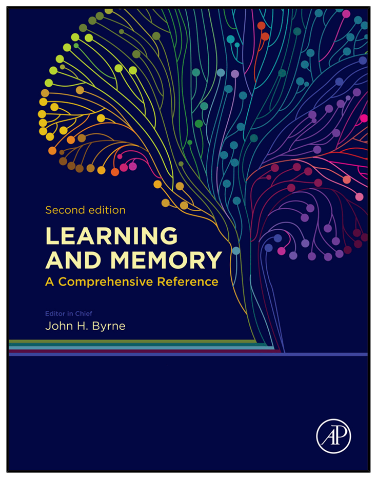

From Voskuilen, C., Ratcliff, R., Fennell, A., McKoon, G. (2017) Diffusion Models of Memory and Decision Making. In: Wixted, J.T. (ed.), *Cognitive Psychology of Memory*, *Vol*. 2 of *Learning and Memory: A Comprehensive Reference, 2nd edition*, Byrne, J.H. (ed.). pp. 227–241. Oxford: Academic Press.
http://dx.doi.org/10.1016/B978-0-12-809324-5.21045-6
ISBN: 9780128051597

Academic Press

2.13 Diffusion Models of Memory and Decision Making

Chelsea Voskuilen, Roger Ratcliff, Alex Fennell, and Gail McKoon, The Ohio State University, Columbus, OH, United States

2.13.1 Introduction

To make a decision about whether something is remembered, one must make some kind of comparison between the thing in question and information stored in memory and then make a decision based on the outcome of that comparison process. Most models of recognition memory are concerned primarily with how information is stored in memory, how the relevant information is extracted from the stored representations, and how these storage and retrieval processes may vary across tasks or stimuli. While these areas of focus are obviously necessary and important, what is often neglected is an account of how the information from memory is used to make a decision about whether a particular thing is remembered or not. We argue that this step of the process is not trivial and omitting it (or using overly simplified models of decision making) can change the conclusions drawn from the data. Additionally, most of the research on memory focuses primarily on the accuracy of memory decisions and ignores the associated response times (RTs). In this chapter, we describe one widely used model of decision making that can account for both accuracy and RTs, Ratcliff's (1978; Ratcliff and McKoon, 2008) diffusion model, and extensions to multichoice decision making and provide several examples of how the application of this modeling approach has led to different conclusions about underlying processes and representations. When a version of the diffusion model is applied to confidence judgment tasks, the results imply that different patterns of response proportions occur because of individual's decision-making preferences, not because more than one memory source or process is used in the decision. When the diffusion model is applied to data from older adults, the results imply that older adults make memory decisions more slowly than younger adults because they are more cautious about making mistakes, not because they have memory deficits. When the diffusion model is used along with multivariate pattern analysis to analyze electroencephalogram (EEG) data, the model allows us to determine which components reflect evidence used by the decision-making process and which do not. When decision-making models such as the diffusion model are used in conjunction with models of memory processes, the additional constraint provided by the decision-making model can help distinguish between competing theories of visual working memory and the strength-based mirror effect. Each of these applications will be discussed in more detail in the following sections.

2.13.2 Diffusion Model

The diffusion model is a model that was designed to account for both response proportions and response time distributions from simple, two-choice tasks in which decisions are made relatively quickly (Ratcliff, 1978; Ratcliff and McKoon, 2008). In the model, evidence in favor of a particular response is accumulated over time from a starting point toward one of two decision boundaries, as shown in **Fig. 1**. The starting point is denoted z and the boundaries are denoted 0 and a such that the distance between the two boundaries is determined by the parameter a. The rate of evidence accumulation is called the drift rate (ν) and is determined by the quality of the information extracted from the stimulus. There is noise in the accumulation of evidence within each trial, represented by the standard deviation (s). RT predictions are obtained by combining the decision time (the time taken for the accumulating evidence to reach one of the boundaries) with a uniformly distributed nondecision component. The nondecision component, which encompasses both encoding, the transformation of the stimulus representation to a decision-related representation, and response output processes, is assumed to be uniformly distributed with mean T_{er} and range s_t. One of the strongest pieces of support for the diffusion model is that it produces the relatively invariant right-skewed shape of RT distributions that is observed in a wide range of experimental data and that the model is incapable of fitting a range of other shapes that are not observed in experimental data (Ratcliff, 2002).

The values of drift rate and starting point (or, equivalently, the starting points of the decision boundaries) vary from trial to trial. The across-trial variability in drift rate is assumed to be normally distributed with a standard deviation of η. The variability in the

Learning and Memory: A Comprehensive Reference, Second Edition, 2017, 227–241

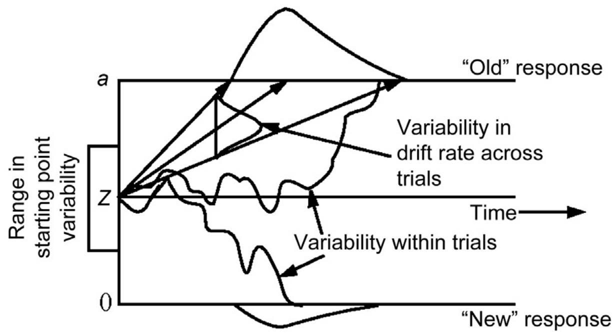

Figure 1 Diffusion model. Evidence accumulation begins at starting point *z* and continues until one of the boundaries (*a* or 0) is reached.

starting point is assumed to be uniformly distributed with a range of s_z. These distributional assumptions are not critical and the use of alternative distributions does not significantly change the estimates of the diffusion model's other parameters (Ratcliff, 2013). The inclusion of these parameters, however, is necessary for the model to be able to produce different patterns of correct and error RTs (Ratcliff and McKoon, 2008). Jones and Dzhafarov (2014) argued that if the forms of the across-trial variability distributions were unconstrained, the model could exactly match any data, but Smith et al. (2014) showed that this argument depends on eliminating within-trial noise in the decision process and so does not apply to the diffusion model.

The diffusion model can produce changes in accuracy and RTs in several ways. Smaller drift rates, representing a lower quality of evidence, will produce slower and less accurate responses. Conversely, larger drift rates will produce faster and more accurate responses. Larger distances between the boundaries (i.e., larger values of *a*) will produce slower and more accurate responses, and smaller distances will produce faster and less accurate responses. Differences in the nondecision component will affect response times but not accuracy.

By fitting both accuracy values and RT distributions, the diffusion model is able to distinguish between changes in the quality of information feeding into the decision process (i.e., the drift rate) and individual differences in how subjects come to a decision based on that information (i.e., different speed/accuracy preferences). In the diffusion model, changes in drift rate or changes in boundary separation affect both accuracy and reaction times, but do so in different and distinguishable ways. A decrease in drift rate will produce a decrease in accuracy and an increase in mean reaction time. A decrease in boundary separation will produce a decrease in accuracy and a decrease in mean reaction time (a speed/accuracy trade-off). Moreover, the parameters affect the shape of the reaction time distribution in different ways. Changes in drift rate primarily affect the spread of the RT distribution (i.e., the distribution is stretched out or compressed such that there is little change in the leading edge of the distribution and more change in the tails) while changes in boundary separation both shift and spread the RT distribution (i.e., the distribution is stretched out or compressed, but also shifted such that there is change in both the leading edge and tails of the distribution). The various patterns of experimental data that are observed in perceptual and cognitive tasks along with how the diffusion model accounts for them are presented in Ratcliff et al. (2015).

This ability to distinguish between changes in evidence and changes in decision-related processes allows one to compare performance across groups while controlling for individual differences in speed/accuracy preferences (i.e., controlling for the fact that individuals may have more or less conservative decision thresholds). The diffusion model allows one to investigate the processing differences giving rise to the observed differences in behavior and determine whether these differences are the result of differences in decision-making preferences (such as decision thresholds) or differences in the ability to extract the information necessary to make a decision. This ability to use the model to discriminate between sources of differences in processing is especially useful when analyzing data from different age groups, as will be discussed later.

The diffusion model has been used to detect changes in task performance that may produce inconsistent behavioral results. In the model, a small change in drift rate can produce small changes in both mean reaction times and accuracy rates that may or may not be significant. Thus examining just the behavioral data in these situations may yield inconclusive results where significant results are obtained only some of the time and for only some of the dependent measures. Because the model is able to disentangle the effects of different components of processing (e.g., response biases, speed/accuracy settings, quality of evidence, etc.), more direct comparisons of these components across subjects can yield significant differences even when behavioral measures (such as RT and accuracy) do not (e.g., White et al., 2009, 2010). White et al. (2010) examined memory for emotional words in dysphoric (i.e., moderately high levels of depressive symptoms) and nondysphoric college students and found that the nondysphoric students had drift rates biased in favor of the positive emotion words while the dysphoric students did not. Most importantly, this bias was not apparent from analyses of raw accuracy or RTs, only from the drift rate analyses.

The diffusion model can also be useful in situations where one dependent variable has reached ceiling or floor because it is still possible to observe and model changes in performance based on the changes in the other dependent variable. For accuracy, performance reaches ceiling at some point such that changes in accuracy are not observed for changes in difficulty. However, there may still be changes in reaction times for these difficulty levels (e.g., subjects may have the same accuracy but faster responses for easier stimuli) such that estimates of drift rate can still reflect changes in difficulty even after accuracy has reached ceiling

(Ratcliff, 2014). This means that the model is able to pick up on differences both across conditions and between subjects even in situations where performance on a task is close to ceiling such that it would be difficult to assess individual differences using only behavioral data.

2.13.3 Signal Detection Theory

Most of the work investigating recognition memory has utilized some form of signal detection theory (SDT) to describe how memory signals are translated into decisions (Atkinson and Juola, 1973; Banks, 1970; Bernbach, 1967; Bower, 1967; Donaldson and Murdock, 1968; Egan, 1958; Grasha, 1970; Kintsch, 1967; Kintsch and Carlson, 1967; Lockhart and Murdock, 1970; Norman and Rumelhart, 1970; Norman and Wickelgren, 1969; Ratcliff et al., 1994, 1992; Wickelgren and Norman, 1966; Yonelinas, 1994). In the signal detection framework, it is assumed that each tested item (or pair of items, as in associative recognition) has some value of memory strength that is normally distributed for each category of tested items (for example, "studied" or "not studied" words). The old/new decision can then be modeled by placing a single criterion on a dimension representing the memory strength of the test items. If the memory strength value for a test item is above the criterion, then an "old" response is made; otherwise, if the memory strength value is below the criterion, then a "new" response is made. Bias toward one of the response choices can be modeled by changes in the placement of the decision criterion, and multiple response options (such as confidence judgments) can be modeled by including additional decision criteria. These decision criteria can then be used to create receiver operating characteristic (ROC) functions, which are plots of the hit rate against the false alarm rate. These hit and false alarm rates are frequently converted to z-scores, resulting in a function called a z-ROC, and the shape of these functions has been used to infer the presence of multiple sources of memory information or memory processes (DeCarlo, 2002, 2003; Hilford et al., 2002; Kelley and Wixted, 2001; Rotello et al., 2004; Yonelinas, 1994; Yonelinas and Parks, 2007).

There are a number of problems with this SDT approach to memory modeling and understanding the cognitive processes underlying decision processes based on memory. First, the SDT approach ignores the reaction time associated with each response. Although there is a well-known relationship between the speed and accuracy with which people make decisions and consistent changes in RT distributions across response types and as a function of experimental manipulations (Atkinson and Juola, 1973; Juola et al., 1971; Ratcliff and Murdock, 1976; Wickelgren, 1977), most memory researchers only collect and analyze accuracy data. To provide a complete account of the decision process, it is important to consider both reaction time distributions and accuracy. A model like the diffusion model can account for both accuracy and RTs, produces RT distribution effects that are consistent with the empirical data (Ratcliff, 1978), and is able to account for speed-accuracy trade-offs by adjusting the boundary separation parameter. This enables the model to make comparisons across individuals who may have different speed–accuracy preferences as will be discussed in the section on aging presented later. Second, the SDT approach often ignores differences between individuals. Signal detection analyses are frequently conducted on data that have been averaged across subjects, so any differences between subjects are ignored or relegated to an appendix. There can be substantial differences in how subjects make responses such that it is not appropriate to only analyze averaged data (Malmberg and Xu, 2006; Ratcliff et al., 1994; Voskuilen and Ratcliff, 2016). Third, the SDT approach assumes that the only source of variability in the decision process is the variability in memory strength between items. This assumption leads to inappropriate conclusions about the z-ROC functions (Ratcliff and Starns, 2009, 2013; Starns et al., 2012a). Fourth, elaborations of SDT often include additional memory processes or additional sources of information to accommodate nonlinear z-ROC functions (DeCarlo, 2002, 2003; Hilford et al., 2002; Kelley and Wixted, 2001; Rotello et al., 2004; Yonelinas, 1994; Yonelinas and Parks, 2007). With the inclusion of reaction time data and individual differences, these additional processes are not always necessary to produce nonlinear z-ROC functions (Ratcliff and Starns, 2013; Voskuilen and Ratcliff, 2016) as will be discussed in the following section on confidence judgments.

All of these problems with SDT can be addressed by using models of the decision-making process such as the diffusion model and variations of it. These models are not memory models in the same way that SDT is not a memory model. A complete description of processing would involve a memory model producing the distributions of memory evidence that feed into the decision-making process (whether SDT or a diffusion model or some alternative decision-making model). However, the diffusion model and variations of it have been able to provide alternative explanations for important behavioral results in the memory literature because empirical effects that have been explained by memory processing have been shown to be explained more naturally by decision processes.

2.13.4 Confidence Judgments and Memory

Ratcliff and Starns (2009, 2013) developed a model that can be seen as a generalization of the diffusion model (RTCON2) that can explain the RTs of confidence judgments as well as the probabilities with which responses in the different confidence categories are made. This model has been applied to item recognition (Ratcliff and Starns, 2009, 2013) and associative recognition tasks (Voskuilen and Ratcliff, 2016) and provides an alternative account for some of the commonly observed z-ROC functions from these tasks. The explanation for the shapes of these functions is based on how subjects set their decision boundaries and is constrained by reaction time data. As such, the shape of the z-ROC function reflects individual differences in how subjects use confidence response scales as opposed to the type of information entering into the decision process from memory.

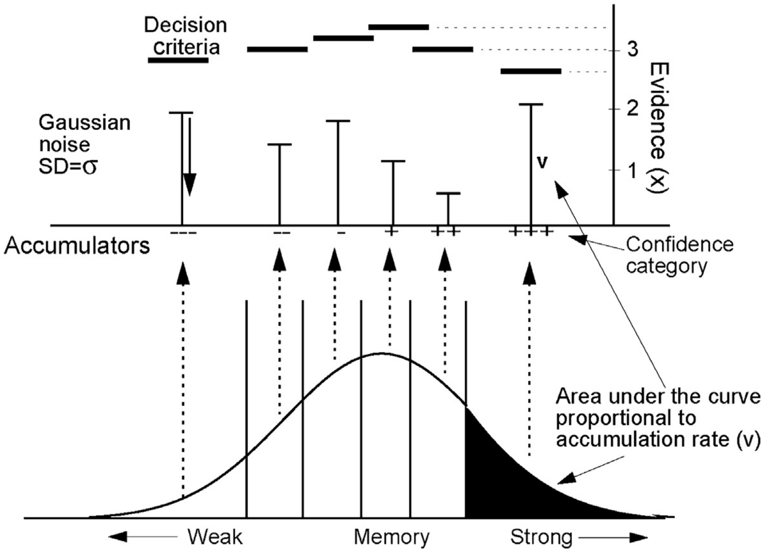

Figure 2 RTCON2 model. The distribution of evidence for an item on a given trial drives six accumulators (one for each confidence category). The proportion of the distribution between the confidence criteria on the memory strength dimension drives the drift rate for each confidence category. When one of the accumulators reaches its decision boundary, the corresponding response is made.

In the RTCON2 model, the decision process consists of racing diffusion processes (one for each confidence category, i.e., each different motor decision – key to be pressed). Evidence on a single trial (i.e., the memory strength for a particular item) is distributed across the evidence strength dimension. These item distributions have a standard deviation of 1 and their mean location varies from trial to trial (as in SDT). The bottom portion of Fig. 2 illustrates how the distribution of evidence for a single item is mapped to the decision process. As in SDT, multiple confidence criteria are used to divide the strength dimension into multiple response regions corresponding to different levels of confidence. Each response region has its own accumulator and decision boundary, as shown in the top portion of Fig. 2, and the diffusion processes race until one of them reaches its decision boundary and the corresponding response is made. The drift rate for each diffusion process is determined by the area of the evidence distribution in each response region and the accumulators are updated according to a constant summed evidence algorithm (see Ratcliff and Starns, 2013 for more detail).

In SDT, confidence is determined solely based on a discrete value of memory evidence and the position of the confidence criteria. In RTCON2, confidence responses are affected by the relative heights of the decision boundaries as well as the evidence strength and confidence criteria. For example, a high-confidence response region may have a higher decision boundary such that more evidence must be accumulated for that response to be selected. The height of the decision boundary would cause that particular response to be selected less often and with a longer reaction time than if that response region had a lower decision boundary, even for items that have a high mean value of evidence. Thus in RTCON2, confidence is not merely a function of accuracy but is also affected by individual differences in how subjects set response thresholds.

Standard SDT with normal distributions of evidence is unable to account for the nonlinear z-ROC functions observed in some associative recognition experiments (Glanzer et al., 2004; Hilford et al., 2002; Kelley and Wixted, 2001; Qin et al., 2001; Slotnick and Dodson, 2005; Slotnick et al., 2000; Wixted, 2007; Yonelinas, 1997, 1999). This has prompted theorists to add extra memory processes (Yonelinas, 1994; Yonelinas and Parks, 2007) or extra sources of information (DeCarlo, 2002, 2003; Hilford et al., 2002; Kelley and Wixted, 2001) to standard SDT to account for these findings. These models are all focused on fitting ROC functions and do not account for RTs. However, RTCON2 can account for nonlinear z-ROC functions through changes in the height of the decision boundaries as opposed to changes in the memory process while also fitting RT distributions.

Three sets of simulated data are plotted in Fig. 3 (these are typical of observed patterns of data). For these simulations, the evidence feeding into the decision was held constant (i.e., all of the simulations use the same drift rate parameters and confidence criteria) while the decision boundaries were varied. For the first data set the boundaries were flat, for the second they were inverted U-shaped, and for the third they were U-shaped (as shown in the third row of Fig. 3). These changes in decision boundaries affect both RTs and response proportions. Confidence categories with relatively smaller decision thresholds yield faster RTs such that the RT quantiles exhibit the same pattern across confidence categories as the decision boundaries (as shown in the first row of Fig. 3). Confidence categories with relatively smaller decision thresholds also yield a larger proportion of responses. In the bottom row of Fig. 3, note that when the low-confidence categories (numbers 3 and 4) had relatively lower thresholds (as in the third column) there were a larger proportion of responses in those categories than when those categories had equal or higher thresholds (as in the first and second columns). However, the drift rate still plays a larger role in response proportions than boundary height (i.e., these simulations were generated with relatively strong drift rates and so the largest proportions of responses are for high-confidence correct responses). When the cumulative hit and false alarm rates for each confidence criteria are standardized and plotted against each other (as in the second row), we see that these z-ROC functions also exhibit the same pattern as the decision boundaries. That

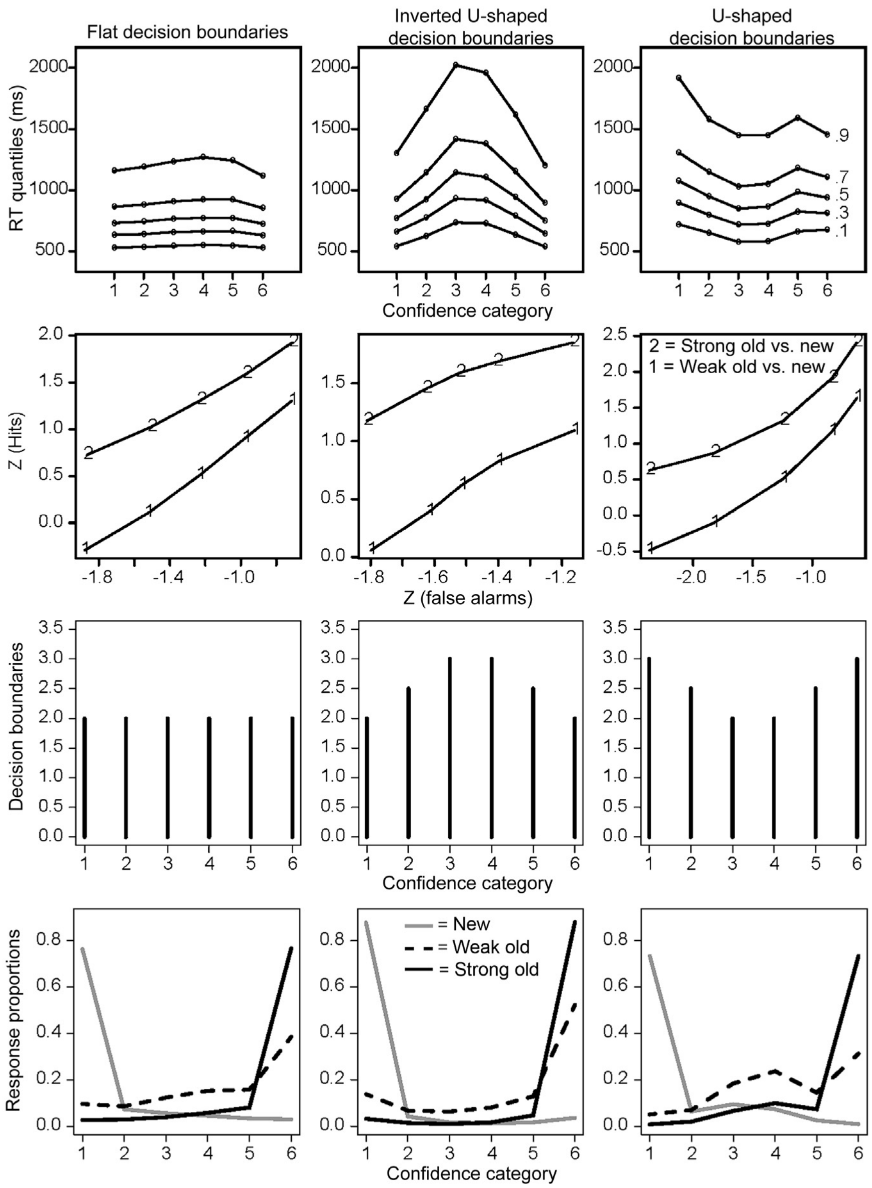

Figure 3 Simulations from RTCON2. Each column shows one set of simulated data. The first row shows the 0.1, 0.3, 0.5, 0.7, and 0.9 RT (response time) quantiles for each confidence category for one of the conditions (the strong old condition). The second row shows the z-ROC functions, the third row shows the values of the decision boundaries used to generate the data, and the fourth row shows the response proportions for each confidence category for each of the three conditions. For all three sets of simulated data, the following parameter values were used: nondecision time was 350 ms; range in nondecision time was 100 ms; within-trial noise was 0.1; scale on drift was 0.04; range in decision boundaries was 1; confidence criteria were 0.0, 0.5, 1.0, 1.5, and 2.0; mean drift rates were −0.5, 1.5, and 2.5 (for new, weak old, and strong old conditions, respectively); and the mean drift rates were normally distributed across trials with a standard deviation of 0.5 (for all conditions). Decision boundary values are shown in the figure.

is, the inverted U-shaped boundaries produced inverted U-shaped z-ROC functions and so on. This example demonstrates the ability of RTCON2 to produce nonlinear z-ROC functions with a single source of information by changing aspects of the decision-making process rather than changing the information upon which the decision is based. It also demonstrates the importance of modeling the decision process as well as the representation of information in memory. Each of these patterns of RT distributions and response proportions was produced using identical evidence distributions. Without a model of the decision-making process, the differences in these patterns of data would be incorrectly interpreted as differences in memory rather than differences in decision-making processes.

RTCON2 has been fit to item recognition data (Ratcliff and Starns, 2013) and associative recognition data (Voskuilen and Ratcliff, 2016) and was able to fit a wide variety of z-ROC shapes from these tasks with a single source of information from memory.

The model's explanation for these different shapes is based primarily on individual differences (i.e., how people set their decision boundaries) rather than additional memory processes or sources of information. In their experiment 3, Voskuilen and Ratcliff (2016) fit RTCON2 to data from 34 subjects from an associative recognition task with confidence judgments. Out of these 34 subjects, exactly half of them had z-ROC functions that were significantly different from linear and the other half had z-ROC functions that were not significantly different from linear. That is, there was considerable variety in the data patterns across subjects. Select fits from this experiment are shown in **Fig. 4** (these subjects were chosen to illustrate the range of patterns the model can fit). The first two rows in **Fig. 4** plot the RT quantiles for each confidence response with the six response keys plotted on the x-axis (the "sure rearranged" category is labeled 1 and the "sure intact" category is labeled 6) and the RT quantiles plotted vertically with each line representing a reaction time quantile. The numbers plotted represent the empirical data and the lines represent predictions from the model. Response categories with fewer than 5 responses are omitted, and only the median quantiles are plotted for categories with between 5 and 10 responses. The third and fourth rows in each figure plot the empirical and predicted z-ROC and ROC curves for each subject. The fifth row plots the decision boundaries for each confidence response and the sixth row plots the response proportions (both empirical data and model predictions) for each confidence response and condition.

The model predictions match the data quite closely for these subjects. As shown in the sixth row, the model is able to reproduce the response proportions from subjects who spread their responses fairly evenly across the confidence categories (subjects 7 and 18) as well as those who used some confidence responses much more often than others (subjects 3 and 13). As shown in the third and fourth rows, the model-predicted ROC curves match the data closely, even for subjects whose performance is near ceiling (subject 13) or floor (subject 12). The model is able to produce both linear z-ROC functions (subjects 7 and 18) and nonlinear z-ROC functions (subject 13). However, while RTCON2 was able to account for most of the response patterns and reaction times in these experiments, the model was not always able to produce the U-shaped z-ROC functions observed for some of the subjects in this experiment. As described previously, RTCON2 is able to produce nonlinear z-ROC functions by changing the relative height of the decision boundaries such that there is a correspondence in the model predictions between the shape of the RT quantiles and the shape of the z-ROC function. Unsurprisingly then, the model had trouble producing U-shaped z-ROC functions when the shapes of the RT quantiles were not consistent with the shapes of the z-ROC functions.

RTCON2 also had difficulty producing some of the patterns of response proportions associated with U-shaped z-ROC functions. In the model, evidence is represented as a normal distribution (with an SD of 1) on some memory strength dimension and the position of this distribution varies across trials (according to another normal distribution with mean μ and SD s). This representation of evidence constrains the possible patterns of response proportions that the model can produce. For example, to produce chance performance for some confidence level, the evidence distributions for "intact" and "rearranged" items must have similar area in that response region. However, such a restriction affects the area of the evidence distributions in all of the other response regions since they are all determined by the location of the normal distribution of evidence. Thus the model has difficulty producing, for example, extreme changes in performance for neighboring response options. This representation of evidence has previously provided a good fit to data (Ratcliff and Starns, 2013). However, as discussed in Ratcliff and Starns (2013), the distribution of memory strength across trials does not need to be a normal distribution and could instead take the form of some distribution predicted by a memory model or a mixture model where the mean of the distribution of memory strength for associative information may depend on other factors (such as item information). Such changes might enable the model to fit more of the U-shaped z-ROC functions. This type of combined modeling approach would allow researchers to take advantage of the ability of RTCON2 to distinguish between the information feeding into a confidence response and individual differences in how the confidence response scale is used.

Much of the research attempting to distinguish between models of memory has been focused on slight differences in the shape of the z-ROC functions and has ignored RT data. The variation in the shapes of the z-ROC has been used to make claims about the number of processes involved in a memory decision, the nature of the evidence involved in the decision, and specific characteristics of the decision process. RTCON2 was designed to account for both RT distributions and response proportions from confidence judgment tasks while making simple assumptions about the representation of evidence in memory. Without assuming more than a single dimension of memory strength, RTCON2 was able to produce a variety of ROC and z-ROC shapes (Ratcliff and Starns, 2013; Voskuilen and Ratcliff, 2016). In our three experiments, some of our subjects had slightly U-shaped z-ROC functions, but other subjects had linear z-ROC functions or inverted U-shaped z-ROC functions. So far, none of the existing memory models can handle the full observed pattern of RTs and response proportions across confidence levels and none of them, to our knowledge, would account for the diversity of z-ROC shapes observed in these experiments. Thus the specific ROC and z-ROC shapes cannot be used solely to infer the nature of evidence from memory but are also indicative of differences in how different subjects choose to set decision boundaries when using confidence response scales.

The diffusion model has also been used to validate an assumption made by signal detection models concerning the variability of memory strengths for old and new items. In recognition memory, the slope of the z-ROC function is generally less than 1. In a signal detection framework, this would occur when the distribution for old items is more variable than the distribution for new items. In contrast, according to Yonelinas' dual-process model (1994, 1997), recognition decisions are made based on two independent decision processes—a continuous familiarity process and an all-or-nothing recollection process. Familiarity is thought to follow a signal detection process where the distributions for old and new items have equal variability, so z-ROC slopes less than 1 occur only when some proportion of the responses are based on recollection. To test these accounts, Starns and Ratcliff (2014) fit a large number of recognition memory data sets with two versions of the diffusion model: one where the variability parameters for evidence for old and new items were constrained to be equal, and one where the two variability parameters were allowed to differ. The model with unequal variability parameters provided the best fit and the estimates for old item variability were larger than the estimates for new

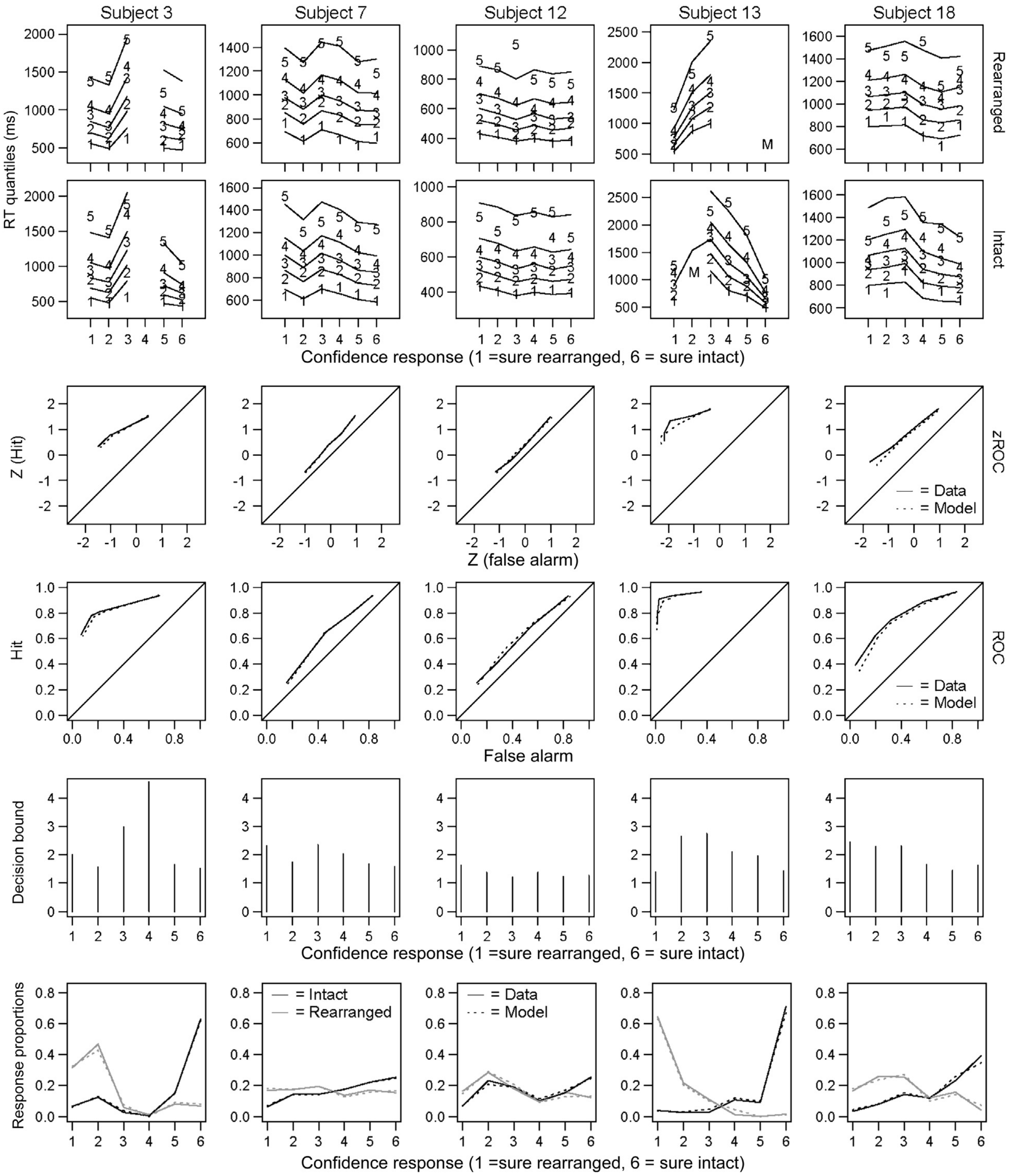

Figure 4 Select model fits from Voskuilen and Ratcliff (2016). The first two rows plot the RT (response time) quantiles for each confidence response with the six response keys plotted on the x-axis (the "sure rearranged" category is labeled 1 and the "sure intact" category is labeled 6) and the RT quantiles plotted vertically with each line representing a reaction time quantile. The numbers plotted represent the empirical data and the lines represent predicted data from the model. In conditions where subjects made between 5 and 10 responses the median RT is plotted as an "M" and the other quantiles are not included. Conditions where subjects made fewer than five responses are omitted from the figure. In conditions where the model predicted between 5 and 10 responses only the median RT is plotted and the other quantiles are not included. Conditions where the model predicted fewer than five responses are omitted from the figure. The third and fourth rows in each figure plot the empirical and predicted z-ROC and ROC curves for each subject. The *solid lines* depict the empirical data and the *dashed lines* depict the model predictions. The fifth row plots the decision boundaries for each confidence response and the sixth row plots the response proportions (both empirical data and model predictions) for each confidence response and condition. The *solid lines* depict the empirical data, the *dashed lines* depict the model predictions, the *black lines* depict responses for "intact" pairs, and the *gray lines* depict responses for "rearranged" pairs.

item variability. Starns et al. (2012a) collected recognition memory data with a probability manipulation (i.e., there were different proportions of "old" items in the test lists) and had subjects emphasize either speed or accuracy across sessions. For both the speed and accuracy sessions, the slopes of the z-ROC functions were less than 1, but the RTs in the speed condition were fast enough (mean RT = 526) that recollection should not have been contributing to the decision process (see McElree et al., 1999). These results are thus inconsistent with the idea that the recollection process is responsible for z-ROC slopes that are less than 1. As in Starns and Ratcliff, a diffusion model with unequal variability parameters fit better than one with equal variability parameters and was able to produce z-ROC slopes less than 1. Both of these studies provide support for the unequal variance assumption of SDT and do so without relying on accuracy-only measures.

2.13.5 Aging and Memory

The diffusion model allows us to determine which components of processing are the sources of differences in RTs. In item recognition, as in other cognitive tasks, older adults consistently make decisions more slowly (Salthouse, 1996; Verhaeghen and Salthouse, 1997) but not necessarily less accurately than younger adults (Balota et al., 2000; Bowles and Poon, 1982; Craik, 2008; Craik and Jennings, 1992; Erber, 1974; Naveh-Benjamin, 2000; Old and Naveh-Benjamin, 2008; Rabinowitz, 1984; Schonfield and Robertson, 1966). This slowdown has often been interpreted as a generalized processing deficit in the central nervous system (e.g., Cerella, 1985; Salthouse, 1996; Deary, 2000). However, using the diffusion model to analyze data from these types of tasks allows us to investigate the processing differences giving rise to these types of observed differences in behavior and determine whether these differences are the result of differences in decision-making preferences (such as decision thresholds) or differences in the ability to extract the information necessary to make a decision.

In item recognition experiments, subjects study lists of items (items may be words, pictures, etc.) and then, during a later test, must distinguish between items that were on the previous study list ("old" items) and items that were not on the previous study list ("new" items). Other studies investigating aging and item recognition have found only small effects of age on accuracy in this task (e.g., Balota et al., 2000; Bowles and Poon, 1982; Craik, 1994; Craik and Jennings, 1992; Erber, 1974; Gordon and Clark, 1974; Kausler, 1994; Naveh-Benjamin, 2000; Neath, 1998, Chap. 16; Old and Naveh-Benjamin, 2008; Rabinowitz, 1984; Schonfield and Robertson, 1966). Ratcliff et al. (2004, 2010) compared older and younger adults' performance on an item recognition task. Consistent with previous research, they found that older adults responded more slowly and were slightly less accurate than younger adults. When their data were fit with the diffusion model, the drift rates for the older adults were not significantly different from the drift rates for the younger adults. Given that drift rate represents the quality of evidence upon which the decision is made, this indicates that the information retrieved from memory by the older adults is not significantly worse than that retrieved by the younger adults. Older adults did, however, have significantly longer nondecision times and significantly higher decision thresholds than younger adults. These differences explain the older adults' slower RTs. In summary, according to the diffusion model analysis, older adults take more time to extract the relevant information from memory and physically make a response (i.e., have a longer nondecision time component) and set more conservative decision thresholds, but the quality of the information they retrieve from memory is comparable to that of younger adults.

In an associative recognition memory experiment, participants study pairs of items and are then asked to distinguish between pairs of items that were previously studied together ("intact") or studied separately ("rearranged"). In contrast to item recognition, performance in associative recognition tasks becomes significantly worse as individuals age (Bastin and Van der Linden, 2005; Buchler and Reder, 2007; Craik and Broadbent, 1983; Craik and McDowd, 1987; Healy et al., 2005; Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2004a, 2004b, 2003; Ratcliff and McKoon, 2015; Ratcliff et al., 2011). Based on these findings, it is generally concluded that aging minimally impacts item information but greatly impacts associative information. However, similar to the situation with item recognition, these studies only measured accuracy. Ratcliff et al. (2011) had three groups of subjects (college age, 60- to 70-year-olds, and 75- to 90-year-olds) who performed both item and associative recognition tasks and recorded both RTs and accuracy. Consistent with previous behavioral findings, accuracy decreased with age, a small amount for the item recognition task and a large amount for the associative recognition task, and RTs increased with age for both of the tasks. The data were fit with the diffusion model and, consistent with previous modeling results, older adults had longer nondecision times and higher decision thresholds than younger adults. For the item recognition task, there were small but significant differences in drift rates across age groups. For the associative recognition task, there were larger significant differences in drift rates. Use of the diffusion model in these studies allows us to determine which slowdowns in RTs reflect a decrease in the quality of information from memory and which slowdowns are the result of differences in decision-making preferences.

This study also collected IQ information and found that IQ scores were positively correlated with drift rates, but not other model parameters (such as boundary separation or nondecision times). Fig. 5 plots mean parameter values from both the item and associative recognition tasks for each age and IQ group. In the first row, we see that drift rate varies both with age and IQ. Older adults and adults with lower IQ have smaller values of drift rate. However, we also see an interaction between age and IQ in the associative recognition task such that IQ has a larger effect on drift rate for younger adults and smaller effect for older adults. For the young adults, there are large differences in the drift rates for the different IQ levels. For the very old adults, there are only small differences in the drift rates for the different IQ levels. This demonstrates that the effect of aging on the quality of information extracted from memory (i.e., the drift rate) is a leveling off with age as opposed to a fixed amount of decline or a "use it or lose it" pattern of decline (although note that this is a cross-sectional design, not longitudinal). That is, for the oldest adults there was essentially no effect of

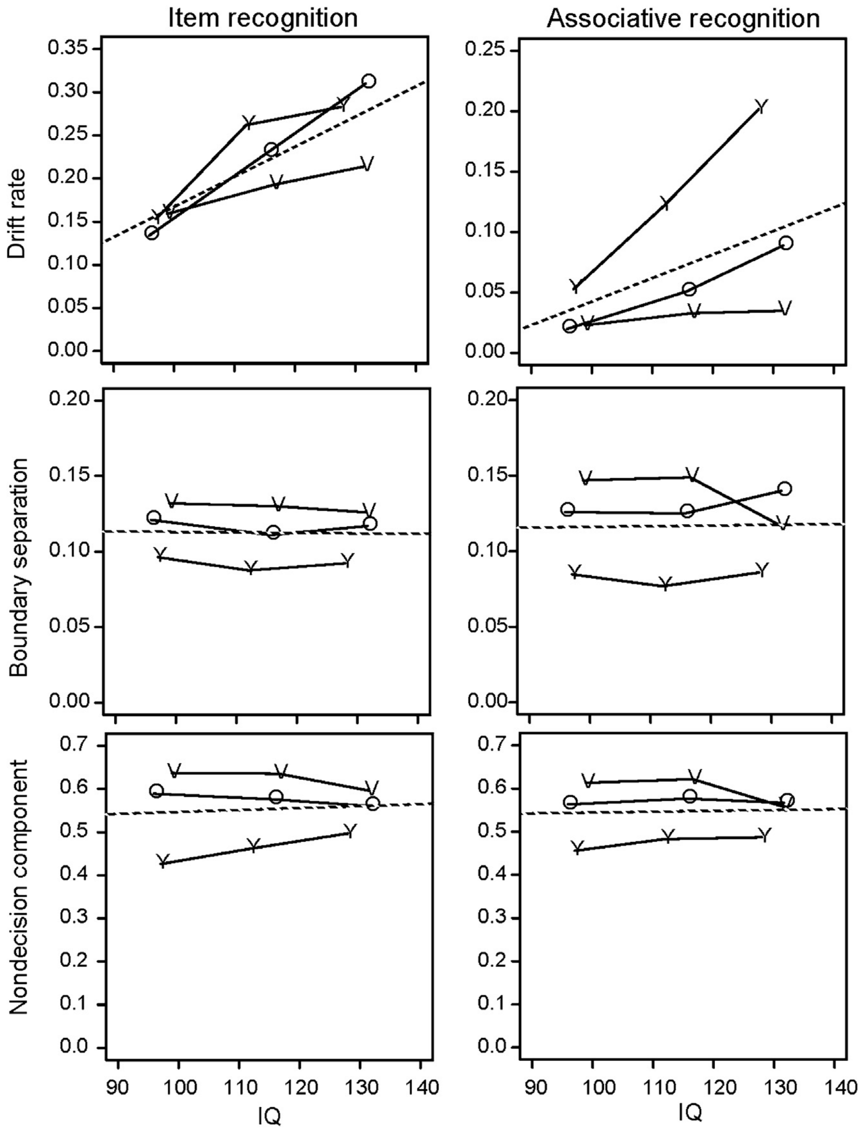

Figure 5 Parameter values from Ratcliff et al., 2011 for each age group (young, old, and very old) and IQ group from both the item and associative recognition tasks. The *dotted lines* plot the averages across the three age groups.

IQ on drift rates in the associative task, as opposed to an effect of similar size as for younger adults (a fixed decline) or a larger effect (a "use it or lose it" pattern of decline, with low IQ adults showing greater decline with age than higher IQ adults). In contrast, for item recognition, IQ has a similar effect on drift rate across all three age groups. In the second and third rows, we see that boundary separation and the nondecision component vary with age but are unaffected by IQ. This pattern of results helps validate the model. When freely estimated, components of processing that we would not expect to be related to IQ are not, and components that we would expect to be related are appropriately correlated with IQ.

By applying the diffusion model, RT and accuracy results can be explained with a single cohesive framework. The model allows us to describe behavioral effects in terms of the processes underlying the decision, which in turn allows us to make more direct comparisons across subjects who may have different speed–accuracy preferences.

2.13.6 Electroencephalogram and Memory

The diffusion model can help distinguish between neural components that are related to the evidence feeding into the decision process and components that are stimulus-related but do not affect the decision. In a perceptual task, Philiastides et al. (2006) found two EEG components that tracked category information about the stimulus. However, when the components were used to sort the data into halves based on the value of each component and the diffusion model was fit to each half, drift rates differed across halves for only the latter component (Ratcliff et al., 2009). This type of analysis helps determine whether information tracked by neural components is being used in the decision-making process.

In the memory literature, there is debate over whether recognition decisions are based on a single continuous memory dimension or a continuous dimension plus a discrete-state process. According to the single-process account, recognition decisions are made based on a single continuous measure of memory strength (although this measure may have more than one source of information; Cohen et al., 2008; Dennis and Humphreys, 2001; Dunn, 2004; Gillund and Shiffrin, 1984; Hintzman, 1984; Shiffrin and Steyvers, 1997; Starns and Ratcliff, 2008; Wixted, 2007). According to Yonelinas' dual-process account, recognition decisions are made based on two independent decision processes—a continuous familiarity process and an all-or-nothing recollection process (Yonelinas, 1994, 1997). Proponents of this dual-process account of recognition memory have argued that a frontocentral event-related potential (ERP) component around 400 ms tracks the familiarity process and a parietal ERP component around 600 ms tracks the recollection process (Eichenbaum et al., 2007; Rugg and Yonelinas, 2003).

Ratcliff et al. (2016) used the diffusion model to determine which EEG components track evidence that is used by the decision process in an item recognition task. In this study, single-trial analysis of EEG signals recorded during the test portions of two item recognition experiments were used to find a component that distinguished between "old" and "new" items. This was done using a multivariate pattern analysis that weights the signals from the array of electrodes to produce a single value that represents how "old" or "new" the EEG signal was for that time window (for more details on this analysis, see Philiastides et al., 2006; Ratcliff et al., 2009). It is important to note that this analysis is based on the EEG data and stimulus category (i.e., "old" or "new") but not on behavioral data. These values were then used to sort the data from each condition into two halves—an "older" half and a "newer" half—and the diffusion model was fit to the data, allowing drift rates to vary across the halves. If the drift rates differ across the halves of the data, this indicates that the EEG signal represents evidence that is used in the decision process. If the drift rates do not differ across the halves of the data, this indicates that the EEG signal does not represent information that is used in the decision process.

Ratcliff et al. (2016) found that the EEG signal distinguished between "old" and "new" items most strongly around 600 ms after stimulus presentation in both experiments. The regressor values from each trial can be used to generate an ROC curve (as shown in **Fig. 6**A and B) and the area under this curve (Az) serves as a measure of the "old"/"new" discriminability. **Fig. 6**C plots the time course of this discriminability measure for each of the two experiments along with a 95% confidence limit (calculated via random permutation). The regressor values from 600 ms were used to sort the data from each condition into "older" and "newer" halves and the diffusion model was fit to the sorted data. Across the two halves, only the drift rate parameters were significantly different and the difference was consistent with the regressor values (e.g., "old" stimuli with "older" regressor values had larger drift rates than "old" stimuli with "newer" regressor values). For comparison, when the data were sorted using the regressor values from an earlier time point (375 ms after stimulus presentation) and fit with the diffusion model, there were no significant differences in drift rates across the two halves. **Fig. 6**C plots the drift rates from the "older" and "newer" halves of the data against each other for each experiment. In the first column, when the data were sorted according to the regressor value at 600 ms after stimulus presentation, the drift rates for the "older" half of the data are larger than the drift rates for the "newer" half of the data. In the second column, when the data were sorted according to the regressor value at 375 ms, the drift rates for the two halves were approximately equal. Although the regressor value at 375 ms significantly discriminated between old and new items (based on the random permutation test) for only one of the experiments, there was a significant difference between the regressor values for old and new items at this time point for both experiments and ERPs at that time point show separation between old and new items. However, when fitting data based on the signal at that time point, there is no difference in drift rates between the two halves. This indicates that the signal at that time point, while it may covary with the stimulus categories, does not represent the information that is used by the decision-making process.

In addition to these split-half analyses, Ratcliff et al. (2016) used the regressor value for each test word as a coefficient for drift rate where drift rate was defined as: $v = v_0 + R^*v_1$, where v_0 was a constant that was different for each condition of the experiment (e.g., words studied once or words studied twice) and R^*v_1 was the contribution of the regressor value (R). Using this expression for v, the diffusion model was fit to the choice and RT for each test word using the maximum likelihood method. The best-fitting value of the v_1 coefficient at 600 ms after a test word was displayed was different from zero; that is, the fit was improved by assuming that drift rate was a function of the regressor compared with the model that assumed a constant drift rate. In other words, the EEG signals indexed information that was used in the decision process.

In prior studies, an observed frontal signal around 400 ms has been interpreted as a familiarity component while a latter parietal signal around 600 ms has been interpreted as a recollection component (Eichenbaum et al., 2007; Rugg and Yonelinas, 2003). According to Yonelinas' dual-process account of recognition, both of these components are used in recognition decisions. In these analyses by Ratcliff et al., however, while there are neural signals that differ for old and new items at both of these time points, only the signal at the latter time point appears to be related to the information used in the decision-making process.

2.13.7 Working Memory

Donkin et al. (2013) used decision-making models and RT distributions to distinguish between competing theories of visual working memory. In their experiments, a small number of colored squares were presented briefly and then masked, and then a single test square was presented in the same location as one of the earlier squares and subjects had to judge whether the

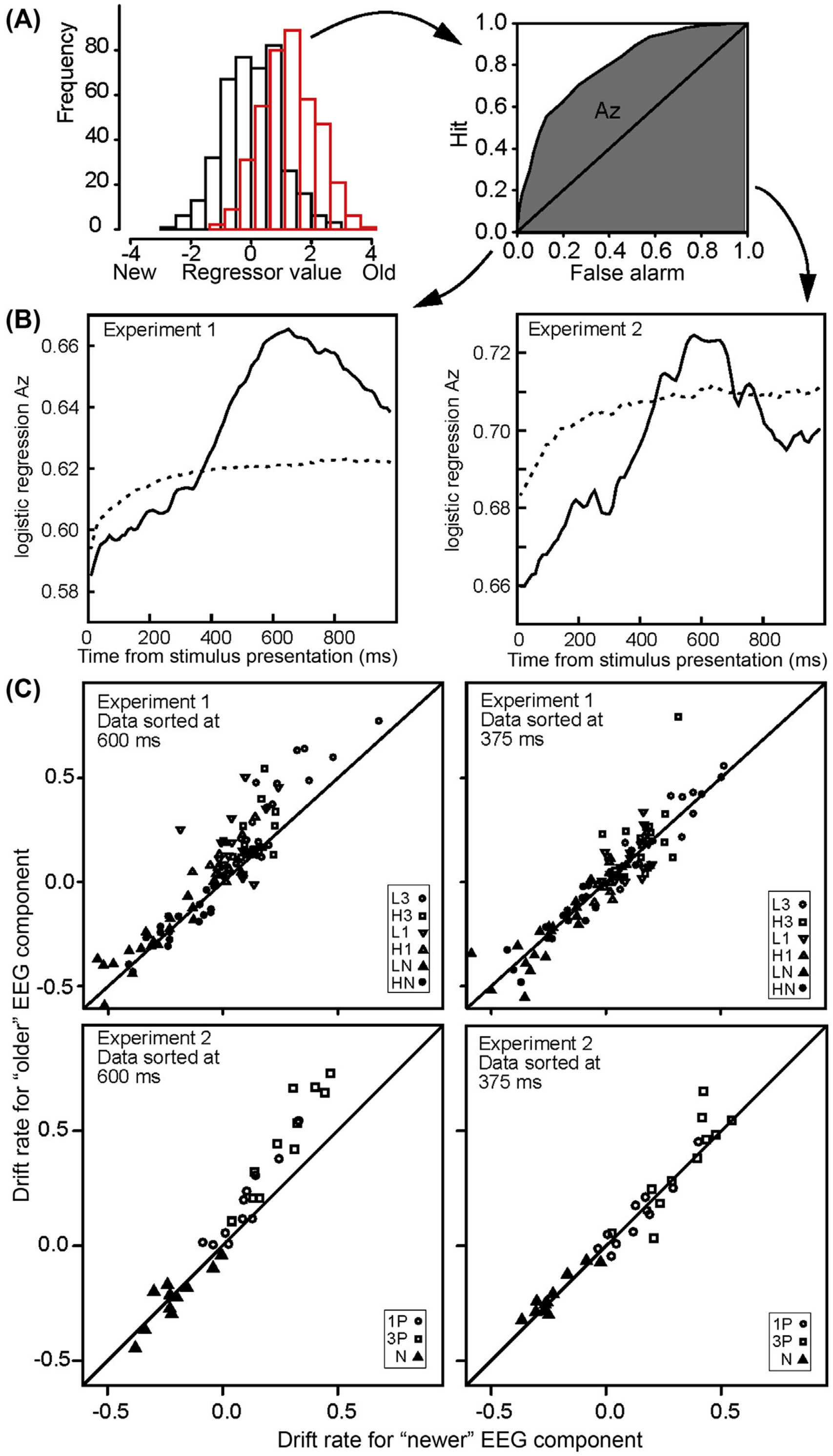

Figure 6 Panel (A) plots the regressor values for old and new test items, which were used to generate the area under the ROC curve (Az). Panel (B) plots the Az value for each time point for each of the two experiments (*solid line*) along with the upper 95% confidence limit (*dashed line*). The regressor values significantly discriminate old and new items when the *solid line* is above the *dashed line*. Panel (C) plots the drift rates for each of the individual subjects and conditions for each of the two experiments from data sorted according to the regressor value at 600 ms (first column) and from data sorted according to the regressor value at 375 ms (second column). The points in each plot are labeled according to the conditions in each experiment: *1*, one time presented; *3*, three times presented; *H*, high-frequency words; *L*, low-frequency words; *N*, new.

test square was of the same color or a different color than the square previously presented in that location. According to a discrete slots account, visual working memory contains some number of slots, which are used to store information about items. Most importantly, these slots are assumed to be all-or-none. That is, if a particular item was stored in one of the available slots, then performance for that item should be perfect (i.e., all of the information about it was retained); if a particular item was not stored in one of the available slots, then performance for that item should be at chance (i.e., none of the information about it was retained). In contrast, according to a continuous resource account of visual working memory, resources can be allotted continuously over items. That is, if there are only a small number of items to be remembered, then more information about each item can be retained, but if a larger number of items must be remembered, then less information about each item can be retained. There is an ongoing debate in the working memory literature between these two accounts, but much of the research examining working memory has focused on accuracy and ignored the RTs associated with these tasks (Alvarez and Cavanagh, 2004; Awh et al., 2007; Barton et al., 2009; Bays et al., 2009, 2011; Bays and Husain, 2008; Cowan, 2001; Cowan and Rouder, 2009; Luck and Vogel, 1997; Pashler, 1988; Rouder et al., 2008; van den Berg et al., 2012; Vogel et al., 2001; Wilken and Ma, 2004; Zhang and Luck, 2008).

Donkin et al. (2013) formalized these two accounts of working memory and used a linear ballistic accumulator (LBA) model (Brown and Heathcote, 2008) to produce accuracy and RT predictions. LBA is a simpler model of decision making and RTs, but its predictions generally match those of the diffusion model and similar parameters across the two models tend to show the same patterns of results (Donkin et al., 2011). When formalized in this way, the two accounts of working memory made similar predictions about accuracy but qualitatively different predictions about RT distributions. The data from their experiments were more consistent with the discrete slots account of working memory, and that model provided a better fit for the majority of their subjects. Using a model of decision making allowed the researchers to generate predictions about RT distributions from the models as well as predictions about accuracy, which in turn made the two accounts of working memory distinguishable.

2.13.8 Strength-Based Mirror Effect

Criss (2010) used the diffusion model to distinguish between two competing explanations for the strength-based mirror effect. The strength-based mirror effect is the finding that hit rates increase and false-alarm rates decrease as studied items are strengthened (e.g., as encoding time increases; Ratcliff et al., 1990; Stretch and Wixted, 1998). There are two popular explanations for this effect. According to the criterion shift account, subjects adopt a more conservative decision criterion when their accuracy is high. According to the differentiation account (which is based on the global matching models; see Shiffrin et al., 1990; Shiffrin and Steyvers, 1997), better encoding conditions will produce more accurate memories such that old and new items will have more dissimilar representations in memory. In signal detection terms, the criterion shift account involves a change in the decision criterion and the differentiation account involves changes in the distributions of memory strength for old and new items.

Criss (2010) collected accuracy and RT data from a recognition memory experiment and fit the data with the diffusion model. She found changes in mean drift rates consistent with a differentiation account (i.e., there was a larger difference between the drift rates for targets and lures for the strong encoding condition than for the weak condition). However, she also found changes in the drift criterion that could be consistent with the criterion shift account.

Additional work by Starns et al. (2010) and Starns et al. (2012b) demonstrated that subjects do adjust decision processes based on the strength of the items in the test list in support of the criterion shift account. According to the differentiation account, the strength-based mirror effect arises because of changes to the memory representations that occur during encoding. Starns et al. (2010, 2012b) held encoding conditions constant, but varied the strength of the items during retrieval (e.g., tested only strong or only weak items from the study period) and found reduced false-alarm rates for stronger test lists but not for stronger study lists. That is, they observed a strength-based mirror effect that was based purely on the strength of the test lists, which is inconsistent with the differentiation claim that the effect occurs because of processes that occur during encoding.

2.13.9 Conclusions

We have described several examples in which application of a model of the decision-making process provided a different account of the data relative to standard methods. In all of these studies, the addition of a formal model of the decision-making process that makes predictions about RT distributions either helped distinguish between theories that were difficult or impossible to distinguish solely from accuracy data, or provided additional support for ideas that were previously based only on accuracy data.

The diffusion model is a model of decision making, not of memory. However, to the extent that memory is studied by asking subjects to make decisions based on information from memory, modeling the decision process makes it possible to distinguish effects of memory from effects of decision making. Attempting to draw conclusions about underlying processes without a model of the decision-making process can lead to incorrect conclusions. Without a model of decision making that allows for individual differences in how people set decision thresholds, differences in RTs may be interpreted as a cognitive deficit rather than cautious behavior. Without a model of confidence judgments and RTs, certain patterns of responses may be interpreted as memory effects rather than differences in decision making. Without a model of how evidence is used to make a decision, patterns of neural activity may incorrectly appear to contain information relevant to a decision. In short, mapping some memory representation into behavior in a memory task requires some kind of decision-making process.

The diffusion model provides one possible account of the process linking underlying cognitive representations and processes to observable choice behavior.

See also: 2.12 Signal Detection Theories of Recognition Memory.

References

Alvarez, G.A., Cavanagh, P., 2004. The capacity of visual short-term memory is set both by visual information load and by number of objects. Psychol. Sci. 15, 106–111.

Atkinson, R.C., Juola, J.F., 1973. Factors influencing speed and accuracy of word recognition. In: Kornblum, S. (Ed.), Attention & Performance, vol. IV. Academic Press, New York.

Awh, E., Barton, B., Vogel, E.K., 2007. Visual working memory represents a fixed number of items regardless of complexity. Psychol. Sci. 18, 622–628.

Balota, D.A., Dolan, P.O., Duchek, J.M., 2000. Memory changes in healthy older adults. In: Tulving, E., Craik, F.I.M. (Eds.), The Oxford Handbook of Memory. Oxford University Press, New York, NY, USA, pp. 395–409.

Banks, W.P., 1970. Signal detection theory and human memory. Psychol. Bull. 74 (2), 81–99. http://dx.doi.org/10.1037/h0029531.

Barton, B., Ester, E.F., Awh, E., 2009. Discrete resource allocation in visual working memory. J. Exp. Psychol. Hum. Percept. Perform. 35, 1359–1367.

Bastin, C., Van der Linden, M., 2005. The effects of aging on the recognition of different types of associations. Exp. Aging Res. 32 (1), 61–77.

Bays, P.M., Husain, M., 2008. Dynamic shifts of limited working memory resources in human vision. Science 321, 851–854.

Bays, P.M., Catalao, R.F.G., Husain, M., 2009. The precision of visual working memory is set by allocation of a shared resource. J. Vis. 9, 1–11.

Bays, P.M., Gorgoraptis, N., Wee, N., Marshall, L., Husain, M., 2011. Temporal dynamics of encoding, storage, and reallocation of visual working memory. J. Vis. 11, 1–15.

van den Berg, R., Shin, H., Chou, W.C., George, R., Ma, W.J., 2012. Variability in encoding precision accounts for visual short-term memory limitations. Proc. Natl. Acad. Sci. U.S.A. 109, 8780–8785.

Bernbach, H.A., 1967. Decision processes in memory. Psychol. Rev. 74 (6), 462–480. http://dx.doi.org/10.1037/h0025132.

Bower, G.H., 1967. A multicomponent theory of memory trace. In: Spence, K.W., Spence, J.T. (Eds.), The Psychology of Learning and Motivation: Advances in Research and Theory, vol. 1. Academic Press, New York.

Bowles, N.L., Poon, L.W., 1982. An analysis of the effect of aging on recognition memory. J. Gerontol. 37 (2), 212–219. http://dx.doi.org/10.1093/geronj/37.2.212.

Brown, S.D., Heathcote, A., 2008. The simplest complete model of choice response time: linear ballistic accumulation. Cogn. Psychol. 57 (3), 153–178.

Buchler, N.E., Reder, L.M., 2007. Modeling age-related memory deficits: a two-parameter solution. Psychol. Aging 22 (1), 104.

Cerella, J., 1985. Information processing rates in the elderly. Psychol. Bull. 98, 67–83.

Cohen, A.L., Rotello, C.M., Macmillan, N.A., 2008. Evaluating models of remember-know judgments: complexity, mimicry, and discriminability. Psychon. Bull. Rev. 15, 906–926.

Cowan, N., 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav. Brain Sci. 24, 87–114.

Cowan, N., Rouder, J.N., 2009. Comment on "dynamic shifts of limited working memory resources in human vision". Science 323, 877.

Craik, F.I.M., 2008. Memory changes in normal and pathological aging. Can. J. Psychiatry 53 (6), 343–345.

Craik, F.I.M., 1994. Memory changes in normal aging. Curr. Dir. Psychol. Sci. 3, 155–158.

Craik, F.I.M., Broadbent, D., 1983. On the transfer of information from temporary to permanent memory. Philos. Trans. R. Soc. Lond. B Biol. Sci. 302 (1110), 341–359.

Craik, F.I.M., Jennings, J.M., 1992. Human memory. In: Craik, F.I.M., Salthouse, T.A. (Eds.), The Handbook of Aging and Cognition. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, England, pp. 51–110.

Craik, F.I.M., McDowd, J.M., 1987. Age differences in recall and recognition. J. Exp. Psychol. Learn. Mem. Cogn. 13 (3), 474.

Criss, A.H., 2010. Differentiation and response bias in episodic memory: evidence from reaction time distributions. J. Exp. Psychol. Learn. Mem. Cogn. 36 (2), 484–499.

Deary, I.J., 2000. Looking Down on Human Intelligence: From Psychometrics to the Brain. Oxford University Press, New York, NY, USA.

DeCarlo, L.T., 2002. Signal detection theory with finite mixture distributions: theoretical developments with applications to recognition memory. Psychol. Rev. 109 (4), 710–721. http://dx.doi.org/10.1037/0033-295X.109.4.710.

DeCarlo, L.T., 2003. An application of signal detection theory with finite mixture distributions to source discrimination. J. Exp. Psychol. Learn. Mem. Cogn. 29 (5), 767–778. http://dx.doi.org/10.1037/0278-7393.29.5.767.

Dennis, S., Humphreys, M.S., 2001. A context noise model of episodic word recognition. Psychol. Rev. 108, 452–477.

Donaldson, W., Murdock, B.B., 1968. Criterion change in continuous recognition memory. J. Exp. Psychol. 76 (3), 325–330. http://dx.doi.org/10.1037/h0025510.

Donkin, C., Brown, S.D., Heathcote, A., Wagenmakers, E.J., 2011. Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? Psychon. Bull. Rev. 18, 61–69.

Donkin, C., Nosofsky, R.M., Gold, J.M., Shiffrin, R.M., 2013. Discrete-slots models of visual working-memory response times. Psychol. Rev. 120 (4), 873–902.

Dunn, J.C., 2004. Remember-know: a matter of confidence. Psychol. Rev. 111, 524–542.

Egan, J.P., 1958. Recognition memory and the operating characteristic. USAF Oper. Appl. Lab. Tech. Note 58-51, 32.

Eichenbaum, H., Yonelinas, A.P., Ranganath, C., 2007. The medial temporal lobe and recognition memory. Annu. Rev. Neurosci. 30, 123–152.

Erber, J.T., 1974. Age differences in recognition memory. J. Gerontol. 29 (2), 177–181. http://dx.doi.org/10.1093/geronj/29.2.177.

Gillund, G., Shiffrin, R.M., 1984. A retrieval model for both recognition and recall. Psychol. Rev. 91, 1–67.

Glanzer, M., Hilford, A., Kim, K., 2004. Six regularities of source recognition. J. Exp. Psychol. Learn. Mem. Cogn. 30, 1176–1195.

Gordon, S.K., Clark, W.C., 1974. Application of signal detection theory to prose recall and recognition in elderly and young adults. J. Gerontol. 29 (1), 64–72. http://dx.doi.org/10.1093/geronj/29.1.64.

Grasha, A.F., 1970. Detection theory and memory processes: are they compatible? Percept. Mot. Skills 30 (1), 123–135. http://dx.doi.org/10.2466/pms.1970.30.1.123.

Healy, M.R., Light, L.L., Chung, C., 2005. Dual-process models of associative recognition in young and older adults: evidence from receiver operating characteristics. J. Exp. Psychol. Learn. Mem. Cogn. 31 (4), 768–788. http://dx.doi.org/10.1037/0278-7393.31.4.768.

Hilford, A., Glanzer, M., Kim, K., DeCarlo, L.T., 2002. Regularities of source recognition: ROC analysis. J. Exp. Psychol. Gen. 131 (4), 494–510. http://dx.doi.org/10.1037/0096-3445.131.4.494.

Hintzman, D.L., 1984. MINERVA 2: a simulation model of human memory. Behav. Res. Methods Instrum. Comput. 16, 96–101.

Jones, M., Dzhafarov, E.N., 2014. Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. Psychol. Rev. 121, 1–32.

Juola, J.F., Fischler, I., Wood, C.T., Atkinson, R.C., 1971. Recognition time for information stored in long-term memory. Percept. Psychophys. 10 (1), 8–14.

Kausler, D.H., 1994. Learning and Memory in Normal Aging. Academic Press, San Diego, CA, USA.

Kelley, R., Wixted, J.T., 2001. On the nature of associative information in recognition memory. J. Exp. Psychol. Learn. Mem. Cogn. 27 (3), 701–722. http://dx.doi.org/10.1037/0278-7393.27.3.701.

Kintsch, W., 1967. Memory and decision aspects of recognition learning. Psychol. Rev. 74 (6), 496–504. http://dx.doi.org/10.1037/h0025127.

Kintsch, W., Carlson, W.J., 1967. Changes in the memory operating characteristic during recognition learning. J. Verbal Learn. Verbal Behav. 6 (6), 891–896. http://dx.doi.org/10.1016/S0022-5371(67)80155-5.
Lockhart, R.S., Murdock, B.B., 1970. Memory and the theory of signal detection. Psychol. Bull. 74 (2), 100–109. http://dx.doi.org/10.1037/h0029536.
Luck, S.J., Vogel, E.K., 1997. The capacity of visual working memory for features and conjunctions. Nature 390, 279–281.
Malmberg, K.J., Xu, J., 2006. The influence of averaging and noisy decision strategies on the recognition memory ROC. Psychon. Bull. Rev. 13 (1), 99–105.
McElree, B., Dolan, P.O., Jacoby, L.L., 1999. Isolating the contributions of familiarity and source information to item recognition: a time course analysis. J. Exp. Psychol. Learn. Mem. Cogn. 25, 563–582.
Naveh-Benjamin, M., 2000. Adult age differences in memory performance: tests of an associative deficit hypothesis. J. Exp. Psychol. Learn. Mem. Cogn. 26 (5), 1170–1187. http://dx.doi.org/10.1037/0278-7393.26.5.1170.
Naveh-Benjamin, M., Guez, J., Shulman, S., 2004a. Older adults' associative deficit in episodic memory: assessing the role of decline in attentional resources. Psychon. Bull. Rev. 11 (6), 1067–1073.
Naveh-Benjamin, M., Guez, J., Kilb, A., Reedy, S., 2004b. The associative memory deficit of older adults: further support using face-name associations. Psychol. Aging 19 (3), 541.
Naveh-Benjamin, M., Hussain, Z., Guez, J., Bar-On, M., 2003. Adult age differences in episodic memory: further support for an associative-deficit hypothesis. J. Exp. Psychol. Learn. Mem. Cogn. 29 (5), 826.
Neath, I., 1998. Human Memory: An Introduction to Research, Data, and Theory. Thomson Brooks/Cole Publishing Co, Belmont, CA, USA.
Norman, D.A., Rumelhart, D.E., 1970. A system for perception and memory. In: Norman, D.A. (Ed.), Models of Human Memory. Academic Press, New York.
Norman, D.A., Wickelgren, W.A., 1969. Strength theory of decision rules and latency in retrieval from short-term memory. J. Math. Psychol. 6 (2), 192–208. http://dx.doi.org/10.1016/0022-2496(69)90002-9.
Old, S.R., Naveh-Benjamin, M., 2008. Differential effects of age on item and associative measures of memory: a meta-analysis. Psychol. Aging 23 (1), 104–118. http://dx.doi.org/10.1037/0882-7974.23.1.104.
Pashler, H., 1988. Familiarity and visual change detection. Percept. Psychophys. 44, 369–378.
Philiastides, M.G., Ratcliff, R., Sajda, P., 2006. Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. J. Neurosci. 26, 8965–8975.
Qin, J., Raye, C.L., Johnson, M.K., Mitchell, K.J., 2001. Source ROCs are (typically) curvilinear: comment on Yonelinas (1999). J. Exp. Psychol. Learn. Mem. Cogn. 27, 1110–1115.
Rabinowitz, J.C., 1984. Aging and recognition failure. J. Gerontol. 39 (1), 65–71. http://dx.doi.org/10.1093/geronj/39.1.65.
Ratcliff, R., 1978. A theory of memory retrieval. Psychol. Rev. 85 (2), 59–108. http://dx.doi.org/10.1037/0033-295X.85.2.59.
Ratcliff, R., 2002. A diffusion model account of reaction time and accuracy in a brightness discrimination task: fitting real data and failing to fit fake but plausible data. Psychon. Bull. Rev. 9, 278–291.
Ratcliff, R., 2013. Parameter variability and distributional assumptions in the diffusion model. Psychol. Rev. 120, 281–292.
Ratcliff, R., 2014. Measuring psychometric functions with the diffusion model. J. Exp. Psychol. Hum. Percept. Perform. 40, 870–888.
Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 20 (4), 873–922. http://dx.doi.org/10.1162/neco.2008.12-06-420.
Ratcliff, R., McKoon, G., 2015. Aging effects in item and associative recognition memory for pictures and words. Psychol. Aging 30, 669–674.
Ratcliff, R., Murdock Jr., B.B., 1976. Retrieval processes in recognition memory. Psychol. Rev. 83, 190–214.
Ratcliff, R., Starns, J.J., 2009. Modeling confidence and response time in recognition memory. Psychol. Rev. 116 (1), 59–83.
Ratcliff, R., Starns, J.J., 2013. Modeling reaction times, choices, and confidence judgments in decision making: recognition memory and motion discrimination. Psychol. Rev. 120, 697–719.
Ratcliff, R., Clark, S.E., Shiffrin, R.M., 1990. List strength effect I: data and discussion. J. Exp. Psychol. Learn. Mem. Cogn. 16, 163–178.
Ratcliff, R., McKoon, G., Tindall, M.H., 1994. Empirical generality of data from recognition memory receiver-operating characteristic functions and implications for the global memory models. J. Exp. Psychol. Learn. Mem. Cogn. 20, 763–785.
Ratcliff, R., Philiastides, M.G., Sajda, P., 2009. Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. Proc. Natl. Acad. Sci. U.S.A. 106, 6539–6544.
Ratcliff, R., Sederberg, P., Smith, T., Childers, R., 2016. A single trial analysis of EEG in recognition memory: tracking the neural correlates of memory strength. Neuropsychologia 93, 128–141.
Ratcliff, R., Sheu, C., Gronlund, S.D., 1992. Testing global memory models using ROC curves. Psychol. Rev. 99 (3), 518–535. http://dx.doi.org/10.1037/0033-295X.99.3.518.
Ratcliff, R., Smith, P.L., McKoon, G., 2015. Modeling regularities in response time and accuracy data with the diffusion model. Curr. Dir. Psychol. Sci. 24, 458–470.
Ratcliff, R., Thapar, A., McKoon, G., 2004. A diffusion model analysis of the effects of aging on recognition memory. J. Mem. Lang. 50, 408–424.
Ratcliff, R., Thapar, A., McKoon, G., 2010. Individual differences, aging, and IQ in two-choice tasks. Cogn. Psychol. 60, 127–157.
Ratcliff, R., Thapar, A., McKoon, G., 2011. Effects of aging and IQ on item and associative memory. J. Exp. Psychol. Gen. 140, 464–487.
Rotello, C.M., Macmillan, N.A., Reeder, J.A., 2004. Sum-difference theory of remembering and knowing: a two-dimensional signal-detection model. Psychol. Rev. 111 (3), 588–616. http://dx.doi.org/10.1037/0033-295X.111.3.588.
Rouder, J.N., Morey, R.D., Cowan, N., Zwilling, C.E., Morey, C.C., Pratte, M.S., 2008. An assessment of fixed-capacity models of visual working memory. Proc. Natl. Acad. Sci. U.S.A. 105, 5975–5979.
Rugg, M.D., Yonelinas, A.P., 2003. Human recognition memory: a cognitive neuroscience perspective. Trends Cogn. Neurosci. 7, 313–319.
Salthouse, T.A., 1996. The processing-speed theory of adult age differences in cognition. Psychol. Rev. 103, 403–428.
Schonfield, D., Robertson, B.A., 1966. Memory storage and aging. Can. J. Psychol. 20 (2), 228–236. http://dx.doi.org/10.1037/h0082941.
Shiffrin, R.M., Steyvers, M., 1997. A model for recognition memory: REM—retrieving effectively from memory. Psychon. Bull. Rev. 4, 145–166.
Shiffrin, R.M., Ratcliff, R., Clark, S., 1990. List strength effect II: theoretical mechanisms. J. Exp. Psychol. Learn. Mem. Cogn. 16, 179–195.
Slotnick, S.D., Dodson, C.S., 2005. Support for a continuous (single-process) model of recognition memory and source memory. Mem. Cogn. 33, 151–170.
Slotnick, S.D., Klein, S.A., Dodson, C.S., Shimamura, A.P., 2000. An analysis of signal detection and threshold models of source memory. J. Exp. Psychol. Learn. Mem. Cogn. 26, 1499–1517.
Smith, P.L., Ratcliff, R., McKoon, G., 2014. The diffusion model is not a deterministic growth model: comment on Jones and Dzhafarov (2013). Psychol. Rev. 121, 679–688.
Starns, J.J., Ratcliff, R., 2008. Two dimensions are not better than one: STREAK and the univariate signal detection model of remember/know performance. J. Mem. Lang. 59, 169–182.
Starns, J.J., Ratcliff, R., 2014. Validating the unequal-variance assumption in recognition memory using response time distributions instead of ROC functions: a diffusion model analysis. J. Mem. Lang. 70, 36–52.
Starns, J.J., Ratcliff, R., McKoon, G., 2012a. Evaluating the unequal-variance and dual-process explanations of zROC slopes with response time data and the diffusion model. Cogn. Psychol. 64 (1–2), 1–34. http://dx.doi.org/10.1016/j.cogpsych.2011.10.002.
Starns, J.J., Ratcliff, R., White, C.N., 2012b. Diffusion model drift rates can be influenced by decision processes: an analysis of the strength-based mirror effect. J. Exp. Psychol. Learn. Mem. Cogn. 38 (5), 1137–1151.
Starns, J.J., White, C.N., Ratcliff, R., 2010. A direct test of the differentiation mechanism: REM, BCDMEM, and the strength-based mirror effect in recognition memory. J. Mem. Lang. 63 (1), 18–34.

Stretch, V., Wixted, J.T., 1998. On the difference between strength-based and frequency-based mirror effects in recognition memory. J. Exp. Psychol. Learn. Mem. Cogn. 24 (6), 1379–1396.
Verhaeghen, P., Salthouse, T.A., 1997. Meta-analyses of age–cognition relations in adulthood: estimates of linear and nonlinear age effects and structural models. Psychol. Bull. 122 (3), 231–249. http://dx.doi.org/10.1037/0033-2909.122.3.231.
Vogel, E.K., Woodman, G.F., Luck, S.J., 2001. Storage of features, conjunctions, and objects in visual working memory. J. Exp. Psychol. Hum. Percept. Perform. 27, 92–114.
Voskuilen, C., Ratcliff, R., 2016. Modeling confidence and response time in associative recognition. J. Mem. Lang. 86, 60–96.
White, C.N., Ratcliff, R., Vasey, M.W., McKoon, G., 2010. Using diffusion models to understand clinical disorders. J. Math. Psychol. 54, 39–52.
White, C., Ratcliff, R., Vasey, M., McKoon, G., 2009. Dysphoria and memory for emotional material: a diffusion model analysis. Cogn. Emot. 23, 181–205.
Wickelgren, W.A., 1977. Speed-accuracy tradeoff and information processing dynamics. Acta Psychol. 41 (1), 67–85. http://dx.doi.org/10.1016/0001-6918(77)90012-9.
Wickelgren, W.A., Norman, D.A., 1966. Strength models and serial position in short-term recognition memory. J. Math. Psychol. 3, 316–347.
Wilken, P., Ma, W.J., 2004. A detection theory account of change detection. J. Vis. 4, 1120–1135.
Wixted, J.T., 2007. Dual-process theory and signal-detection theory of recognition memory. Psychol. Rev. 114, 152–176.
Yonelinas, A.P., 1994. Receiver-operating characteristics in recognition memory: evidence for a dual-process model. J. Exp. Psychol. Learn. Mem. Cogn. 20 (6), 1341–1354. http://dx.doi.org/10.1037/0278-7393.20.6.1341.
Yonelinas, A.P., 1997. Recognition memory ROCs for item and associative information: the contribution of recollection and familiarity. Mem. Cogn. 25 (6), 747–763.
Yonelinas, A.P., 1999. The contribution of recollection and familiarity to recognition and source-memory judgments: a formal dual-process model and an analysis of receiver operating characteristics. J. Exp. Psychol. Learn. Mem. Cogn. 25, 1415–1434.
Yonelinas, A.P., Parks, C.M., 2007. Receiver operating characteristics (ROCs) in recognition memory: a review. Psychol. Bull. 133 (5), 800–832. http://dx.doi.org/10.1037/0033-2909.133.5.800.
Zhang, W., Luck, S.J., 2008. Discrete fixed-resolution representations in visual working memory. Nature 453, 233–235.