophysics*, *74*(2), 454–465.

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review*, *16*(6), 1026–1036.

Einstein, A. (1905). On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heart. *Annalen Der Physik*, *17*, 549–560.

Feller, W. (1968). *An Introduction to Probability Theory, Vol. 1*. Wiley, New York.

Gondan, M., Blurton, S. P., & Kesselmeier, M. (2014). Even faster and even more accurate first-passage time densities and distributions for the Wiener diffusion model. *Journal of Mathematical Psychology*, *60*, 20–22.

Grasman, R. P., Wagenmakers, E.-J., & van der Maas, H. L. (n.d.). EZ2: An extension of the EZ-diffusion model. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.80.5838&rep=rep1&type=pdf

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, *109*(2), 340.

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4052662/

Heitz, R. P., & Schall, J. D. (2013). Neural chronometry and coherency across speed–accuracy demands reveal lack of homomorphism between computational and neural mechanisms of evidence accumulation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1628), 20130071.

Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 598.

Horn, S. S., Bayen, U. J., & Smith, R. E. (2011). What can the diffusion model tell us about prospective memory? *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *65*(1), 69.

Jennings, J. R., Wood, C. C., & Lawrence, B. E. (1976). Effects of graded doses of alcohol on speed-accuracy tradeoff in choice reaction time. *Perception & Psychophysics*, *19*(1), 85–91.

Luce, R. D. (1986). *Response times*. Oxford University Press. Retrieved from https://books.google.ca/books?hl=en&lr=&id=WSmpNN5WCw0C&oi=fnd&pg=PA1&dq=Luce,+1986+response+times&ots=XqHSPaa2iP&sig=jV2FVkBqAGK__u-KDgYAiNmOzfM

Ma, Z. S., Krings, A. W., & Millar, R. C. (2009). Introduction of first passage time (FPT) analysis for software reliability and network security. In *Proceedings of the 5th Annual Workshop on Cyber Security and Information Intelligence Research: Cyber Security and Information Intelligence Challenges and Strategies* (p. 63). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1558680

Myung, I. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology*, *44*(1), 190–204.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*(1), 90–100.

Navarro, D. J., & Fuss, I. G. (2009). Fast and accurate calculations for first-passage times in Wiener diffusion models. *Journal of Mathematical Psychology*, *53*(4), 222–230.

Pachella, R. G., & Pew, R. W. (1968). Speed-accuracy tradeoff in reaction time: Effect of discrete criterion times. *Journal of Experimental Psychology*, *76*(1), 19–24.

Pew, R. W. (1969). The speed-accuracy operating characteristic. Retrieved from http://deepblue.lib.umich.edu/handle/2027.42/33038

Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, *6*(10), 421–425.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*(3), 446.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review*, *9*(2), 278–291.

Ratcliff, R. (2008). The EZ diffusion method: Too EZ? *Psychonomic Bulletin & Review*, *15*(6), 1218–1228.

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

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333.

Ratcliff, R., Thapar, A., & Mckoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & Psychophysics*, *65*(4), 523–535.

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging, practice, and perceptual tasks: a diffusion model analysis. *Psychology and Aging*, *21*(2), 353.

Ratcliff, R., Thapar, A., & McKoon, G. (2007). Application of the diffusion model to two-choice tasks for adults 75-90 years old. *Psychology and Aging*, *22*(1), 56.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481.

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, *106*(2), 261.

Rouder, J. N., Lu, J., Speckman, P., Sun, D., & Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin & Review*, *12*(2), 195–223.

Rouder, J. N., Sun, D., Speckman, P. L., Lu, J., & Zhou, D. (2003). A hierarchical Bayesian statistical framework for response time distributions. *Psychometrika*, *68*(4), 589–606.

Shepp, L. A. (1967). A first passage problem for the Wiener process. *The Annals of Mathematical Statistics*, 1912–1914.

Siegert, A. J. (1951). On the first passage time probability problem. *Physical Review*, *81*(4), 617.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161–168.

Spaniol, J., Voss, A., Bowen, H. J., & Grady, C. L. (2011). Motivational incentives modulate age differences in visual perception. *Psychology and Aging*, *26*(4), 932.

Standage, D., Blohm, G., & Dorris, M. C. (2014). On the neural implementation of the speed-accuracy trade-off. *Frontiers in Neuroscience*, *8*. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4131279/

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

Tuerlinckx, F., Maris, E., Ratcliff, R., & De Boeck, P. (2001). A comparison of four methods for simulating the diffusion process. *Behavior Research Methods, Instruments, & Computers*, *33*(4), 443–456.

Vandekerckhove, J. (2009). Extensions and applications of the diffusion model for two-choice response times. *Status: Published*. Retrieved from https://lirias.kuleuven.be/handle/1979/2658

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*, *14*(6), 1011–1026.

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, *40*(1), 61–72.

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. (2008). A Bayesian approach to diffusion process models of decision-making. In *Proceedings of the 30th annual conference of the cognitive science society* (pp. 1429–1434). Retrieved from https://lirias.kuleuven.be/handle/123456789/198884

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods*, *16*(1), 44.

Van Ravenzwaaij, D., & Oberauer, K. (2009). How to use the diffusion model: Parameter recovery of three methods: EZ, fast-dm, and DMAT. *Journal of Mathematical Psychology*, *53*(6), 463–473.

Vickers, D., Burt, J., Smith, P., & Brown, M. (1985). Experimental paradigms emphasizing state or process limitations: I Effects on speed-accuracy tradeoffs. *Acta Psychologica*, *59*(2), 129–161.

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: A practical introduction. *Experimental Psychology*, *60*(6), 385.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*(7), 1206–1220.

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767–775.

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology*, *52*(1), 1–9.

Wabersich, D., & Vandekerckhove, J. (2014). The RWiener Package: an R Package Providing Distribution Functions for the Wiener Diffusion Model. *A Peer-Reviewed, Open-Access Publication of the R Foundation for Statistical Computing*, 49.

Wabersich, D., Wabersich, M. D., & Suggests, M. (2014). Package "RWiener." Retrieved from ftp://mirror.csclub.uwaterloo.ca/CRAN/web/packages/RWiener/RWiener.pdf

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, *21*(5), 641–671.

Wagenmakers, E.-J., van der Maas, H. L., Dolan, C. V., & Grasman, R. P. (2008). EZ does it! Extensions of the EZ-diffusion model. *Psychonomic Bulletin & Review*, *15*(6), 1229–1235.

Wagenmakers, E.-J., Van Der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22.

Wakuta, K. (2000). A first-passage problem with multiple costs. *Mathematical Methods of Operations Research*, *51*(3), 419–432.

Whelan, R. (2010). Effective analysis of reaction time data. *The Psychological Record*, *58*(3), 9.

White, C. N., Ratcliff, R., & Starns, J. J. (2011). Diffusion models of the flanker task: discrete versus gradual attentional selection. *Cognitive Psychology*, *63*(4), 210–238.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*(1), 67–85.

Wood, C. C., & Jennings, J. R. (1976). Speed-accuracy tradeoff functions in choice reaction time: Experimental designs and computational procedures. *Perception & Psychophysics*, *19*(1), 92–102.

Zandt, T. V. (2002). Analysis of response time distributions. *Stevens' Handbook of Experimental Psychology*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/0471214426.pas0412/full

Tables and figures

Table 1.

Parameters for the Wiener diffusion model

Symbol	Parameter	Explanation
v	Drift rate	Information accumulation
a	Boundary separation	Speed-accuracy tradeoff
z	Starting point	Bias towards a response
T_{er}	Non-decision time	Encoding time and response output

Table 2.

Results from the accompanying R script showing performance for each condition

	Pc	VRT	MRT
Speed performance	0.63	0.035	0.53
Accuracy performance	0.75	0.47	1.22

Table 3.

Results from the accompanying R script showing fitted parameters for each condition

Boundary separation	Non-decision time	Starting point	Drift	Condition
0.95	0.31	0.49	0.58	speed
2.04	0.30	0.50	0.56	accuracy

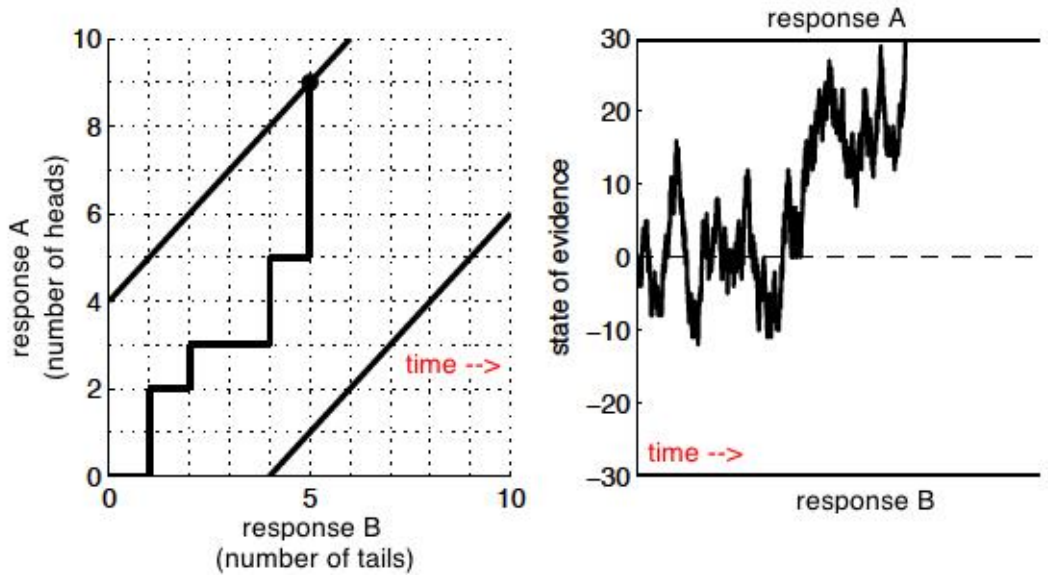


Figure 1. Image with simple two-decision sequential sampling models. A discrete-time random walk model is shown on the left, and a continuous-time diffusion model on the right. On the left image, the axis for time is rotated 45 degrees and runs diagonally. Figure based on Navarro and Fuss (2009).

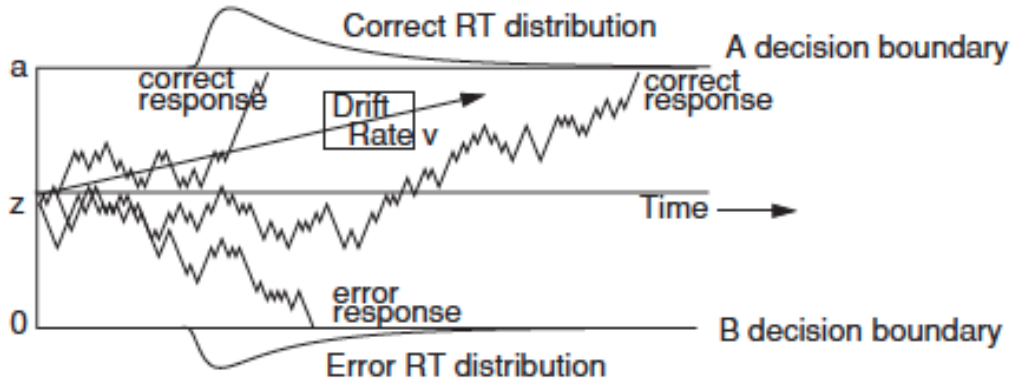


Figure 2. Ratcliff's diffusion decision model. Three sample information accumulation paths (mean slope *v*) are demonstrated. Each path begins at starting point z, and continues until it ends at the upper (*a*) or lower threshold (0). Figure reprinted from Ractliff and McKoon (2008).

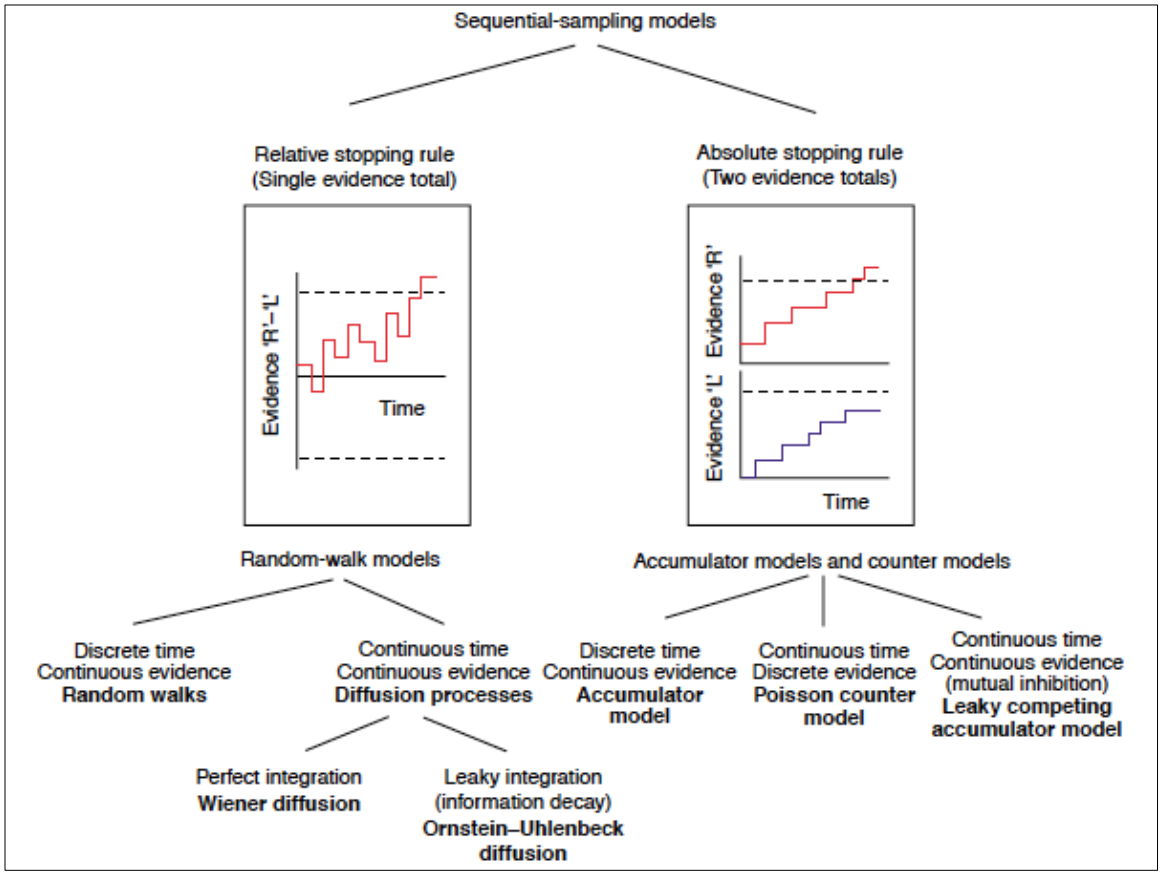
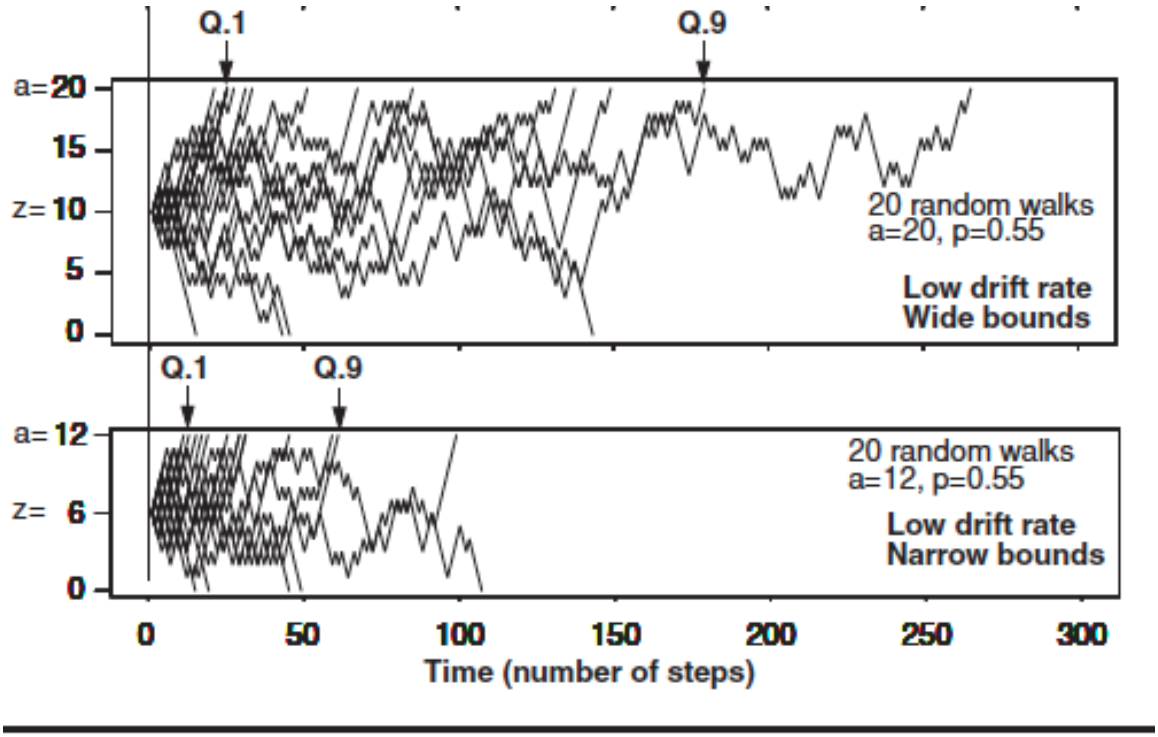
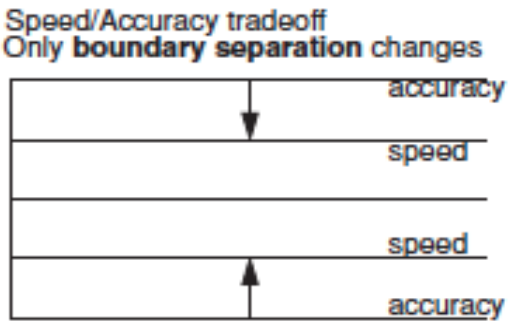


Figure 3. A taxonomy of sequential sampling models (Smith & Ratcliff, 2004). The chosen model will vary depending on the theoretical process under investigation.





F*igure 4.* Demonstration of how boundary separation depicts the speed-accuracy tradeoff. The image above shows a high (*a*=20) and low (*a*=12) boundary separation. The image below shows how increasing the boundary separation causes increases in accuracy, but at the expense of decreasing the speed of response. Figure reprinted from Ratcliff and McKoon (2008).

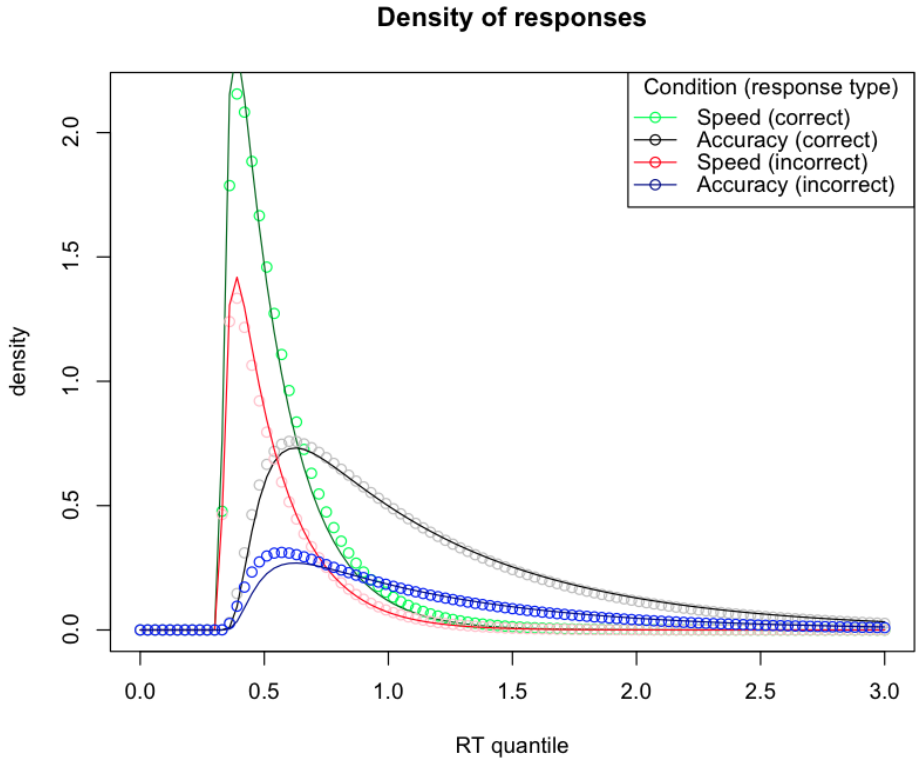


Figure 5. Image with cdfs of RTs for correct and incorrect responses in each experimental condition. The points are the obtained data, and the lines are model predictions.

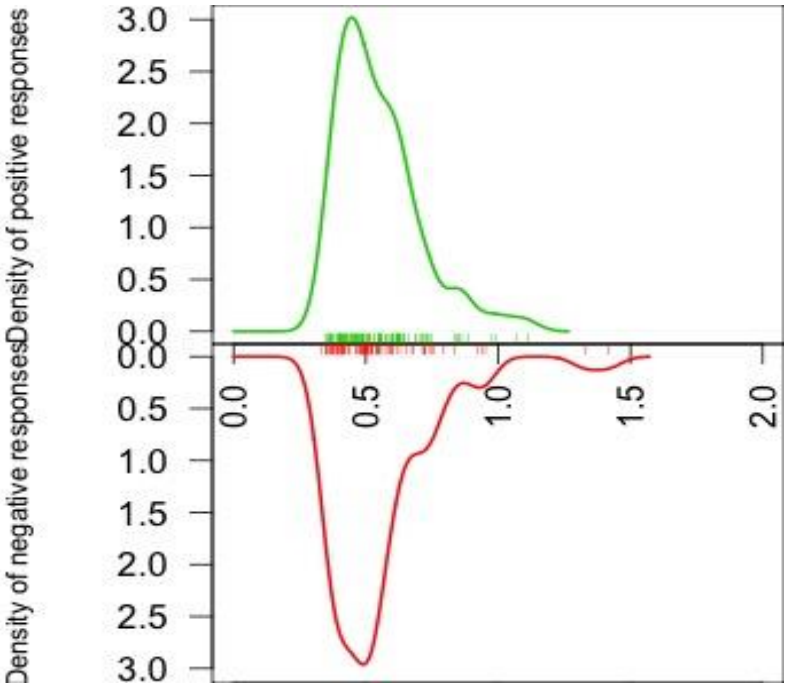


Figure 6. Plot using the RWiener package function with RTs for one experimental condition. Correct responses are shown above in green and incorrect responses below in red.

Appendix

Code used to assist with R script

EZ-diffusion model code from Wagenmakers, van der Maas, and Grasman (2007)

```
get.vaTer <- function(Pc, VRT, MRT, s=0.1)
  {
  s2 <- s^2
  # The default value for the scaling parameter s equals 0.1
  if (Pc == 0)
    cat("Oops, Pc == 0!\n")
  if (Pc == 1)
    cat("Oops, Pc == 1!\n")
  # If Pc equals 0, .5, or 1, the method will not work, and
  # an edge correction is required.
  L <- qlogis(Pc)
  # The function "qlogis" calculates the logit.
  x <- L*(L*Pc^2 - L*Pc + Pc - .5)/VRT
  v <- sign(Pc-.5)*s*x^(1/4)
  # This gives drift rate.
  a <- s2*qlogis(Pc)/v
  # This gives boundary separation.
  y <- -v*a/s2
  MDT <- (a/(2*v)) * (1-exp(y))/(1+exp(y))
  Ter <- MRT - MDT
  # This gives nondecision time.
  return(list(v, a, Ter))
}
```

Code from the RWiener package help page (Wabersich, 2014) used for guidance in the script

```
# estimate parameters with optim
# onm <- optim(c(1,.1,.1,1),wiener_deviance,dat=dat, method="Nelder-Mead")
# est <- optim(onm$par,wiener_deviance,dat=dat, method="BFGS",hessian=TRUE)

# compare two models with deviance wiener_deviance(c(3,.3,.5,0), dat)
wiener_deviance(c(3,.3,.5,0.5), dat)
# log-likelihood difference wiener_likelihood(c(3,.3,.5,0), dat)-
wiener_likelihood(c(3,.3,.5,0.5), dat)

Likelihood-Ratio-Test and Wald-Test
# wiener_likelihood(est1$par,dat1)+wiener_likelihood(est1$par,dat2)
# combined loglike
# model_ll <- function(pars,delta,dat1,dat2) {
#   wiener_likelihood(pars,dat1)+
#     wiener_likelihood(c(pars[1:3],pars[4]+delta),dat2)
# }
# ## Likelihood ratio test
# model_ll(est1$par,1,dat1,dat2)
# # compute Likelihood ratio
# LR <- 2*model_ll(est1$par,0,dat1,dat2) + 2*model_ll(est1$par,1,dat1,dat2)
# # compare with critical X^2(1) quantile, alpha=0.05
# LR > qchisq(0.95,1)
```

```
# # get p-value from X^2(1)
# pchisq(LR,1, lower.tail=FALSE)
```