Using diffusion models to understand clinical disorders☆

Corey N. White *, Roger Ratcliff, Michael W. Vasey, Gail McKoon

The Ohio State University, United States

A R T I C L E   I N F O

A B S T R A C T

Sequential sampling models provide an alternative to traditional analyses of reaction times and accuracy in two-choice tasks. These models are reviewed with particular focus on the diffusion model (Ratcliff, 1978) and how its application can aid research on clinical disorders. The advantages of a diffusion model analysis over traditional comparisons are shown through simulations and a simple lexical decision experiment. Application of the diffusion model to a clinically relevant topic is demonstrated through an analysis of data from nonclinical participants with high- and low-trait anxiety in a recognition memory task. The model showed that after committing an error, participants with high-trait anxiety responded more cautiously by increasing their boundary separation, whereas participants with low-trait anxiety did not. The article concludes with suggestions for ways to improve and broaden the application of these models to studies of clinical disorders.

Techniques and models from cognitive psychology are being used with increasing frequency in investigations of psychopathology and clinical disorders (e.g., McFall, Treat, & Viken, 1997; McNally & Reese, 2009; Treat & Dirks, 2007). These methods and models play a significant role in elucidating the abnormal cognitive processes that are associated with such disorders. In this article we demonstrate how a theory of cognitive processing can enhance cognitive–clinical interactions and lead to a better understanding of the cognitive effects of psychopathologies like depression and anxiety. The focus is on the use of sequential sampling models to analyze data from two-choice response time (RT) tasks. The article is structured as follows: We briefly review some areas in which two-choice tasks have been employed to investigate cognitive processing in anxiety and depression. We then show through simulations and a simple lexical decision experiment how a sequential sampling model, Ratcliff's diffusion model (Ratcliff, 1978; Ratcliff, Van Zandt, & McKoon, 1999), can improve analyses of two-choice tasks by decomposing accuracy and RT distributions into distinct components of processing. Application of the diffusion model to a clinically relevant topic is then demonstrated through the analysis of recognition memory data from subclinical participants with high- and low-trait anxiety to assess changes in decision criteria that result from committing an error. The article concludes with discussion of areas of potential improvement for the application of sampling models like the diffusion model to inform clinical research.

1. Two-choice tasks and clinical research

The focus of this article is two-choice tasks to which sequential sampling models can be applied. These tasks involve a fast (typically less than two seconds), one-process decision and the collection of RTs and accuracy. This includes, but is not limited to, discrimination (e.g., brightness or numerosity; Ratcliff & Rouder, 1998), recognition memory (Ratcliff, 1978; Ratcliff, Thapar, & McKoon, 2004; Spaniol, Voss, & Grady, 2008), lexical decision (Ratcliff, Gomez, & McKoon, 2004), stop-signal (Verbruggen & Logan, 2009), implicit association (Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007) and perceptual matching (Ratcliff, 1981; Van Zandt, Colonius, & Proctor, 2000). These tasks cover a range of decision types to which sampling models can be applied, including yes/no, old/new, same/different, categorization, two-alternative forced choice, and response signal.

Within the realm of research on psychopathology and clinical populations, two-choice tasks are commonly employed to investigate processing differences between patients and healthy controls. For example, these tasks have been instrumental in showing that individuals with high levels of anxiety show preferential attention for threatening information (see Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007, for a

meta-analytic review). In a modified probe detection task, a threatening word (e.g., cancer) and a neutral word (e.g., chair) are shown at different locations on a screen, and one of the words is replaced by a probe that participants must detect. Anxious individuals show faster RTs when the probe replaces the threatening word compared to the neutral word, suggesting that they preferentially attend to threat (e.g., Mogg & Bradley, 1999; Mogg, Bradley, De Bono, & Painter, 1997). Similar results have been demonstrated with obsessive compulsive disorder (Lavy, van Oppen, & van den Hout, 1994), posttraumatic stress disorder (McNally, Kaspi, Riemann, & Zeitlin, 1998), social anxiety disorder (Rheingold, Herbert, & Franklin, 2003), and panic disorder (McNally, Riemann, & Kim, 1990), suggesting that preferential processing of threat is common to many anxiety disorders (see Bar-Haim et al., 2007). It is thought that this bias for threat is involved in both the etiology and maintenance of anxious states (Mathews, 1990; Mathews & Mackintosh, 1998), making it an important component of anxiety. Research in this domain has led to the development of several models that account for the association between threat bias and anxiety (Bishop, 2007; Frewen, Dozois, Joanisse, & Neufeld, 2008; Mathews & Mackintosh, 1998; Mogg & Bradley, 1998; Weierich, Treat, & Hollingworth, 2008).

Researchers using two-choice tasks to study depression have found a slightly different pattern of processing differences. Whereas high anxiety is associated with a bias to process threatening information, depressive symptoms are more closely linked with abnormal emotional processing. Nondepressed individuals typically show a bias for positive over negative emotional information, but depressed individuals either lack that advantage, leading to unbiased processing of positive and negative emotional information (Siegle, Granholm, Ingram, & Matt, 2001), or show a bias for negative over positive emotional information (Power, Cameron, & Dalgleish, 1996). Further, unlike high anxiety, depressive symptoms are associated with deficits on many cognitive tasks. Depressed patients have shown slower RTs and lower accuracy on two-choice recognition memory tasks (Hilbert, Niederehe, & Kahn, 1976) and greater interference on Stroop tasks (Lemelin et al., 1996).

Two-choice tasks have also been employed to assess the efficacy of antidepressants. In one study, patients taking the antidepressant drug reboxetine were better at a simple two-choice identification task than patients taking a placebo (Ferguson, Wesnes, & Schwartz, 2003; see also Hindmarch, 1998), suggesting that the antidepressant mitigated the processing deficit. A recent study also showed that for healthy volunteers, reboxetine increased processing of positive emotional words in a manner that could potentially reverse negative biases in depression (Norbury, MacKay, Cowen, Goodwin, & Harmer, 2008).

Other psychological disorders have been studied with two-choice tasks, including obsessive compulsive disorder (Ruchsow et al., 2005), schizophrenia (Williams & Hemsley, 1986), hypochondriasis (Lecci & Cohen, 2007), borderline personality disorder (Nigg, Silk, Starvo, & Miller, 2005), and posttraumatic stress disorder (Masten et al., 2008). A complete review of the different methodologies and findings from studies such as these is beyond the scope of this article, but this sample illustrates the very active domain of clinical research with two-choice tasks.

For each of the studies mentioned above, data analyses involved comparisons of average RTs and/or accuracy values. Indeed, this approach has served researchers well in studies of clinical disorders and cognitive processing, as evidenced by the brief review above. However, there are situations in which comparisons of RTs or accuracy cannot sufficiently identify processing differences between groups or conditions. This problem can be overcome by using sequential sampling models to augment analyses of two-choice tasks. Section 2 provides an overview of these models and the ways in which they can improve analyses of two-choice tasks.
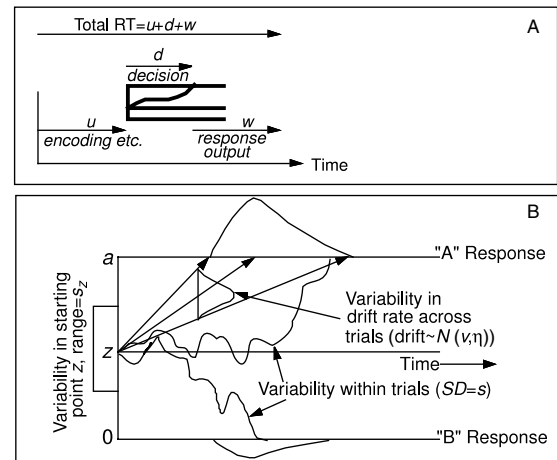

Fig. 1. An illustration of the diffusion model. Panel A shows the total response process, including encoding and response output. Panel B shows the diffusion process for the decision component of the response process. Parameters of the model are: a, boundary separation; z, starting point; T_{er}, mean value of the nondecision component of reaction time; η, SD in drift across trials; s_z, range of the distribution of starting point (z) across trials; v, drift rate; p_0, proportion of contaminants; and s, SD in variability in drift within trials..

2. Sequential sampling models

Sequential sampling models describe the processes involved in making fast, two-choice decisions. The models were developed to account for the entire data set associated with two-choice paradigms, namely accuracy and the distributions of RTs for correct and error responses. Several models have been developed in this class, including the Linear Ballistic Accumulator model (Brown & Heathcote, 2008), the Leaky Competing Accumulator model (Usher & McClelland, 2001), the Poisson Counter model (Smith & Van Zandt, 2000), and the Diffusion model (Ratcliff, 1978; Ratcliff et al., 1999). We focus on Ratcliff's diffusion model because it has been widely employed as an analytical tool (see Ratcliff & McKoon, 2008), and has been shown to fit behavioral data as well as or better than competing models (Ratcliff & Smith, 2004). However, the primary structure and components of the diffusion model are found in various degrees in all of models in this class.

The diffusion model is a theory of simple, two-choice decisions. The model assumes that noisy evidence is accumulated over time until a criterial amount has been reached, at which point a response is initiated. Fig. 1 shows a schematic of the model. Panel A shows the entire response process. The stimulus is encoded (u), a decision is reached (d), and the response is executed (w). The model does not explain encoding or response execution, but it incorporates a parameter, Ter, to account for the time these processes take ($u + w$ in Fig. 1A). The focus of the model is the diffusion–decision process, shown in Panel B. In the model, noisy evidence is accumulated from a starting point, z, to one of two boundaries, a or 0. The two boundaries represent the two possible decisions, such as yes/no, word/nonword, etc. Once the process reaches a boundary the corresponding response is initiated. The inherent noise in the accumulation of information, represented by the nonmonotonic paths in Fig. 1B, produces the characteristic right skew of empirical RT distributions.

The primary components of the model are boundary separation (a), drift rate (v), starting point (z), and nondecision processing (Ter). Each has a straightforward psychological interpretation. The position of the starting point, z, indexes response bias. If an individual is biased towards a response (e.g., through different payoffs), their starting point will be closer to the corresponding boundary, meaning that less evidence is required to make that response. This will lead to faster and more probable responses

at that boundary compared to the other. The separation between the two boundaries, a, indexes response caution or speed/accuracy settings. A wide boundary separation reflects a cautious response style. In this case, the accumulation process will take longer to reach a boundary, but it is less likely to hit the wrong boundary by mistake, producing slow but accurate responses. Drift rate, v, indexes the quality of evidence from the stimulus. If the stimulus is easily classified, it will have a high rate of drift and approach the correct boundary quickly, leading to fast and accurate responses. The noise in the evidence accumulation, s, acts as a scaling parameter of the model (i.e., if it were doubled, the other parameters could be doubled to produce the same pattern of data), and is set to a fixed value of .1.

There is variability in the values of some of these components based on the assumption that they fluctuate from trial to trial in the course of an experiment. Such variability is necessary for the model to correctly account for the relative speeds of correct and error responses. *Eta* is the across-trial variability in drift rate, s_z is across-trial variability in starting point, and s_t is across-variability in *Ter*. The model also includes an assumption about contaminants (e.g., lapses in attention) and estimates the proportion of contaminant responses, p_o. For the mathematical details of the diffusion model, readers are directed to Ratcliff and Tuerlinckx (2002) or Ratcliff and Smith (2004).

3. Advantages of a diffusion model analysis

There are several advantages of the diffusion model over traditional analyses of RTs and/or accuracy. The first, and perhaps the most important, advantage of the diffusion model stems from its ability to decompose behavioral data into processing components. The model can be fit to behavioral data to separate out the different component values described above, allowing researchers to compare values of response caution, response bias, nondecision time, and stimulus evidence. With this approach researchers can better identify the source(s) of differences between groups of subjects. For example, older adults (60–90 year old) are often slower than college students in two-choice tasks, which has been taken by some to reflect a general decline or slowdown in processing (e.g., Myerson, Ferraro, Hale, & Lima, 1992). However, Ratcliff, Thapar, and McKoon (2006) used the diffusion model to show that, in tasks such as recognition memory and brightness discrimination, older adults are slower because of longer nondecision time and wider boundary separation (i.e., they are more cautious). Importantly, older adults did not have lower drift rates than young adults, suggesting that they still acquire the same quality of information from a stimulus. Thus there was no age-related impairment in discrimination or recognition memory. The diffusion model allowed a more detailed examination of the differences between older and young adults, challenging the general slowing hypothesis and providing an alternative account of the data.

This approach can easily be extended to studies of psychopathology. Suppose we performed an experiment with depressed patients and healthy controls, and found that the patients were slower overall (e.g., Lemelin, Baruch, Vincent, Everett, & Vincent, 1997; Pisljar, Pirtosek, Repovs, & Grgic, 2008; Rogers et al., 2000). With the diffusion model, we could determine if the RT difference was due to more cautious responding (boundary separation), poorer evidence from the stimulus (drift rates), or slower motor response (nondecision time).

A detailed understanding of processing differences associated with pathologies like depression can potentially lead to better assessment and treatment. Targeted cognitive treatment has been employed to reverse biased information processing and decrease levels of anxiety or depression. In one study, individuals characterized by excessive cognitive worry were trained to selectively direct attention *away* from threatening words (Hazen, Vasey, & Schmidt, 2009). Participants performed several sessions of a modified probe discrimination task that included threat words paired with neutral words. Importantly, the probe replaced the neutral word rather than the threat word on almost every trial, so over time participants were implicitly trained to attend away from the threat words. This simple training regiment significantly reduced threat bias and levels of anxiety and depression compared to a sham training condition. In a related vein, Lang, Moulds, and Holmes (2009) had participants watch a depressing film, then trained half of the group to have a more positive appraisal of emotional events and half to have a more negative appraisal. The group that was trained with the positive emotional bias had fewer depressive intrusions and were less impacted by the negative film. Studies such as these show that mitigating or reversing biased processing of information appears to be a promising treatment for depression and anxiety. In this regard, detailed understanding of the relationship between cognitive biases and psychopathology can lead to more effective treatment.

By fitting RTs and accuracy jointly, the diffusion model can aid with the identification of different types of bias. It is well known that differences in accuracy or RTs can be due to discriminability or response bias. For example, in a recognition memory task individuals must determine whether test words were previously studied or new. Suppose that one group of participants had more hits than another (i.e., they correctly recognized more studied words). This could reflect stronger memory for the first group, or instead a bias to respond "old". If participants in the first group also responded "old" to many of the unstudied lures, it would suggest the results were due to bias. Analyses of this sort benefit greatly from the use of signal detection theory (SDT; Green & Swets, 1974), which uses accuracy values from each condition to distinguish between discriminability and bias. However, SDT cannot differentiate between two types of bias that can produce similar patterns of accuracy, response bias and memorial bias. Response bias refers to a shift of the decision criterion, where individuals require more or less evidence to make one of the responses. This corresponds to the starting point, z, in the diffusion model. Memorial bias refers to a shift in the memory-strength distribution underlying the decision, where a class of stimuli provide more or less evidence for the response. This corresponds to the drift criterion in the diffusion model (a direct analog of the criterion in SDT, see Ratcliff & McKoon, 2008), which essentially reflects a shift in drift rates for each condition (e.g., the drift rate for old items increases by the same amount that the drift rate for new items decreases). Response biases and memorial biases produce similar changes in accuracy, thus SDT cannot differentiate between them. But by including RTs into the analysis, the diffusion model can separate the effects of these biases and identify which is responsible for the data (Spaniol et al., 2008; Voss, Rothermund, & Brandtstadter, 2008).

This approach can be applied to studies of psychopathology. Several studies have used affective decision tasks (e.g., "Is this word threatening or not?", "Is this picture emotionally positive or negative?") to explore biases for threatening or emotional information in individuals with anxiety or depression. In these tasks, participants with high anxiety are more likely to classify words as threatening compared to low anxiety participants. Several studies have analyzed this effect and concluded that it is simply a response bias and does not reflect any differences in the threat value of the words or pictures themselves (Becker & Rinck, 2004; Manguno-Mire, Constans, & Geer, 2005; Windmann & Kruger, 1998). A diffusion model analysis of such data could augment our understanding of this bias by determining whether it is due to response bias, perceptual bias, or both. Results of such a study could inform clinicians who wish to employ cognitive training treatments similar to the ones previously discussed. If the tendency to classify items or events as

threatening stemmed from a response bias, clinicians might focus on training patients to overcome this bias by actively classifying items as nonthreatening. If, on the other hand, the bias stemmed from a perceptual bias, clinicians might train patients to associate items with safety, thus reducing the threat value of the items. Further, it is conceivable that these two types of bias are indicative of different subtypes of anxiety, meaning that studies of this sort could improve assessment, classification, and treatment.

The diffusion model can identify different decision components because it utilizes all of the behavioral data. The model uses accuracy and RT distributions for correct and error responses, so all of the data are explained. In contrast, a comparison of mean RTs does not account for potential differences in accuracy, and an SDT analysis does not account for difference in RTs. Since each aspect of the data is affected by the different components of the decision process, each one contains useful information that should be used to inform analyses.

The final advantage of the diffusion model that we explore in this article involves the quality of evidence extracted from the stimuli. Stimulus evidence refers to how strongly an item indicates a response. In lexical decision, for example, participants must classify letter strings as words or nonwords. Commonly encountered words, like tree, produce a strong lexical match and thus strong evidence for the *word* response, whereas rare words, like aardvark, produce a weak lexical match and thus weak evidence for the *word* response. This difference in lexical evidence would be reflected by faster RTs, higher accuracy, and higher drift rates for the common words.

Stimulus evidence is often the primary focus of researchers employing two-choice tasks, and drift rates provide a more direct index of it than either RTs or accuracy. The reasoning is as follows: RTs and accuracy are used to index stimulus evidence, but they are both affected by the other components of the decision process. Individual differences in decision components like boundary separation essentially add noise to RTs and accuracy because they do not reflect differences in evidence strength. This decreases the sensitivity of these measures when used to assess stimulus evidence. In contrast, when the diffusion model is applied the drift rate parameter indexes stimulus evidence, and the effects of the other components are separated into the corresponding parameters. In other words, individual differences in components like boundary separation affect RTs and accuracy, but not drift rates. As a result, drift rates are better able to detect small differences in stimulus evidence that might not be apparent with comparisons of RTs or accuracy. This point is illustrated in Section 4.

4. Measures of stimulus evidence

We present the results from a simple lexical decision experiment and a series of simulations that were designed to assess the sensitivity of the dependent measures that can be used to assess differences in stimulus evidence. Although lexical decision is used, the results are meant to apply to two-choice tasks in general.

4.1. Experiment 1

Experiment 1 was a simple lexical decision task. Of primary interest is how well each dependent measure of stimulus evidence, drift rates, RTs, and accuracy, truly reflects the lexical advantage for high frequency words over low frequency words. In the experiment, different components of the decision process were manipulated through instructions to demonstrate how the dependent measures are affected. Separate groups of participants were given either speed/accuracy instructions or response bias (e.g., more words than nonwords) instructions. With this design, we can assess how robust each measure is against the effects of instruction, which should not affect stimulus evidence, and how well each measure captures the lexical advantage of high frequency words.

The effects of speed/accuracy and bias instructions have been explored previously with the diffusion model (e.g., Ratcliff, 1985; Ratcliff & Rouder, 1998; Voss, Rothermund, & Voss, 2004; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008), but to our knowledge no one has explored the effects of these instructions on within-subject comparisons of stimulus evidence. The results of this experiment are not directly related to research on psychopathology but are meant to illustrate a generic situation in which researchers are interested in a processing difference between two types of stimuli.

Procedure. A basic lexical decision task was performed in which participants were shown single strings of letters and asked to determine whether they were words or not. The stimuli were displayed on a CRT of a Pentium IV class PC, and RTs and accuracy were collected from the keyboard. Participants were told to press the "/" key if the string was a word and the "z" key if it was a nonword. They were originally instructed to respond quickly and accurately. Letter strings were presented until a response was made, with a 200 ms interval before presentation of the next string. To discourage guessing, the word "ERROR" was displayed for 750 ms after an incorrect response. Participants first completed a practice block of 30 words and 30 nonwords.

Participants in the speed/accuracy condition were told to consider the pace at which they performed the practice block to be their normal pace. For each subsequent block they were instructed to go at their normal pace, or to emphasize speed or emphasize accuracy. They were also informed that emphasizing speed might lead to more errors, and emphasizing accuracy might lead to slower responses. At the beginning of each block, participants were informed whether it was a speed, accuracy, or normal pace block. Participants completed six blocks of each type in random order. Each block consisted of 42 nonwords and 42 words, the latter of which were split into half high frequency words and half low frequency words.

For participants in the bias condition, blocks of trials were constructed that contained different proportions of words and nonwords. After completing the practice block, subjects were informed that some blocks would contain more words than nonwords, some more nonwords than words, and some an even number of words and nonwords. At the beginning of each block, subjects were informed that it would be a "word," "nonword," or "even" block. Word blocks contained 60 words (30 each of high and low frequency) and 24 nonwords. Nonword blocks contained 60 nonwords, and 24 words (12 each of high and low frequency). Even blocks contained 42 words (21 each of high and low frequency) and 42 nonwords. Subjects completed 6 of each block in random order, for a total of 18 blocks.

Participants. Ohio State University students completed the experiment for credit in an introductory psychology course. There were 18 participants that received speed/accuracy instructions and 18 participants that received bias instructions.

Stimuli. The stimuli were high and low frequency words and nonwords. The high frequency word pool consisted of 866 words with frequencies from 78 to 10,600 per million (mean 287.49, $SD = 476$; Kucera & Francis, 1967). The low frequency word pool consisted of 899 low frequency words with frequencies of 4 and 5 per million (mean 4.41, $SD = 0.17$). Nonwords were created from a separate word pool by randomly replacing all of the vowels with other vowels (except for u after q), producing pronounceable nonwords. Words and nonwords were randomly chosen from each pool without replacement for each participant.

Model Fitting. All responses faster than 300 ms or slower than 3 s were discarded (less than .6% of the total data). The model was fit to each participant's data using a X^2 minimization routine (for other methods, see Ratcliff & Tuerlinckx, 2002; Voss et al., 2004; Voss & Voss, 2007). The data entered into the routine were the number

Table 1
Best fitting parameter values from Experiment 1.

	a	z/a	Ter	η	s_z	s_t	p_o	v_{HF}	v_{LF}	$v_{nonword}$	$X^2(23)$
Instruction											
Speed	0.104	0.506	0.415	0.049	0.072	0.175	0.002	0.384	0.152	−0.224	45.1 (16)
	(.02)	(.05)	(.04)	(.06)	(.02)	(.05)	(.01)	(.10)	(.06)	(.06)	
Normal	0.133	0.482	0.453	0.135	0.047	0.178	0.001	0.410	0.193	−0.245	38.8 (12)
	(.02)	(.05)	(.03)	(.08)	(.04)	(.05)	(.01)	(.12)	(.06)	(.08)	
Accuracy	0.156	0.495	0.485	0.09	0.061	0.179	0.008	0.387	0.189	−0.253	39.6 (19)
	(.04)	(.04)	(.07)	(.06)	(.05)	(.08)	(.01)	(.12)	(.07)	(.06)	
Biased-word	0.133	0.614	0.417	0.057	0.041	0.173	0.001	0.392	0.152	−0.231	41.7 (14)
	(.03)	(.05)	(.04)	(.06)	(.04)	(.05)	(.01)	(.13)	(.05)	(.06)	
Neutral	0.141	0.518	0.434	0.101	0.060	0.174	0.002	0.457	0.179	−0.238	43.1 (21)
	(.04)	(.05)	(.03)	(.08)	(.05)	(.07)	(.01)	(.17)	(.09)	(.08)	
Biased-nonword	0.131	0.341	0.439	0.067	0.063	0.169	0.001	0.477	0.216	−0.221	40.6 (17)
	(.04)	(.07)	(.03)	(.06)	(.03)	(.06)	(.01)	(.15)	(.09)	(.08)	

Note. Parameter values are averages across participants (SDs in parenthesis). a = boundary separation; z/a = position of starting point relative to the boundaries (values above .5 indicate bias to respond "word" and values below .5 indicate bias to respond "nonword"); Ter = nondecision component; η = variability in drift across trials; s_z = variability in starting point; s_t = variability in nondecision component; p_o = probability of outlier responses, v_{HF} = drift rate for high frequency words; v_{LF} = drift rate for low frequency words; $v_{nonword}$ = drift rate for nonwords.

Table 2
Behavioral data averaged across participants for Experiment 1.

Speed/Acc instruction	Mean RTs			Bias Instruction	Mean RTs		
	Correct	Error	Acc		Correct	Error	Acc
Speed block				**Word-bias block**			
HF	531	470	.886	HF	558	571	.985
	(45)	(84)	(.08)		(64)	(231)	(.01)
LF	605	562	.723	LF	696	783	.884
	(65)	(92)	(.09)		(136)	(196)	(.05)
Nonword	604	547	.859	Nonword	752	634	.874
	(71)	(96)	(.06)		(156)	(111)	(.07)
Normal block				**Even block**			
HF	610	546	.951	HF	609	555	.957
	(52)	(219)	(.04)		(75)	(101)	(.03)
LF	714	742	.831	LF	752	741	.797
	(98)	(159)	(.07)		(129)	(213)	(.08)
Nonword	715	772	.906	Nonword	722	792	.915
	(93)	(190)	(.05)		(122)	(340)	(.05)
Accuracy block				**Nonword-bias block**			
HF	698	593	.971	HF	633	478	.923
	(95)	(342)	(.03)		(66)	(201)	(.05)
LF	824	872	.848	LF	748	629	.731
	(135)	(239)	(.08)		(138)	(143)	(.12)
Nonword	815	923	.952	Nonword	654	787	.956
	(132)	(265)	(.03)		(117)	(208)	(.02)

Note. Standard deviations are shown in parenthesis. HF = high frequency words; LF = low frequency words; Acc = accuracy.

of observations, accuracy, and correct and error RT distributions for each condition. All data were fit simultaneously. The RT distributions were approximated by the .1, .3, .5 (median), .7, and .9 quantiles of each distribution, providing a summary of the distribution shape that is robust against the effects of outlier responses (see Ratcliff & Tuerlinckx, 2002, for justification). For a given set of parameter values, the predicted quantiles from the diffusion model are compared against the empirical quantiles, producing a X^2 value. The parameters are then adjusted using a SIMPLEX routine to minimize this value. For example, suppose a participant had accuracy of 88% and .1 and .3 quantiles of the RT distribution for correct responses of 440 ms and 480 ms, respectively. This means that 17.6% $(.3 - .1 * .88)$ of the responses for that condition fall between 440 ms and 480 ms. This value is compared against the predicted value from the diffusion model, and the difference is minimized through parameter adjustment.

Each block type (e.g., speed, normal, or accuracy) was fit independently to allow the parameters to capture changes in criteria. The average parameter values and X^2 values from the fitting routines are shown in Table 1. The degrees of freedom for X^2 value are given by $(K*11)-M$, where K is the number of conditions and M is the number of free parameters. The obtained X^2

values were all larger than the critical value (35), showing significant misses between the model and the data. However, because the sensitivity of the X^2 increases with the number of observations, even a small deviation would produce significant values (see Ratcliff, Thapar, & McKoon, 2004). Overall, the fit quality was comparable to previous applications of the model (Ratcliff, Thapar, Gomez, & McKoon, 2004; White, Ratcliff, Vasey, & McKoon, 2009) and captured the data well. Inspection of the model predictions from the best fitting parameters (not presented) showed that the prediction errors were small and symmetrically distributed around 0 (see White, Ratcliff, Vasey, & McKoon, in revision, for more details).

4.2. Results

The best fitting diffusion model parameters and behavioral data averaged across participants are shown in Tables 1 and 2, respectively. The data suggest that the instructions were effective. With speed/accuracy instructions, mean RTs for each condition were smallest for speed blocks, larger for normal blocks, and largest for accuracy blocks. Accuracy values showed the opposite pattern. With bias instructions, responses were fastest and most
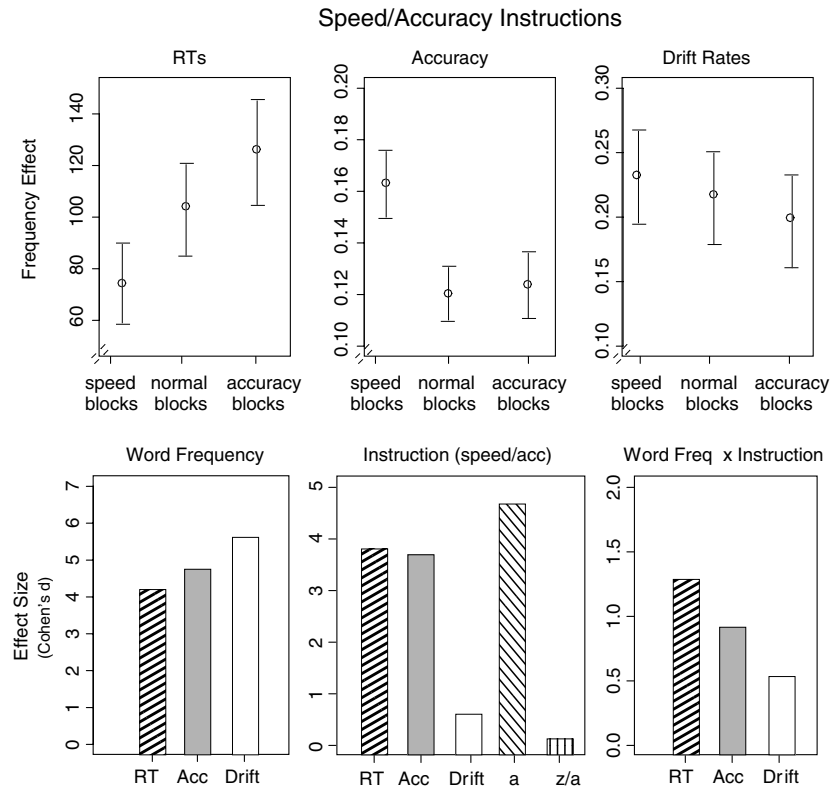
Fig. 2. Results from Experiment 1 for speed/accuracy instructions. The top panel shows the word frequency effect, defined as the difference between high and low frequency words for each measure. The bottom panel shows the effect sizes (from Cohen's *d*) from the ANOVAs for each measure. Error bars are 2 SEs.

probable when participants were biased to that response, and slowest and least probable when they were biased to the opposite response.

To assess the robustness and sensitivity of the dependent measures of lexical evidence, repeated measures ANOVAs were performed with frequency (high, low) and instruction (speed, normal, accuracy, or word, neutral, nonword) as within factors. Figs. 2 and 3 show the data and the effect sizes for each comparison. The ideal dependent measure would reflect only differences in lexical processing, which would be reflected by a large main effect of frequency, but no effect of instruction and no interaction between frequency and instruction. Repeated measures ANOVAs were also performed with instruction as the within factor on boundary separation (a) and the relative starting point (z/a) to assess how well they track the effects of instruction. Table 3 shows the results of the ANOVAs.

For speed/accuracy instructions, the effect of word frequency was largest for drift rates, followed by accuracy and RTs. Both RTs and accuracy comparisons showed significant main effects of the speed/accuracy instructions and significant interactions between frequency and instruction, whereas drift rates did not. The RT difference between high and low frequency words was largest for the accuracy blocks, followed by the normal and speed blocks, whereas the opposite was found for accuracy (see Fig. 2). In contrast, the drift rate difference between high and low frequency words did not reliably vary as a function of instruction. Instead, the effects of instruction were reflected in the diffusion model analysis by a large main effect of instruction for boundary separation (a).

The results for the bias instructions were similar and are shown in Fig. 3. Again, the main effect of word frequency was largest for drift rates, followed by accuracy and RTs. All three measures showed significant effects of the bias instructions, though drift rates were least affected. The interaction between instruction and word frequency was significant for accuracy and marginally

significant for RTs, but not for drift rates. As Fig. 3 shows, the effects of response bias were reflected in the diffusion model analysis by a large main effect of instruction for the relative starting point (z/a).

These results show that drift rates are more sensitive to stimulus evidence and less affected by speed/accuracy and response bias instructions compared to accuracy and mean RTs. Importantly, these were within-subject comparisons, showing that the effects of response criteria on RTs and accuracy cannot be completely eliminated by using a participant as their own baseline.

4.3. Previous applications of the diffusion model to study psychopathology

Although drift rates were most sensitive to the word frequency effect in the above experiment, the effect sizes were quite large, meaning that all three measures showed a reliable difference between the conditions. However, in situations with small effect sizes the extra sensitivity of the drift rates can be critical. We showed this recently by using the diffusion model to analyze lexical decision data from high- and low-anxious participants. As previously mentioned, it is well established that high-anxious individuals show biased processing of threat (e.g., Fox, 1993). Accordingly, it was predicted that they should be faster at identifying threat words compared to neutral words in lexical decision. However, several studies using RTs as dependent measures failed to find such an advantage (Hill & Kemp-Wheeler, 1989; MacLeod & Mathews, 1991; Mathews & Milroy, 1994). The failure to find threat bias in this task was taken to support the hypothesis that threat bias only occurs when there are multiple inputs competing for attention, a condition which the lexical decision task lacks (see MacLeod & Mathews, 1991).

Several models have been formulated to explain preferential processing of threatening information in anxious individuals. In

Table 3
ANOVA results from Experiment 1.

Speed/Accuracy instruction	Accuracy	Mean RT	Drift rates	a	z/a
Main effect: Word frequency	$F(1, 17) = 97.9$ $p < .001$	$F(1, 17) = 66.8$ $p < .001$	$F(1, 17) = 118.5$ $p < .001$	–	–
Main effect: Instruction	$F(2, 34) = 55.2$ $p < .001$	$F(2, 34) = 59.5$ $p < .001$	$F(2, 34) = 1.08$ ns	$F(2, 34) = 76.3$ $p < .001$	$F(2, 347) = 1.33$ ns
Interaction	$F(2, 34) = 3.32$ $p = .048$	$F(2, 34) = 6.8$ $p = .003$	$F(2, 34) = 1.16$ ns	–	–
Bias instruction					
Main effect: Word frequency	$F(1, 17) = 102.8$ $p < .001$	$F(1, 17) = 44.5$ $p < .001$	$F(1, 17) = 135.9$ $p < .001$	–	–
Main effect: Instruction	$F(2, 34) = 30.4$ $p < .001$	$F(2, 34) = 40.9$ $p < .001$	$F(2, 34) = 4.45$ $p = .021$	$F(2, 34) = 1.2$ ns	$F(2, 347) = 107.5$ $p < .001$
Interaction	$F(2, 34) = 19.8$ $p < .001$	$F(2, 34) = 2.67$ $p = .083$	$F(2, 34) = 0.86$ ns	–	–

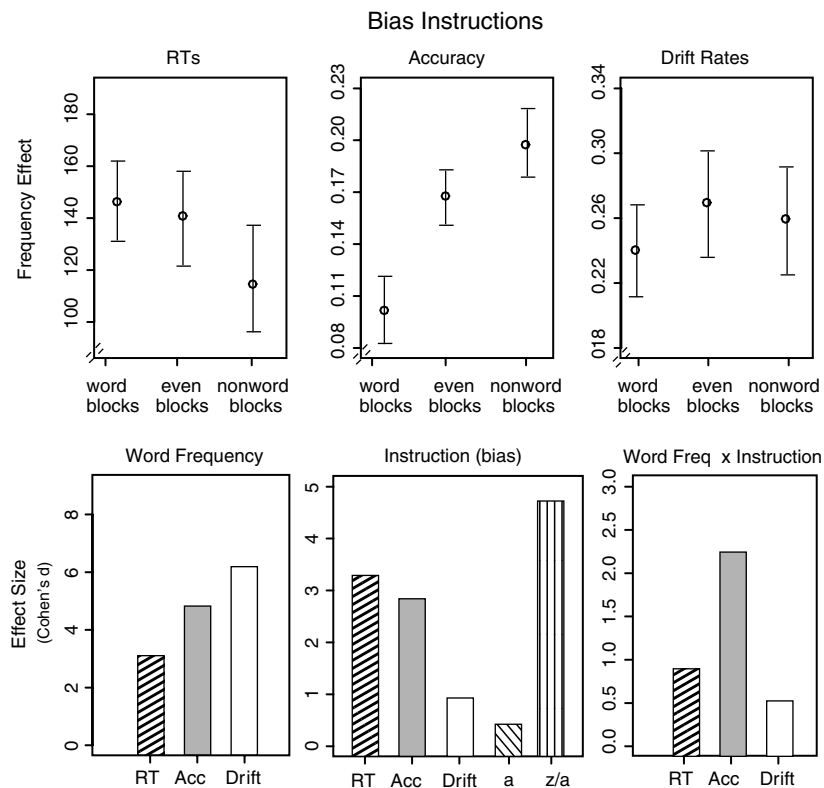Note. a = boundary separation; z/a = position of starting point relative to the boundaries.

Fig. 3. Results from Experiment 1 for response bias instructions. The top panel shows the word frequency effect, defined as the difference between high and low frequency words for each measure. *Word* refers to blocks with more words than nonwords, *nonword* refers to blocks with more nonwords, and *even* refers to blocks with an even number of words and nonwords. The bottom panel shows the effect sizes (from Cohen's *d*) from the ANOVAs for each measure. Error bars are 2 SEs.

relation to the diffusion model, these models provide a descriptive account of why anxious individuals have facilitated processing (i.e., higher drift rates) for threatening items compared to nonthreatening items. Many of these models have been adjusted, explicitly or implicitly, to account for the null findings from lexical decision tasks and the role of processing competition among inputs (e.g., Bishop, 2007; Mathews & Mackintosh, 1998). However, in their simplest form many of these models still predict biased processing of threat, regardless of whether there is more than one input to compete for attention (see White et al., in revision). Thus we hypothesized that although input competition might magnify the effect of the threat bias, making it easier to detect, there could still be small effects of threat bias without competition. To test this, individuals with high- and low-trait anxiety performed a single-string lexical decision task (without input competition) with threatening and matched

nonthreatening words presented among many neutral fillers. Threat bias was defined as an advantage for the threatening words over the nonthreatening words (i.e., higher accuracy, higher drift rates, or faster RTs). Across three separate subject groups, the behavioral measures showed only weak, nonsignificant trends hinting at a threat bias for anxious participants. In contrast, the diffusion model analysis showed a threat bias for high-anxious participants that replicated with each participant group (White et al., in revision). These results challenge the hypothesis that processing competition is necessary to demonstrate threat bias in anxious individuals, allowing for more parsimonious models of anxiety.

The previous study illustrates a situation in which the diffusion model advanced our understanding of psychopathology, but there has also been at least one instance in which a study of psychopathology advanced our understanding and use of the diffusion model itself. We previously used the diffusion model to investigate

Table 4
Simulated and recovered parameter values from the diffusion model and effect sizes from between-condition comparisons.

	a	z	Ter	η	s_z	s_t	p_o	v_{nonthr}	v_{threat}	v_{filler}	$v_{nonword}$	$X^2 (33)$
Simulation value	0.14 (.03)	0.07 (.01)	0.45 (.03)	0.10	0.06	0.20	0.001	0.27	0.33	0.30	−0.30	
Recovered 860 total obs	0.141 (.047)	0.070 (.023)	0.456 (.047)	0.128 (.062)	0.068 (.050)	0.198 (.043)	0.003 (.007)	0.306 (.093)	0.355 (.101)	0.338 (.078)	−0.338 (.086)	28.7 (11.7)
Recovered 120 total obs	0.188 (.075)	0.092 (.038)	0.458 (.058)	0.173 (.116)	0.053 (.062)	0.220 (.078)	0.001 (.016)	0.370 (.129)	0.441 (.152)	0.400 (.145)	−0.411 (.175)	20.79 (15.3)

Effect size (d)	Mean RT	Accuracy			Drift rate (120 obs)	Drift rate (860 obs)
Between	1.64 (.87)	1.42 (.67)			1.82 (.64)	2.86 (.97)
Within	2.76 (.85)	2.10 (.83)			2.93 (1.12)	3.42 (1.16)

Note. Recovered values are averages across simulated subjects (SDs in parenthesis). "860 total obs" indicates that there were 400 observations in each filler condition; "120 total obs" indicates that there were only 30 observations in each (see text for details). a = boundary separation; z = starting point; Ter = nondecision component; η = variability in drift across trials; s_z = variability in starting point; s_t = variability in nondecision component; p_o = probability of outlier responses, v = drift rate for each of the following conditions: threat words (threat), nonthreat words (nonthr), filler words (filler), and nonwords (nonword); d = effect size of comparison between threat and nonthreat words (from within or between-subjects comparisons).

emotional processing in dysphoric (i.e., moderately high levels of depressive symptoms) and nondysphoric college students (White et al., 2009). The goal was to assess differences in memory and lexical processing of positive and negative emotional words, which were presented among many neutral filler words. However, the emotional word pools used in the experiments only contained 30 words each. This left relatively few observations (especially for errors) to use in fitting the model, which would result in noisy parameter estimates. To remedy this, the model was fit to all conditions simultaneously, including the neutral filler conditions with hundreds of observations. The only parameter that was allowed to vary across condition was drift rate. Estimates for the other parameters, like nondecision time and boundary separation, were weighted by the number of observations for each condition. In this manner the filler conditions with many observations were used to constrain the fitting process, allowing the drift rates for the emotional words to be better estimated. In other words, the drift rates for the positive and negative emotional words were estimated based on boundary separation and nondecision estimates that were derived mostly from the filler conditions. The results of three experiments showed a bias for positive emotional words in the nondysphoric participants, but not in the dysphoric participants (White et al., 2009), consistent with previous research (Bradley & Mathews, 1983; Matt, Vazquez, & Campbell, 1992). Importantly, this difference in emotional bias was not significant when the diffusion model was fit only to the emotional conditions with few observations, nor was it significant in comparisons of RTs or accuracy. Since there are often a limited number of critical stimuli for use in studies of psychopathology, this approach provides a method for increasing sensitivity without requiring more critical stimuli. Next we present several simulations that were designed to illustrate this technique and the advantages of the diffusion model described above.

4.4. Simulations

There are two main goals of these simulations. The first is to demonstrate how filler conditions with a large number of observations can improve fits to conditions with relatively few observations (Ratcliff, 2008; White et al., 2009). The second goal is to show how each dependent measure, drift rates, accuracy, and RTs, is affected by individual differences in response components. In line with the work reviewed above, we simulated data from the diffusion model with parameter values similar to those obtained in a lexical decision experiment (White et al., in revision). The simulations were designed to reflect an experiment where a group of participants with high anxiety show a processing difference between threatening and nonthreatening words. There

were four conditions in the simulated experiment: threatening words, nonthreatening words, filler words, and filler nonwords. Importantly, the simulated conditions differed in the number of observations to reflect situations with a limited number of critical stimuli. The threat/nonthreat conditions had 30 observations each, whereas the filler conditions had 400 observations each. The total number of observations (860) reflects the number that could be obtained in a 45 min experiment. We set the drift rates in the simulations to reflect an advantage for the threat words over the nonthreat words. The drift rates were 0.33 for threat words, 0.27 for nonthreat words, 0.30 for filler words, and −0.30 for nonwords (the negative value indicates response at the bottom boundary). The remaining parameter values are shown in Table 4. In practice, there are differences across individuals in response caution, nondecision time, and response bias, so values of these components were drawn from normal distributions (means and standard deviations shown in Table 4) for 150 simulated subjects. The range of values chosen for this simulation was taken from the same experiment as the simulation values (White et al., in revision). In real experiments, drift rates would vary across participants as well, but they were kept constant in these simulations to focus on the effects of the other response components.

4.4.1. Using filler conditions to constrain fits

To assess the benefit of including the filler conditions with many observations, the diffusion model was fit to the simulated data in two ways. In one set of fits, the filler conditions only had 30 observations each (120 total), so each condition had relatively few observations. In the other set of fits, all of the observations were used to constrain the model, so the filler conditions had 400 observations each (860 total). When fitting each set of simulated data, only drift rate was allowed to vary between conditions. The contribution of each condition to the model fits was weighted by the number of observations in that condition, as described above. Thus in the fits with 860 total observations, the filler words and nonwords with 400 observations each contributed heavily to the estimates for boundary separation, nondecision time, response bias, and the variability parameters. The threat/nonthreat data with only 30 observations each determined the drift rates for the respective conditions, but did not greatly affect the estimates for the other parameters. Conversely, in the fits with only 120 observations, each condition contributed equally to the parameter estimation. The simulated and recovered parameter values are shown in the top portion of Table 4.

For the model fits with only 120 total observations, the estimates for a, eta, and the drift rates were too large (z was also inflated, but it remained fairly stable relative to a). The inflation of the

parameter estimates occurred because the small number of observations per condition meant that many simulated subjects had too few errors to properly constrain the fits of the model. With fewer than 5 errors in a condition, the quantiles for the error RTs for that condition cannot be accurately estimated. As a result, the variability components tend to be overestimated, leading to larger component estimates to compensate (Ratcliff & Tuerlinckx, 2002). This problem is most pronounced in tasks with high accuracy like lexical decision. Interestingly, although the drift rate estimates were inflated, the difference between the threat/nonthreat conditions appears relatively consistent with the difference used in the simulations.

In contrast, the model fits with 860 total observations accurately recovered the parameter values that were used to generate the data (within one standard deviation). Importantly, the drift rates for the threat/nonthreat conditions were accurately estimated, even though there were still relatively few observations in those conditions. With the data for all four conditions fit simultaneously, the filler conditions with 400 observations constrained the parameter values for all of the components other than the drift rates for the threat/nonthreat conditions. Consequently, the drift rates for these conditions were accurately estimated because the other decision components were mostly determined by the other responses.

Using the constraints from large-n conditions on parameter estimates for small-n conditions is an important step forward in application of the diffusion model. The constraints allow the model to estimate drift rates for conditions that require low numbers of observations, increasing the number of situations in which the model can be applied. The discovery of this approach was the direct result of research using the diffusion model to study psychopathology, and researchers in that domain are most likely to benefit from it.

4.4.2. Sensitivity of dependent measures

The other aim of the simulations was to show the different sensitivities of the dependent measures used to assess stimulus evidence, reinforcing the results from the lexical decision experiment reported above. We compared the measures that could be used to detect the difference between threat and nonthreat words: drift rates, mean RTs, and accuracy. Although 150 subjects were simulated, studies with patient populations are often limited in sample size, so 30 simulated subjects were randomly sampled from the total pool. For each sample, drift rates, predicted mean RTs, and predicted mean accuracy values for the threat and nonthreat conditions were calculated and then submitted to t-tests (both within and between). This process was repeated 2000 times (with replacement) to ensure stable results. The bottom portion of Table 4 shows the mean effect size (from Cohen's d) for each measure. Consistent with the results from Experiment 1, the difference between threat and nonthreat words was better detected by the drift rates than by the RT means or accuracy. This difference in sensitivity is highly reliable. For the between-subject comparisons, the 95% confidence intervals on the advantage of drift rates (860 obs.) compared to accuracy (in units of Cohen's d) are 1.26–1.62 for between-subject tests and 1.11–1.53 for within, and the interval for drift rates compared to mean RTs is .98–1.46 for between and .58–.86 for within. Drift rates are better able to capture the difference because they are determined by the full set of data for each condition, and because between-subject differences in response criteria are absorbed into the other components. Consistent with the results of Experiment 1, this was true even for within-subject comparisons.

4.4.3. Detecting differences between groups

The final simulation built on the previous one and was designed to mimic a full experimental situation where researchers are trying to detect processing differences between groups. The method was the same as above, except in this case data were simulated for two separate groups. In this simulation, one group, which we will denote as the high anxiety group, had a moderate advantage for threat words over nonthreatening words (i.e., threat bias), whereas the other group (low anxiety) had only a small advantage for threat words. Thus high anxiety is associated with a larger threat bias than low anxiety. The goal of the simulation was to show how well each measure detected this difference between the groups.

Table 5 shows the parameter values that were used to simulate the data. For the high anxiety group, the values were the same as used in the previous simulations, with a higher drift rate for threat (.33) compared to nonthreat words (.27). For the low anxiety group, there was only a small advantage for threat words (.28) over nonthreat words (.26). All of the other parameter values were the same as in the previous simulation, except that across-subject variability in drift rate was incorporated. Thus, in this set of simulations there were individual differences in a, Ter, z/a, and v, which reflects a more realistic experimental situation than the previous simulation that did not include variability in drift rates across simulated subjects.

The procedure was as follows: data were simulated for 150 subjects in each group, then the diffusion model was fit back to the simulated data (using all 860 observations). Out of the total set of simulated subjects, 20 were chosen from each anxiety group. Then, for recovered drift rates, simulated accuracy values, and simulated mean RTs, a mixed ANOVA was performed with condition (threat, nonthreat) as the within factor and group (high anxiety, low anxiety) as the between factor. This process was repeated 2000 times with replacement. The values for each measure are shown at the bottom of Table 5. The primary focus is on the interaction term of the ANOVA, which would be used to detect differential threat bias between the anxiety groups. For each run of the process, we recorded the p-value of the interaction and calculated the percentage of runs that showed a significant interaction ($p < .05$) for each dependent measure. Drift rates detected the true difference 83.8% of the time, compared to 57.2% for RTs and 38.1% for accuracy values.[1] Thus with relatively small differences between groups, mean RTs or accuracy values can be too insensitive to detect the difference.

These simulations reinforce the experimental work reviewed above. When the diffusion model provides an adequate account of the behavioral data, drift rates are more sensitive than accuracy or RTs in detecting differences between conditions. Further, the inclusion of filler conditions with many observations can improve fit quality for conditions with few observations. Since research on psychopathology often involves small effects and a limited number of critical stimuli, the diffusion model provides a promising alternative to analyses of RTs or accuracy. Next we present a new analysis in which the diffusion model is used to assess differences between individuals with high- and low-trait anxiety.

5. Experiment 2: Anxiety and error reactivity

Experiment 2 demonstrates an approach to fitting the diffusion model to compare processing between two groups of participants,

---

[1] The advantage of RTs over accuracy in these simulations is due to ceiling effects. In Experiment 1, accuracy was more sensitive to the word frequency effect because accuracy for low frequency words was low enough that ceiling effects were reduced. To show this, we repeated the between-group simulation with lower values of drift for threat and nonthreat words. The threat and nonthreat drift rates were .16 and .10, respectively, for the high anxiety group, and .12 and .10 for the low anxiety group. With this simulation, the percentage of detected group differences was 69% for drift rates, 32% for accuracy, and 15% for RTs, indicating that accuracy values are more sensitive than RTs unless there are ceiling effects.

Table 5
Results from group simulations.

	a	z	Ter	η	s_z	s_t	p_o	v_{nonthr}	v_{threat}	v_{filler}	$v_{nonword}$	X^2
Low anxiety: Simulation value	0.14 (.03)	0.07 (.01)	0.45 (.03)	0.10	0.06	0.20	0.001	0.26 (.07)	0.28 (.07)	0.30 (.07)	−0.30 (.07)	
Recovered value	0.145 (.04)	0.072 (.02)	0.456 (.04)	0.118 (.06)	0.064 (.04)	0.206 (.03)	0.003 (.01)	0.267 (.11)	0.286 (.11)	0.321 (.11)	−0.330 (.10)	30.3 (14)
High anxiety: Simulation value	0.14 (.03)	0.07 (.01)	0.45 (.03)	0.10	0.06	0.20	0.001	0.27 (.07)	0.33 (.07)	0.30 (.07)	−0.30 (.07)	
Recovered value	0.149 (.04)	0.075 (.02)	0.451 (.04)	0.121 (.06)	0.067 (.05)	0.206 (.03)	0.003 (.01)	0.299 (.12)	0.363 (.14)	0.333 (.10)	−0.331 (.11)	29.2 (12)

	Low anxiety		High anxiety		% Detection of between-group difference
Dependent measure	Threat	Nonthreat	Threat	Nonthreat	
Mean RTs	701 (117)	710 (108)	668 (106)	703 (128)	57.2
Accuracy	.919 (.07)	.909 (.09)	.950 (.06)	.924 (.09)	38.1
Drift rates	0.286 (.11)	0.267 (.11)	0.363 (.14)	0.299 (.12)	83.8

Note. Recovered values are averages across simulated subjects (SDs in parenthesis). Detection of between-group differences refers to the percentage of times (out of each sample of 20 participants) that each measure showed a significant difference between the high and low anxiety groups (see text for details). a = boundary separation; z = starting point; Ter = nondecision component; η = variability in drift across trials; s_z = variability in starting point; s_t = variability in nondecision component; p_o = probability of outlier responses; v = drift rate for each of the following conditions: threatening words (threat), nonthreatening words (nonthr), filler words (filler), and nonwords (nonword).

those with high- or low-trait anxiety. The experiment and simulations discussed above focused primarily on the advantage of drift rates as measures of stimulus evidence, but Experiment 2 focuses on how the diffusion model can aid researchers who are more interested in decision criteria. We focus on response style following correct and error responses in a recognition memory task. Individuals have been shown to have slower responses after an error than after a correct response (e.g., Hajcak, McDonald, & Simons, 2003), and individuals with high negative affect, which is common to both depression and anxiety, show a larger slowdown after errors than individuals with low negative affect (Robinson, Meier, Wilkowski, & Ode, 2007). This slowdown is thought to reflect increased caution so as to avoid the negative affect associated with additional negative feedback (Holroyd & Coles, 2002). The diffusion model allows us to test this directly, as any differences in caution should be reflected by increased boundary separation.

5.1. Method

Participants performed a recognition memory task in which they studied lists of words and then had to decide "old" or "new" according to whether test words had been studied or not. They were instructed to press the "/" key if the word had been studied and the "z" if it had not. Each participant completed 12 study lists of 26 words and 12 test lists of 52 words (half old and half new). Study words were presented for 1200 ms, and test words were shown until a response was made. Participants were instructed to respond quickly and accurately. After an incorrect response, the word "ERROR" was displayed for 750 ms before the next trial.

5.2. Stimuli

The stimuli were drawn randomly without replacement from the same pools of high and low frequency words as in Experiment 1. For the analyses presented here all conditions were collapsed into four conditions: old items after a correct response, new items after a correct response, old items after an error, and new items after an error.

5.3. Measure

The Spielberger Trait Anxiety Inventory (STAI; Spielberger, 1985) was used to assess anxiety level. This 20 item self-report

Table 6
Behavioral data from Experiment 2 averaged across participants.

	Low anxiety		High anxiety	
	Targets	Lures	Targets	Lures
Mean correct RT				
Post-correct	744 (92)	772 (114)	750 (99)	796 (116)
Post-error	740 (124)	774 (126)	766 (91)	809 (114)
Mean error RT				
Post-correct	847 (135)	842 (152)	887 (210)	809 (212)
Post-error	785 (166)	742 (164)	859 (154)	802 (133)
Accuracy				
Post-correct	.699 (.10)	.766 (.11)	.719 (.07)	.758 (.12)
Post-error	.687 (.10)	.708 (.16)	.691 (.11)	.703 (.16)
d'				
Post-correct	1.32 (.46)		1.35 (.49)	
Post-error	1.12 (.50)		1.14 (.56)	

Note. Standard deviations are shown in parenthesis. Post-Error refers to responses following errors; Post-Correct refers to responses following correct responses.

questionnaire is commonly used to assess subclinical levels of anxiety. Higher scores on the questionnaire indicate higher levels of trait anxiety.

5.4. Participants

There were 120 total participants in the experiment who received credit in an introductory psychology class. The upper and lower thirds of STAI scores were used to group participants. Low anxiety participants ($n = 42$) had a mean STAI score of 31.4, and high anxiety participants ($n = 42$) had a mean STAI score of 46.4.

5.5. Results

All responses faster than 250 ms or longer than 3 s were excluded (less than .8% of the data). The results from the experiment are shown in Table 6. Accuracy and d' were lower after an error than after a correct response, but there were no difference between high and low-anxious participants ($Fs < 1$). Unlike other studies (e.g., Robinson et al., 2007) there were no significant increases in mean RTs following error responses for either group ($F < 1$). Thus the analysis of behavioral data shows no differences in post-error responses as a function of anxiety.

Table 7
Best fitting parameter values from the diffusion model in Experiment 2.

Low anxiety	a	z/a	Ter	η	s_z	s_t	p_o	v_{target}	v_{lure}	$X^2(12)$
Post-correct	0.125 (.03)	0.544 (.09)	0.538 (.05)	0.233 (.09)	0.036 (.04)	0.173 (.10)	0.002 (.01)	0.128 (.09)	−0.249 (.11)	20.8 (8.3)
Post-error	0.121 (.03)	0.594 (.13)	0.516 (.06)	0.183 (.12)	0.028 (.03)	0.223 (.11)	0.007 (.02)	0.078 (.11)	−0.200 (.18)	24.3 (14)
High anxiety										
Post-correct	0.125 (.03)	0.554 (.08)	0.551 (.05)	0.235 (.08)	0.042 (.04)	0.176 (.07)	0.005 (.02)	0.153 (.09)	−0.229 (.11)	21.1 (10)
Post-error	0.136 (.04)	0.603 (.12)	0.515 (.06)	0.194 (.11)	0.042 (.04)	0.197 (.11)	0.007 (.03)	0.091 (.10)	−0.206 (.14)	19.81 (11)

Note. Recovered values are averages across subjects (SDs in parenthesis). a = boundary separation; z/a = position of starting point relative to the boundaries (values above .5 indicate a bias to respond "old"); Ter = nondecision component; η = variability in drift across trials; s_z = variability in starting point; s_t = variability in nondecision component; p_o = probability of outlier responses; v_{target} = drift rate for targets; v_{lure}=drift rate for lures. Post-correct refers to trials following a correct response. Post-error refers to trials following an error.
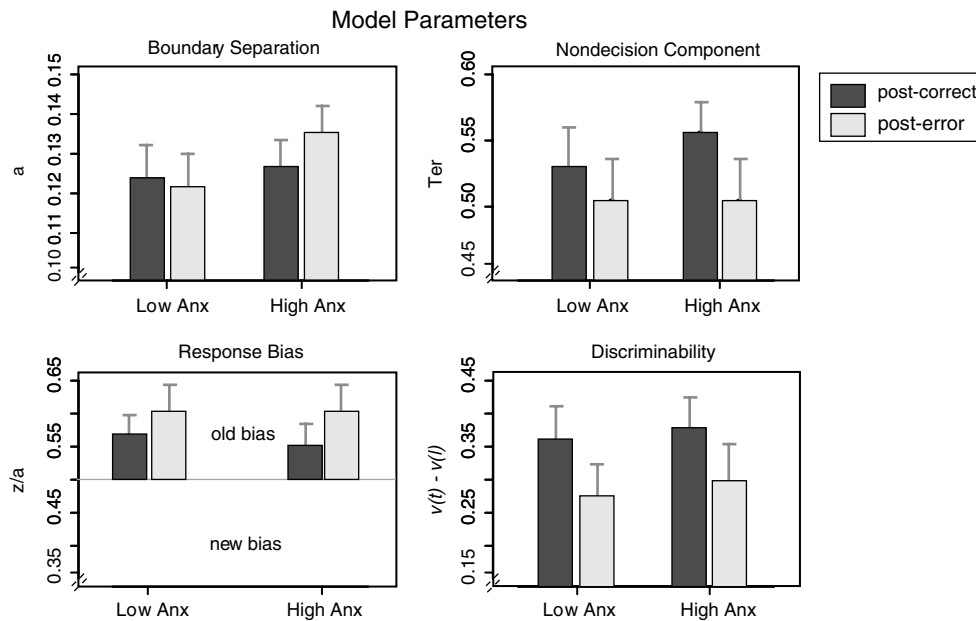

Fig. 4. Best fitting model parameters from Experiment 2 averaged across subjects. Dark bars are the best fitting parameters for responses following a correct response and light bars are for trials following an error. $v(t)$ = drift rate for targets; $v(l)$= drift rate for lures; a = boundary separation; z/a= relative position of starting point between boundaries; Ter = nondecision time. Values of response bias above .5 indicate a bias to respond "old." Error bars are 2 SEs.

5.6. Model fitting and results

The diffusion model was fit to each participant's data, resulting in parameter estimates for every participant that could be used in t-tests and ANOVAs in the same manner as RTs or accuracy. For this particular experiment, the model was fit separately to post-correct and post-error responses. Since each condition in the experiment had sufficient numbers of observations, there was no need to use fillers to constrain the model fitting like shown in the simulations. The results are shown in Fig. 4 and Table 7. The Chi-squared values shown in Table 7 were all near the critical value (21), suggesting good fits to the data. All comparisons used mixed ANOVAS with anxiety group (high, low) as the between factor and trial type (post-correct, post-error) as the within factor.

For both anxiety groups, nondecision time (Ter) was shorter after an error than after a correct response, $F(1, 82) = 25.28$, $MSE = .002, p < .001$. This result was unexpected, and could reflect an increase in impulsive responses following an error. However, the changes in Ter did not differ as a function of anxiety (interaction: $F(1, 82) = 1.88, MSE = .002, n.s.$). The starting point parameter showed an overall bias to respond "old" that was more pronounced after an error than after a correct response, $F(1, 82) = 8.25, MSE = .010, p < .01$, but the bias did not reliably differ between the two groups ($F < 1$).

Consistent with predictions, high anxiety participants increased their boundary separation after an error, whereas low anxiety participants did not. The interaction between anxiety group and trial type was significant for boundary separation, $F(1, 82) = 5.73$, $MSE = .0004, p = .018$, but there were no main effects of group or trial type (see Fig. 4). Subsequent comparisons showed a significant increase in boundary separation following errors for high-anxious participants, $t(41) = 2.17, p = .03$, and a nonsignificant decrease for low-anxious participants, $t(41) = -1.15, p = .25$. Drift rates were used to assess discriminability, which was operationalized as the difference between drift rates for old (positive values) and new (negative values) words, with a larger difference indicating better quality of evidence and thus better discriminability. There was an overall decrease in discriminability after errors for both groups (main effect of trial type: $F(1, 82) = 24.9, MSE = .014$, $p < .001$), but there was no main effect of anxiety group or interaction between trial type and anxiety group ($Fs < 1$).

The results of the diffusion model analysis are consistent with Robinson et al. (2007), showing that high-anxious participants were more cautious after committing an error. As mentioned above, this increase in caution has been suggested to stem from a desire to avoid the negative feedback associated with committing an error. Avoidance of negative or threatening information is a major component of high anxiety (e.g., Mathews, 1990), thus this

finding is consistent with typical high anxiety behavior. Since there were no group differences in RTs or accuracy, this difference in response caution between high- and low-anxious participants would not have been apparent without use of the diffusion model.

6. Discussion

We have identified and demonstrated ways in which sequential sampling models like the diffusion model can augment and improve studies of clinical populations and psychological disorders. The diffusion model allows researchers to compare different components of the decision process to identify the loci of processing differences between conditions or groups. There remain several areas in which future work is needed to expand the application of these models to studies of psychopathology, to which we turn next.

6.1. Broadening the application of sequential sampling models

Fitting processing models to data is more complex and involved than traditional methods of separately analyzing RTs and accuracy. The additional complexity and time required to use these models might discourage researchers from incorporating them into their research. Fortunately, there have been several attempts to provide user-friendly programs to aid in implementing the diffusion model. One approach is the EZ diffusion model of Wagenmakers, van der Maas, and Grasman (2007). This is a simplified version of the full diffusion model which extracts estimates for *Ter*, *a*, and v from behavioral data. However, it has been argued that the assumptions necessary for application of the EZ diffusion model, like having an unbiased starting point, are often not met in usual applications of the model (Ratcliff, 2008; but see Wagenmakers, van der Maas, Dolan, & Grasman, 2008). Additionally, the EZ model is more sensitive to outliers and less efficient at parameter recovery than the X^2 method. In light of these limitations, we recommend that the EZ diffusion model only be used for early data exploration, not for comparisons of parameter values across groups or conditions.

Another approach to broadening the application of the diffusion model comes from the development of statistical packages for implementation. Vandekerckhove and Tuerlinckx (2008) offer a MATLAB toolbox, DMAT, that allows researchers to implement the model in a fairly flexible manner. Voss and Voss (2007) have also developed a software package, *fast-dm,* that allows researchers to flexibly estimate diffusion model parameters from behavioral data.

There is also room for advancement in the types of data to which the diffusion model can be applied. One area for improvement involves fitting the model to data with relatively few observations, since the model parameters cannot be accurately estimated with too few data points. As mentioned above, practical constraints often limit the number of observations that can be obtained in studies with clinical populations. For example, if researchers were interested in determining if posttraumatic stress disorder patients had better or worse memory for information related to their traumatic event, they might have a limited number of words or pictures that sufficiently represent the event and are salient to the patients. We showed above how filler conditions with many observations can be used to improve the model fits for conditions with relatively few observations, providing a more accurate estimate of the drift rates for the critical stimuli. However, this procedure still requires the collection of many observations, which might not be feasible when studying patient populations. Patients with psychological disorders might lack the attentional capacity, motivation, or ability to perform hundreds of experimental trials, meaning there would not be enough observations to accurately fit the diffusion model. This is an important limitation of complex models like the diffusion model.

One method to deal with few observations is to fit the model to data averaged across participants. Unfortunately, it is well known that fitting averaged data can lead to biases and distortions in parameter estimates. Recently, Cohen, Sanborn, and Shiffrin (2008) showed that, for certain models, there are conditions in which fitting group data is superior to fitting individual data, particularly when each participant has very few observations. In support of this approach, Ratcliff, Thapar, and McKoon (2004) found that across several data sets the fits to averaged group data were consistent with the average of fits to individual participant's data. However, the data sets in that study had a moderate to large number of observations for each participant. In light of this, we do not currently recommend this approach since the behavior of the fitting methods have not been investigated when there are relatively few observations for each participant.

Another future direction for sequential sampling models like the diffusion model involves application to new experimental paradigms. Although these models have been shown to account for data from a range of two-choice tasks (see Ratcliff & McKoon, 2008), there are several tasks to which sequential sampling models have not yet been applied, such as the previously discussed probe detection task. Future research is necessary to determine whether and in what manner sequential sampling models can account for data from such paradigms. Since there are many two-choice tasks being used to investigate differences between groups or clinical populations, there is great promise in developing processing models to augment analysis of these paradigms.

6.2. Sources of evidence and complex decisions

In most applications, models like the diffusion model are agnostic about what sources of information contribute to the evidence used in the decision process. The model does not specify how the drift rate is determined, other than to say that it represents the quality of evidence for a response (but see Ratcliff, 1981; Smith & Ratcliff, 2009, for models that integrate the decision process with models of encoding processes). In this regard there is nothing inherent in the diffusion model that is related to a particular clinical disorder. It is only through thoughtful experimental design that the components of the model become meaningful to clinical researchers. For example, if differences in drift rates reflect enhanced or impaired memory for threatening information, it can help researchers better understand abnormal processing in anxiety, and potentially help identify individuals who are at risk of developing anxiety disorders. Thus while sequential sampling models provide a processing account of the decision, they only provide a descriptive account of the information feeding the decision process.

However, although it has not been described until this point, there is a sequential sampling model that provides a complimentary approach to the diffusion model, decision field theory (DFT; Busemeyer & Townsend, 1993; Roe, Busemeyer, & Townsend, 2001). DFT has a similar structure as the diffusion model, but it is meant to account for longer, more complex decisions between more than two options. A schematic of the model is shown in Fig. 5. In the model, the preference state for each option is determined by evaluating the relative valences of the options. If an option has a positive valence, this leads to approach behavior and increases the chance of selecting that option. In contrast to the diffusion model, there is only one response boundary and each option has its own accumulation path. Whichever path reaches the boundary first is selected as the response. Importantly, more than one stimulus attribute can be used to contrast the options, and the valence of each option at any given time is determined by the valences for whichever attribute is being attended to at that time. If, for example, an individual were choosing between two cars that differed
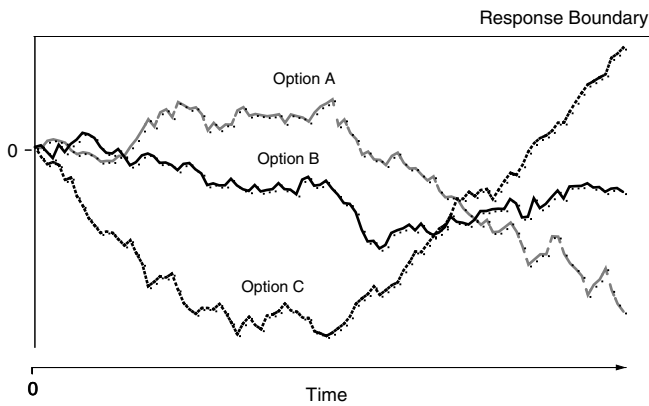
Response Boundary

Option A

0 -

Option B

Option C

0

Time

Fig. 5. An illustration of Decision Field Theory. Each path represents the preference state for one of the three options in the decision (A, B, or C). The first path to reach the top boundary is selected for response. See text for details.

in price and quality, focusing on quality might produce a preference for car A, while focusing on price might produce a preference for car B. The decision process in DFT is dynamic, because over the course of the decision the buyer might first focus on quality and then switch to price. Further, these attributes might have different importance, meaning a greater proportion of deliberation time is spent focusing on one over the other.

DFT provides a promising framework for researchers investigating psychopathology and clinical disorders. Busemeyer, Townsend, and Stout (2002) incorporated a dynamic model of needs into DFT. In this version of the model, the valence of a decision is determined by attentional values and motivational values. Motivational values are determined by current needs and how well each option satisfies those needs. The authors showed how DFT could account for the results of a study by Goldberg, Lerner, and Tetlock (1999), in which induced emotion led individuals to select harsher punishments for offenders. In the Goldberg et al. study, participants who watched a film where a criminal went unpunished were more likely to select a harsher punishment for a later, unrelated crime. DFT was able to account for the selection of stricter punishment by assuming that viewing the film increased one's need for punishment relative to their need for compassion. In this regard, the role of needs, emotions, and motivation can be accounted for in a decision process model (see also Busemeyer, Dimperio, & Jessup, 2007).

Similar to the diffusion model, DFT can identify the sources of processing differences more precisely than can be done with behavioral analyses. The addition of constructs like motivation to decision models can greatly advance research with clinical populations. Such advancements in decision models like DFT and the diffusion model will improve our understanding of the relationship between clinical disorders and abnormal cognitive, motivational, and decisional processes.

7. Conclusion

The experiments and simulations discussed in this article demonstrate the utility of sequential sampling models like the diffusion model in efforts to understand how cognitive processing differs as a function of psychopathology. These models can decompose behavioral data into meaningful psychological components, allowing researchers to assess potential differences in response caution, response bias, stimulus evidence, and encoding and response time. Differences in RTs and/or accuracy can be localized to the component(s) of processing that are responsible. Further, sequential sampling models can identify processing differences even when there are no apparent differences in the behavioral data, as the present analysis of anxiety and error responses demonstrated.

In this regard, sequential sampling models can provide alternative explanations to those based on behavioral data alone.

References

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, *133*, 1–24.

Becker, E. S., & Rinck, M. (2004). Sensitivity and response bias in fear of spiders. *Cognition and Emotion*, *18*, 961–976.

Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: An integrative account. *Trends in Cognitive Sciences*, *11*, 307–316.

Bradley, B., & Mathews, A. (1983). Negative self-schemata in clinical depression. *British Journal of Clinical Psychology*, *22*, 173–181.

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

Busemeyer, J. R., Dimperio, E., & Jessup, R. K. (2007). Integrating emotional processes into decision-making models. In *Integrated models of cognitive systems* (pp. 213–229). Oxford University Press.

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

Busemeyer, J. R., Townsend, J. T., & Stout, J. C. (2002). Motivational underpinnings of utility in decision making: Decision field theory analysis of deprivation and satiation. In *Emotional cognition: From brain to behaviour* (pp. 197–219). Amsterdam, Netherlands: John Benjamins Publishing Company.

Cohen, A. L., Sanborn, A. N., & Shiffrin, R. M. (2008). Model evaluation using grouped or individual data. *Psychonomic Bulletin & Review*, *15*, 692–712.

Ferguson, J. M., Wesnes, K. A., & Schwartz, G. E. (2003). Reboxetine versus paroxetine versus placebo: Effects on cognitive functioning in depressed patients. *International Clinical Pharmacology*, *18*, 9–14.

Fox, E. (1993). Attentional bias in anxiety: Selective or not? *Behaviour Research and Therapy*, *31*, 487–493.

Frewen, P. A., Dozois, D. J. A., Joanisse, M. F., & Neufeld, R. W. J. (2008). Selective attention to threat versus reward: Meta-analysis and neural-network modeling of the dot-probe task. *Clinical Psychology Review*, *28*, 307–337.

Goldberg, J. H., Lerner, J. S., & Tetlock, P. E. (1999). Rage and reason: The psychology of the intuitive prosecutor. *European Journal of Social Psychology*, *41*, 260–274.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. Oxford, England.

Hajcak, G., McDonald, N., & Simons, R. F. (2003). Anxiety and error-related brain activity. *Biological Psychology*, *64*, 77–90.

Hazen, R. A., Vasey, M. W., & Schmidt, N. B. (2009). Attentional retraining: A randomized clinical trial for pathological worry. *Journal of Psychiatric Research*, *43*, 627–633.

Hilbert, N. M., Niederehe, G., & Kahn, R. L. (1976). Accuracy and speed of memory in depressed and organic aged. *Educational Gerontology*, *1*, 131–146.

Hill, A. B., & Kemp-Wheeler, S. M. (1989). The influence of anxiety on lexical and affective decision time for emotional words. *Personality and Individual Differences*, *10*, 1143–1149.

Hindmarch, I. (1998). The behavioural toxicity of antidepressants: Effects on cognition and sexual function. *International Clinical Psychopharmacology*, *13*, S5–S8.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709.

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the implicit association test: A diffusion-model analysis. *Journal of Personality and Social Psychology*, *93*, 353–368.

Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lang, T. J., Moulds, M. L., & Holmes, E. A. (2009). Reducing depressive intrusions via a computerized cognitive bias modification of appraisals task: Developing a cognitive vaccine. *Behaviour Research and Therapy*, *47*, 139–145.

Lavy, E. H., van Oppen, P., & van den Hout, M. A. (1994). Selective processing of emotional information in obsessive compulsive disorder. *Behaviour Research and Therapy*, *32*, 243–246.

Lecci, L., & Cohen, D. (2007). Altered processing of health threat words as a function of hypochondriacal tendencies and experimentally manipulated control beliefs. *Cognition & Emotion*, *21*, 211–224.

Lemelin, S., Baruch, P., Vincent, A., Everett, J., & Vincent, P. (1997). Distractibility and processing resource deficit in major depression: Evidence for two deficient attentional processing models. *Journal of Nervous and Mental Disease*, *185*, 542–548.

Lemelin, S., Baruch, P., Vincent, A., Laplante, L., Everett, J., & Vincent, P. (1996). Attention disturbance in clinical depression: Deficient distractor inhibition or processing resource deficit? *Journal of Nervous and Mental Disease*, *18*, 114–121.

MacLeod, C., & Mathews, A. (1991). Biased cognitive operations in anxiety: Accessibility of information or assignment of processing priorities? *Behaviour Research and Therapy*, *29*, 599–610.

McFall, R. M., Treat, T. A., & Viken, R. J. (1997). Contributions of cognitive theory to new behavioral treatments. *Psychological Science*, *8*, 174–176.

Manguno-Mire, G. M., Constans, J. I., & Geer, J. H. (2005). Anxiety-related differences in affective categorizations of lexical stimuli. *Behaviour Research and Therapy*, *43*, 197–213.

Masten, C. L., Guyer, A. E., Hodgdon, H. B., McClure, E. B., Charney, D. D., Ernst, M., et al. (2008). Recognition of facial emotions among maltreated children with high rates of posttraumatic stress disorder. *Child Abuse & Neglect*, *32*, 139–153.

Mathews, A. (1990). Why worry? The cognitive function of anxiety. *Behaviour Research and Therapy*, *28*, 455–468.

Mathews, A., & Mackintosh, B. (1998). A cognitive model of selective processing in anxiety. *Cognitive Therapy and Research*, *22*, 539–560.

Mathews, A., & Milroy, R. (1994). Processing of emotional meaning in anxiety. *Cognition and Emotion*, *8*, 535–553.

Matt, G. E., Vazquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review*, *12*, 227–255.

McNally, R. J., Kaspi, S. P., Riemann, B. C., & Zeitlin, S. B. (1998). Selective processing of threat cues in posttraumatic stress disorder. *Journal of Abnormal Psychology*, *99*, 398–402.

McNally, R. J., & Reese, H. E. (2009). Information-processing approaches to understanding anxiety disorders. In *Oxford handbook of anxiety and related disorders* (pp. 136–152). New York: Oxford University Press.

McNally, R. J., Riemann, B. C., & Kim, E. (1990). Selective processing of threat cues in panic disorder. *Behaviour Research and Therapy*, *28*, 407–412.

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy*, *36*, 809–848.

Mogg, K., & Bradley, B. P. (1999). Some methodological issues in assessing attentional biases for threatening faces in anxiety: A replication study using a modified version of the probe detection task. *Behaviour Research and Therapy*, *37*, 595–604.

Mogg, K., Bradley, B. P., De Bono, J., & Painter, M. (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behaviour Research and Therapy*, *35*, 297–303.

Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. *Psychology and Aging*, *7*, 257–270.

Nigg, J. T., Silk, K. R., Starvo, G., & Miller, T. (2005). Disinhibition and borderline personality disorder. *Development and Psychopathology*, *17*, 1129–1149.

Norbury, R., MacKay, C. E., Cowen, P. J., Goodwin, G. M., & Harmer, C. J. (2008). The effects of reboxetine on emotional processing in healthy volunteers: An fMRI study. *Molecular Psychiatry*, *13*, 1011–1020.

Pisljar, M., Pirtosek, Z., Repovs, G., & Grgic, M. (2008). Executive dysfunction in late-onset depression. *Psychiatria Danubina*, *20*, 231–235.

Power, M. J., Cameron, C. M., & Dalgleish, T. (1996). Emotional priming in clinically depressed subjects. *Journal of Affective Disorders*, *38*, 1–11.

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

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, *88*, 552–572.

Ratcliff, R. (1985). Theoretical interpretations of speed and accuracy of positive and negative responses. *Psychological Review*, *92*, 212–225.

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

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*, 159–182.

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

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

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

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, *19*, 278–289.

Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*, *50*, 408–424.

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

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

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, *106*, 261–300.

Rheingold, A. A., Herbert, J. D., & Franklin, M. E. (2003). Cognitive bias in adolescents with social anxiety disorder. *Cognitive Therapy and Research*, *27*, 639–655.

Robinson, M. D., Meier, B. P., Wilkowski, B. M., & Ode, S. (2007). Introversion, inhibition, and displayed anxiety: The role of error reactivity processes. *Journal of Research in Personality*, *41*, 558–578.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, *108*, 370–392.

Rogers, M. A., Bradshaw, J. L., Phillips, J. G., Chiu, E., Vaddadi, K., Presnel, I., et al. (2000). Parkinsonian motor characteristics in unipolar major depression. *Journal of Clinical and Experimental Neuropsychology*, *22*, 232–244.

Ruchsow, M., Gron, G., Reuter, K., Spitzer, M., Hermle, L, & Kiefer, M. (2005). Error-related brain activity in patients with obsessive-compulsive disorder and in healthy controls. *Journal of Psychophysiology*, *19*, 298–304.

Siegle, G. J., Granholm, E., Ingram, R. E., & Matt, G. E. (2001). Pupillary and reaction time measures of sustained processing of negative information in depression. *Biological Psychiatry*, *49*, 624–636.

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

Smith, P. L., & Van Zandt, T. (2000). Time-dependent Poisson counter models of response latency in simple judgment. *British Journal of Mathematical and Statistical Psychology*, *53*, 293–315.

Spaniol, J., Voss, A., & Grady, C. L. (2008). Aging and emotional memory: Cognitive mechanisms underlying the positivity effect. *Psychology and Aging*, *23*, 859–872.

Spielberger, C. D. (1985). Assessment of state and trait anxiety: Conceptual and methodological issues. *Southern Psychologist*, *2*, 6–16.

Treat, T. A., & Dirks, M. A. (2007). Integrating clinical and cognitive science. In *Psychological clinical science: Papers in honor of Richard M. McFall*. New York: Psychology Press.

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

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

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, *7*, 208–256.

Verbruggen, F., & Logan, G. D. (2009). Proactive adjustments of response strategies in the stop- signal paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 835–854.

Voss, A., Rothermund, K., & Brandtstadter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, *44*, 1048–1056.

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

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

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

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

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

Weierich, M. R., Treat, T. A., & Hollingworth, A. (2008). Theories and measurement of visual attentional processing in anxiety. *Cognition & Emotion*, *22*, 985–1018.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). *Anxiety enhances threat processing without competition among multiple inputs: A diffusion model analysis*. Manuscript (in revision).

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2009). Dysphoria and memory for emotional material: A diffusion model analysis. *Cognition & Emotion*, *23*, 181–205.

Williams, R. M., & Hemsley, D. R. (1986). Choice reaction time performance in hospitalized schizophrenic patients and depressed patients. *European Archives of Psychiatry & Neurological Sciences*, *236*, 169–173.

Windmann, S., & Kruger, T. (1998). Subconscious detection of threat as reflected by an enhanced response bias. *Consciousness and Cognition*, *7*, 603–633.