aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Current Directions in Psychological Science
2015, Vol. 24(6) 458–470

DOI: 10.1177/0963721415596228
cdps.sagepub.com
SAGE

Modeling Regularities in Response Time and Accuracy Data With the Diffusion Model

Roger Ratcliff[1], Philip L. Smith[2], and Gail McKoon[1]
[1]Department of Psychology, The Ohio State University, and [2]Melbourne School of Psychological Sciences, The University of Melbourne

Abstract
Diffusion models for simple two-choice decision making have achieved prominence in psychology and neuroscience. The standard model views decision making as a process in which noisy evidence is accumulated until one of the two response criteria is reached, at which point the associated response is made. The criteria represent the amount of evidence needed to make a decision, and they reflect the decision maker's response biases and speed–accuracy trade-off settings. In this article, we review the regularities in experimental data that a model must explain. These include the relation between accuracy and mean response times, the shapes of response-time distributions for correct and error responses and how they change with experimental variables, and individual differences in response time and accuracy. These relations are sometimes overlooked by researchers, but, taken together, they provide extremely strong tests of models.

Keywords
diffusion model, response time and accuracy, response-time distributions, individual differences

Decision making plays a central role in the human ability to translate perception into action. Deciding to stop or proceed when a traffic light changes or deciding to treat a person in the street as a stranger or an acquaintance are the kinds of two-choice cognitive tasks we perform many times every day. These decisions do not require prolonged deliberation or reflection of the kind associated with decision making in more complex domains. Instead, they involve a rapid matching of a perceptual representation of the world to our stored knowledge about the world in memory. This allows us to identify the things in our immediate surroundings and to determine how we should respond to them.

Many everyday decisions involve a choice between two alternatives, as the preceding examples illustrate. In the laboratory, much of what we have learned about human cognition likewise has come from tasks in which we ask people to make two-choice decisions about stimulus materials. Laboratory studies typically measure either response time (RT) or the probabilities with which people choose alternatives (*choice probabilities*). RTs measure the speed with which decisions are made; choice probabilities measure how similar the decision alternatives are to each other and how easily they can be distinguished. In many tasks, choice probability is a measure of response accuracy, and it is a function of task difficulty. Typically, researchers are interested in how and why choice probabilities or RTs change across experimental conditions as aspects of the task, such as stimulus discriminability or instructions about speed and accuracy, are varied.

This article provides an overview of regularities in two-choice data that must be explained and describes how one particular model, *Ratcliff's diffusion model*, does so (Ratcliff, 1978; Ratcliff & McKoon, 2008). The diffusion model is a member of a larger class of sequential sampling models that were developed in mathematical psychology and have become influential in neuroscience (e.g., Forstmann, Ratcliff, & Wagenmakers, 2015; Gold &

Corresponding Author:
Roger Ratcliff, Department of Psychology, The Ohio State University, Columbus, OH 43210
E-mail ratcliff.22@osu.edu

Shadlen, 2007; Hanes & Schall, 1996; Shadlen & Kiani, 2013; Smith & Ratcliff, 2004). More than any other model of speeded decision making, the diffusion model has been successfully applied to a wide variety of cognitive tasks. These include tasks that investigate the basic cognitive processes involved in perception, memory, and language and tasks that investigate cognitive deficits for special subject populations, such as children with ADHD, the elderly, sleep-deprived individuals, and dyslexics (e.g., Mulder et al., 2010; Ratcliff, Thapar, & McKoon, 2001, 2003, 2004, 2010; Ratcliff & Van Dongen, 2009; Zeguers et al., 2011). The model has been used in these various domains as a theoretical tool to help understand basic cognitive processing and as a psychometric tool to characterize cognitive deficits. In applications, the model allows researchers to distinguish between performance differences due to the quality of the information entering the decision process from a stimulus and differences due to the amount of information that must be accumulated before a response is made.

One of the aims of decision models is to provide an account of the mechanisms that underlie RT and accuracy. Although many studies use only one of these variables, they covary in regular and predictable ways, as numerous experiments have shown and as we detail later. Responses in easy conditions of a task are both faster and more accurate than responses in difficult conditions, and the functions that relate RTs and accuracy have a common form across conditions within experiments, across different experiments, and across subject populations. Error RTs are sometimes shorter than correct RTs and sometimes longer; which of these patterns is found depends on the experimental conditions in predictable ways. Distributions of RTs have the same characteristic shape across the conditions of an experiment, experiments, and subject populations, and in all cases, the shapes change in the same ways as the difficulty of a task is varied. Below, we illustrate these regularities with two subject populations, young and older adults, and three experimental tasks: two low-level, perceptual tasks (brightness and distance discriminations) and one higher-level task (recognition memory). We chose these three tasks to emphasize that the same relationships between RT and accuracy and the same shapes and behaviors of RT distributions are found in both low-level and higher-level tasks. We also describe how the diffusion model predicts these relationships and, equally importantly, how it cannot predict others.

The Diffusion Model

Figure 1a summarizes the main properties of the diffusion model. Decisions are made by accumulating noisy evidence to one of two response criteria that represent the amounts of evidence needed to make each response. The time taken to accumulate the evidence determines the decision-time component of RT, and the criterion that is reached first determines the alternative that is chosen. The model assumes that evidence is continuously distributed and accumulates in continuous time as a single total that can take on both positive and negative values. Positive evidence is evidence for one response, and negative evidence is evidence for the other. Mathematically, the process of evidence accumulation is modeled as a Brownian motion, or Wiener, diffusion process, and the decision criteria are modeled as absorbing boundaries (i.e., at which the process terminates or is "absorbed"). Brownian motion is the cumulative sum of a large number of identically distributed independent random variables. It is obtained by allowing the magnitudes of the variables and the times between changes to the cumulative sum to become small. In the model, the values of the random variables represent the instantaneous values of the evidence at a given time. The absorbing barriers represent the amounts of evidence needed to make each response: Evidence accumulation ceases when the process reaches an absorbing boundary. The trial-to-trial trajectories of the accumulating evidence have the highly irregular appearance shown in Figure 1.

To make a decision, an individual sets the two criteria, or decision boundaries, one above the starting point (z) and one below it (at 0 and a, respectively). The average rate at which evidence accumulates is known as *drift rate* (v in Fig. 1a). Its absolute magnitude depends on the discriminability of the alternatives: Higher discriminability leads to larger drift rates, which in turn lead to more rapid evidence accumulation. The sign and magnitude of drift rate are usually assumed to remain constant during a trial, but the noise in the accumulation process leads to variability in RTs and accuracy. This noise is illustrated in Figure 1a and is controlled by a parameter of the model.

In typical applications, responses at the upper criterion with positive drift rate are correct responses, and responses at the lower criterion are errors. The distance between the two criteria reflects an individual's speed–accuracy trade-off settings: The criteria are moved closer together to increase speed and moved farther apart to increase accuracy. The location of the starting point relative to the criteria determines bias toward one or the other response.

Figure 1b shows the components of processing in the model that together comprise RT. In addition to the decision time (d), there is the time required to encode the stimulus and form a representation of the attributes that are the basis of the discrimination (e) and the time required to output a motor response (r). The durations of e and r are added to form a nondecision time with duration T_{er}, which is estimated as a single parameter when the model is fit to data.

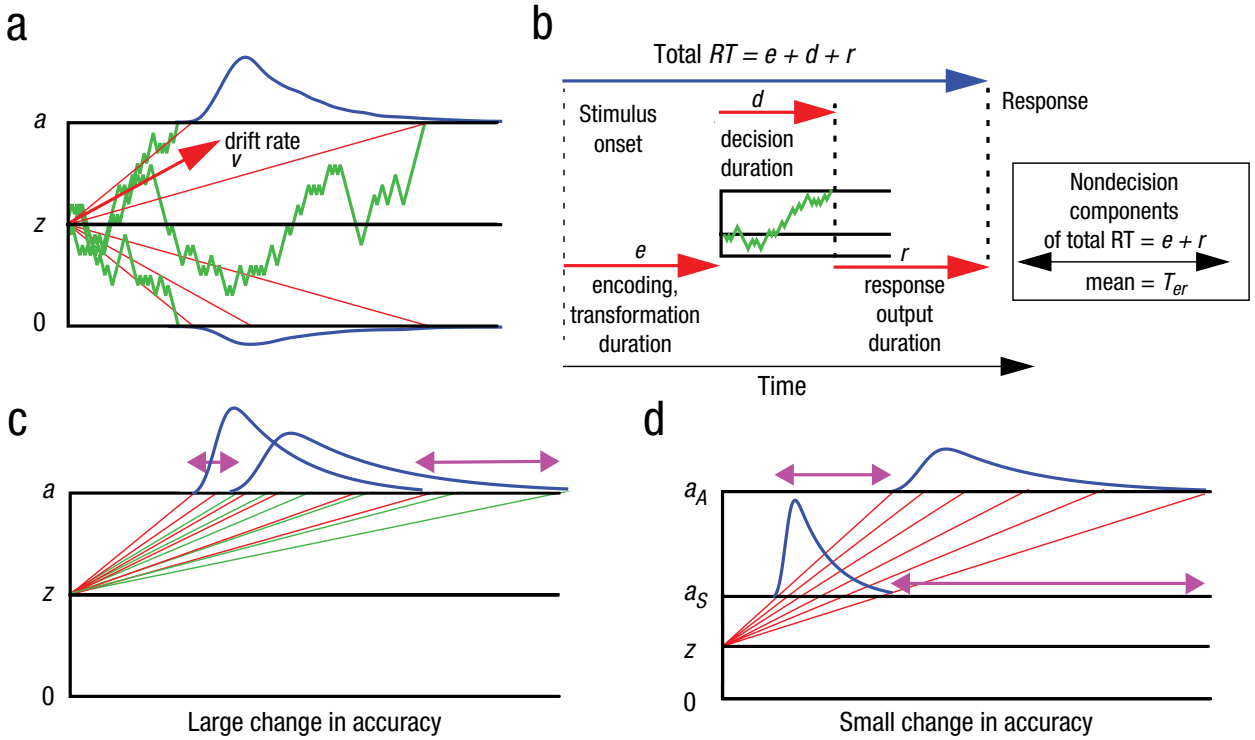


Fig. 1. An illustration of the diffusion model and how predictions change with drift rate and boundary separation. Panel a shows a diffusion model with sample paths with decision boundaries 0 and *a*, starting point at *z*, and mean drift rate *v*. The red lines are illustrative of average paths for different decision times showing how equal differences in the rate of approach (i.e., vertical differences) lead to a skewed response-time (RT) distribution (note that real average paths are not straight lines; see Figure 11A, Ratcliff, Cherian, & Segraves, 2003). Panel b shows the components of the decision, including the duration of the decision process (*d*), the duration of stimulus encoding and transformation of the stimulus to extract decision-related information to provide drift rate (*e*), and the duration of response output (*r*). The sum of *e* and *r* is T_{er}, which is the time for all of the nondecision components of the response. Panel c shows the effect of a change in drift rate on RT distributions; the red average paths (higher drift rate) lead to a narrower spread in the decision-time distribution versus the green average paths (lower drift rate), which lead to a wider spread in the decision-time distribution; the magenta arrows show the differences in the fastest and slowest processes. Panel d shows the effect of changing the decision boundary (*aS* to *aA*) on RT distributions that might occur, for example, by changing from speed to accuracy instructions.

The noisiness of the evidence-accumulation process allows the model to capture the inherent variability of decision making: People do not necessarily make the same response from one presentation of a stimulus to the next, and the time taken to respond likewise varies from one presentation to another. Choice probabilities and distributions of RTs for correct and error responses are the primary expression of this variability. Figures 1c and 1d show, respectively, how the model predicts changes in RT distributions as stimulus discriminability changes and as decision criteria change. Figure 1c shows that changing discriminability, represented by a change in drift rate, leads to a small shift in the leading edge of the RT distribution and a large increase in its spread. Figure 1d shows that increasing the criteria for accuracy (i.e., shifting from prioritizing speed, *aS*, to prioritizing accuracy, *aA*) leads to a larger shift in the leading edge of the distribution together with an increase in spread.

With criteria equidistant from the starting point and no variability across trials in any model parameters, the model predicts identical distributions for correct and error RTs. With variability in drift rate across trials, errors are slow because higher drift rates lead to faster responses with a lower probability of error and lower drift rates lead to slower responses with a higher probability of error. A mixture of higher and lower drift rates therefore leads to errors that are slower than correct responses (see Ratcliff & McKoon, 2008, Fig. 4). For similar reasons, across-trial variability in starting point produces errors faster than correct responses. When a process starts near the correct criterion, there will be few errors, and they will be slow because the process has to travel a long distance to reach the error criterion. When a process starts near the error criterion, there will be more errors, and they will be fast because the process has to travel only a short distance to reach the error criterion. A

mixture of higher and lower starting points leads to errors that are faster than correct responses.

Occasionally, a crossover pattern is found in which there are slow errors when instructions to subjects stress accurate responding and fast errors when instructions stress fast responding. This pattern is found in results from the brightness-discrimination task reported below and in the lexical decision data of Wagenmakers, Ratcliff, Gomez, and McKoon (2008). The crossover pattern is predicted by the diffusion model through a combination of drift rate and starting-point variability. When the separation between the criteria is small, performance is dominated by variability in starting point, whereas when the separation is large, performance is dominated by variability in drift rate. This leads to fast and slow errors, respectively, with only criterion separation varying between conditions.

The components of processing in the model are described by a number of free parameters that must be estimated from data. In addition to the drift-rate, criterion-separation, and nondecision-time parameters, there are parameters controlling the across-trial variability in each of the components. This might appear to endow the model with a large amount of flexibility, but it is in fact highly constrained. Ratcliff (2002) used simulations to show that the model was unable to fit sets of artificial but plausible RT distributions that differed from those found experimentally in various ways. The model also provides a high degree of data reduction and so is very parsimonious. A set of distributional data with 100 to 150 degrees of freedom can typically be fit using the model with eight to ten free parameters (depending on the experimental design). This represents a data-reduction factor of around 12:1 or 15:1.

There are three packages available for fitting the full diffusion model to data (fast-dm, Voss & Voss, 2007; Diffusion Model Analysis Toolbox [DMAT], Vandekerckhove & Tuerlinckx, 2008; Hierarchical Bayesian estimation of the Drift-Diffusion Model [HDDM], Wiecki, Sofer, & Frank, 2013; see also Wagenmakers, Van Der Maas, & Grasman, 2007). Ratcliff and Childers (2015) conducted simulations that compared the different packages in the quality of the fits provided, the number of observations needed to fit the model, and how well individual differences were recovered.

Hidden Regularities in Data From Two-Choice Decision Experiments

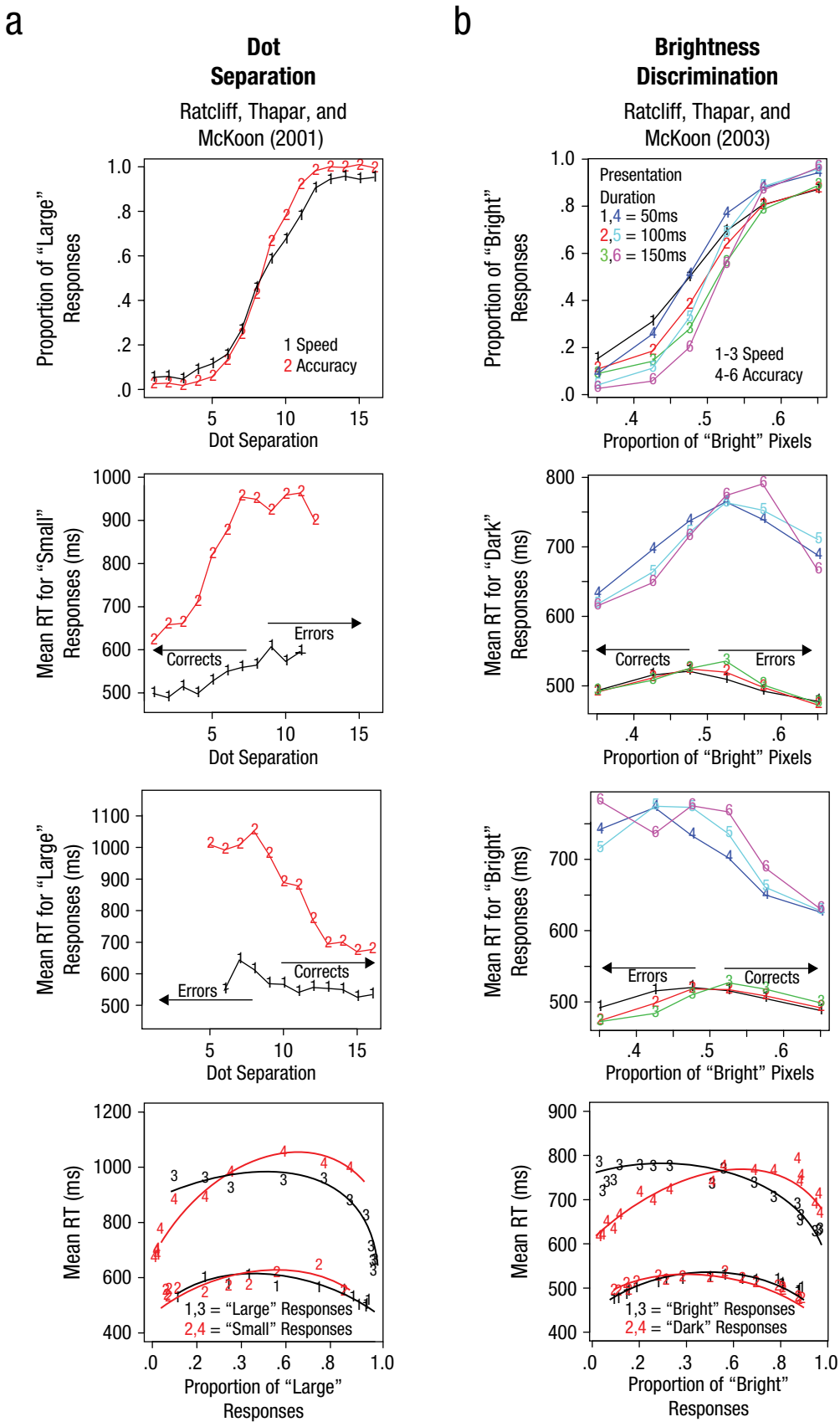
The regularity of data from two-choice tasks is compelling, and it places strong constraints on models that are developed to explain it. Here, we illustrate four regularities with the data from the three experiments mentioned above in which participants judged the distance between two dots as "large" or "small" (Ratcliff et al., 2001), judged the brightness of a random array of pixels as "dark" or "bright" (Ratcliff et al., 2003), and judged a word as "new" or "old" according to whether it had appeared in a preceding list of to-be-remembered words (Ratcliff et al., 2004). For the brightness task, the exposure duration of the stimuli was varied in addition to brightness (the proportion of white vs. black pixels). For the memory task, the subjects in one group were college-aged and in the other they ranged from 60 to 75 years old; word frequency and the number of repetitions were varied. All of the experiments were run in blocks of trials in which either speed or accuracy of responding was emphasized. The speed-versus-accuracy manipulation was crossed with the parametric manipulations of stimulus or memory discriminability.

The first of the regularities is the relationship between mean RT and choice probability—specifically, how they covary as stimulus discriminability and instructions to subjects are changed. The second is the shape of RT distributions, and the third is how their means, dispersion (variance), and shape change as a function of stimulus discriminability and instructions. The fourth is how, across subjects, accuracy and RT are not significantly correlated; knowing that a person's performance on a task is fast or slow does not allow you to predict whether the person is accurate or inaccurate.

Mean RTs and choice probabilities

The top row of Figure 2 shows plots of psychometric functions for the probabilities of "large," "bright," and "old" responses as they varied with stimulus discriminability, speed/accuracy instructions, and, for the brightness task, exposure duration. The second row shows plots of mean RTs for correct responses and errors for "small," "dark," and "new" responses, and the third shows them for "large," "bright," and "old" responses (all plots in Fig. 2 show data averaged over subjects).

The bottom row shows latency-probability plots. It is these plots that directly lay out all of the relations among choice probabilities, stimulus discriminability, speed/accuracy instructions, and mean RTs (the information that is given separately in the above three rows). The top two lines for each latency-probability plot represent responses under accuracy instructions, and the bottom two responses under speed instructions. The black lines represent "large," "bright," and "old" responses, and the red lines represent "small," "dark," and "new" responses. What is compelling from the latency-probability plots is that the points all fall on very regular, inverted U-shaped functions. With these functions, it is easy to see how mean RT changes with accuracy. It is also easy to see whether mean RTs for correct responses are faster or



(continued)

Fig. 2. *(continued)*

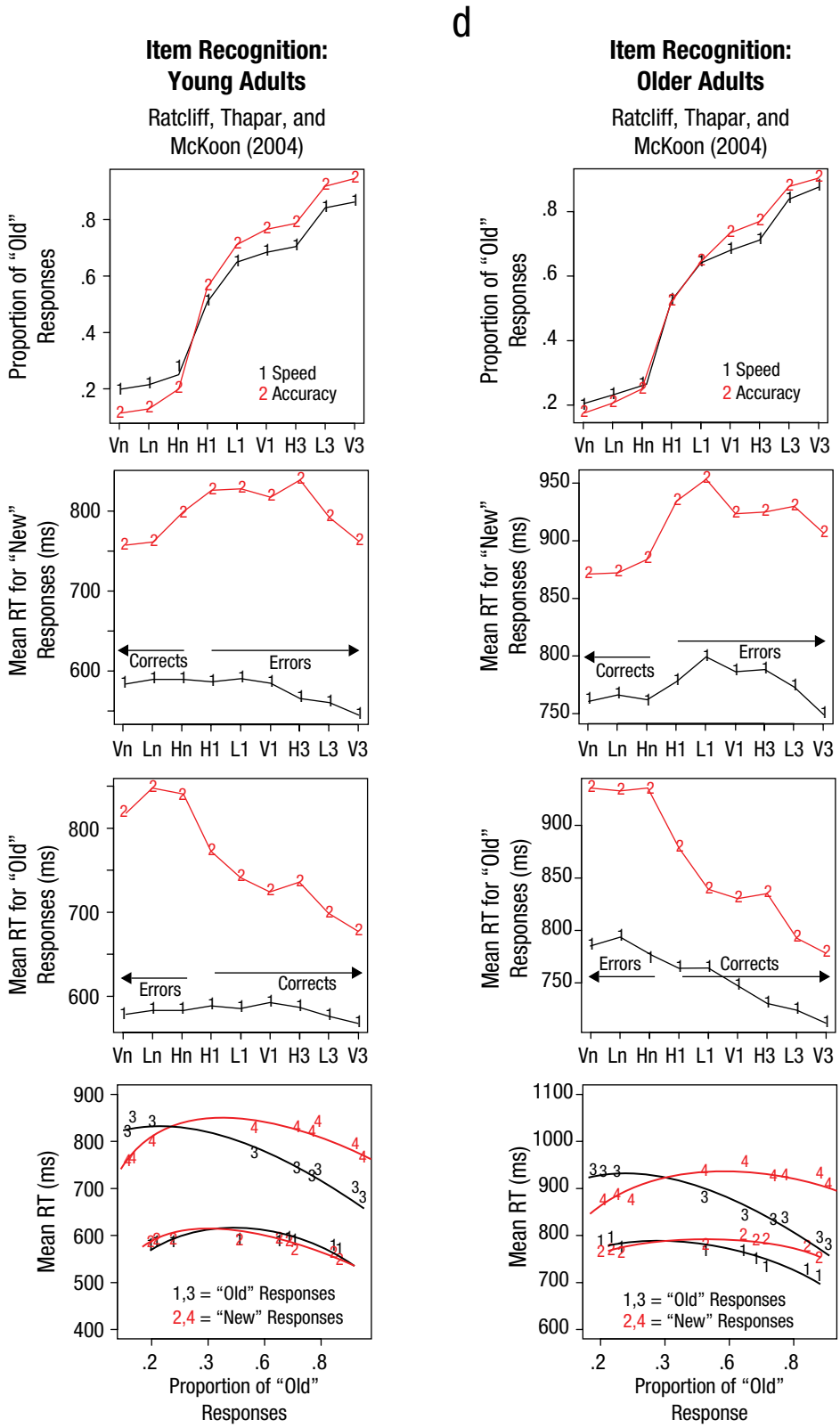


Fig. 2. Plots of accuracy and mean response times (RTs) against the dependent variables for four experiments (each experiment is in a vertical column). The graphs in the first row plot the proportion of responses against the independent variable. The graphs in the second and third rows show mean RTs in the two stimulus conditions for each experiment. For the first three rows, the response plotted is shown in the *y*-axis label. The fourth row shows plots of mean RT against response proportions for each condition (the lines are hand drawn; for fits, see Fig. 4). Notations on the *x*-axes of panels c and d represent conditions for the item-recognition experiments: V = very low-frequency words, L = low-frequency words, H = high-frequency words, N = new, 1 = one presentation, 3 = three presentations.

slower than mean RTs for errors, a relationship that provides, as we described above, a crucial test of decision-making models.

In the brightness-discrimination task, there is a cross-over, in that errors are slower than correct responses under accuracy instructions but faster than correct responses under speed instructions. For the other experiments, errors are mostly slower than correct responses.

RT distributions

Figure 3a shows one sample subject's RT distributions from the brightness-discrimination task for easy- and difficult-condition correct and error responses under speed and accuracy instructions. The distributions behave in the same ways for the dot-separation task and the memory task. First, all the distributions have the same unimodal, right-skewed shape, and second, they all change in the same way as the difficulty of a decision increases. The tails of the distributions become more spread out, indicating an increased number of slow responses, and the leading edges change very little. This is true for both correct and error responses, and it is true with both speed and accuracy instructions.

Another way, and perhaps a more transparent way, to represent the regularities in the shapes of RT distributions is to use quantile-quantile (Q-Q) plots. One condition of an experiment is selected—usually the easiest condition—and then the quantiles of the RT distributions from all the other conditions are plotted against the quantiles of the selected condition. The Q-Q plots in Figure 3b were constructed using the .1, .3, .5, .7, and .9 quantiles averaged over subjects. These quantiles represent, respectively, the fastest 10% of responses, the fastest 30%, the median RT, the slowest 30%, and the slowest 10%. The figure shows four sets of Q-Q plots, one for the speed-instructions condition of the brightness-discrimination task, one for the accuracy-instructions condition of the brightness-discrimination task, one for the speed-instructions condition of the recognition-memory task, and one for the accuracy-instructions condition of the recognition-memory task. The resulting functions are close to linear, which means that the shapes of the RT distributions are similar across all the conditions: One distribution can be obtained from another by equating their location parameters and changing the underlying time scale. The near linearity of Q-Q functions is a regularity that a model must capture, and the diffusion model does so. The inserts in the figure show the Q-Q functions derived from fits of the diffusion model to the data. The Q-Q plots for the dot-separation task, not shown, have the same linearity.

From the histograms in Figure 3, the changes in RT distributions due to changes in difficulty and speed/accuracy instructions are somewhat difficult to see (especially because these are for only one subject in the experiment). This can be remedied by representing the data in quantile-probability plots, which are like latency-probability plots except that they show entire RT distributions, not just mean RTs.

Figure 4a shows an RT distribution with arrows marking the quantiles. To construct a quantile probability function, equal-area rectangles are drawn between the pairs of quantiles below .9 and above .1 and half the areas outside the .1 and .9 quantiles. These rectangles capture the shape of the RT distribution.

Figure 4b illustrates how histograms are translated into quantile-probability functions. The *x*-axis represents the proportion of responses of one kind—for example, the proportion of "bright" responses. The responses with choice probabilities above .5 represent correct responses (for bright stimuli), and the probabilities below .5 represent errors (for dark stimuli). The two columns on the right might correspond to two conditions for which "bright" was the correct response, one more difficult with a lower probability of a correct response and the other easier with a higher probability of a correct response. The two columns on the left might correspond to two conditions for which the correct response was "dark," one with a higher probability of a "bright" error response than the other.

Figures 4c and 4d show the quantile-probability functions for the brightness-discrimination and item-recognition tasks with young subjects. To produce these functions, the quantile RTs and choice probabilities for each condition were computed for each subject and then averaged, giving 18 conditions in the brightness task and 9 conditions in the memory task. The numbers on the functions are the data points, and the *x*s and lines through them are fits of the diffusion model. The model predicts that there will be small changes in the leading edges of the distributions and large spreads in the tails as difficulty increases.

Correlations between an individual's accuracy and response time

The common-sense hypothesis would be that individuals with higher accuracy have shorter RTs (i.e., if someone is better at a task, he or she will be faster). This predicts significant negative correlations between an individual's RT and accuracy, but we have repeatedly found none (Ratcliff et al., 2010; Ratcliff, Thompson, & McKoon, 2015). For the data we have used for this article, Figure 5 shows mean RTs plotted against accuracy for the dot-separation, brightness-discrimination, and recognition-memory experiments. Each point on the plots represents one subject. The correlations range from −.08 to .21, more positive than negative. The implication of this lack

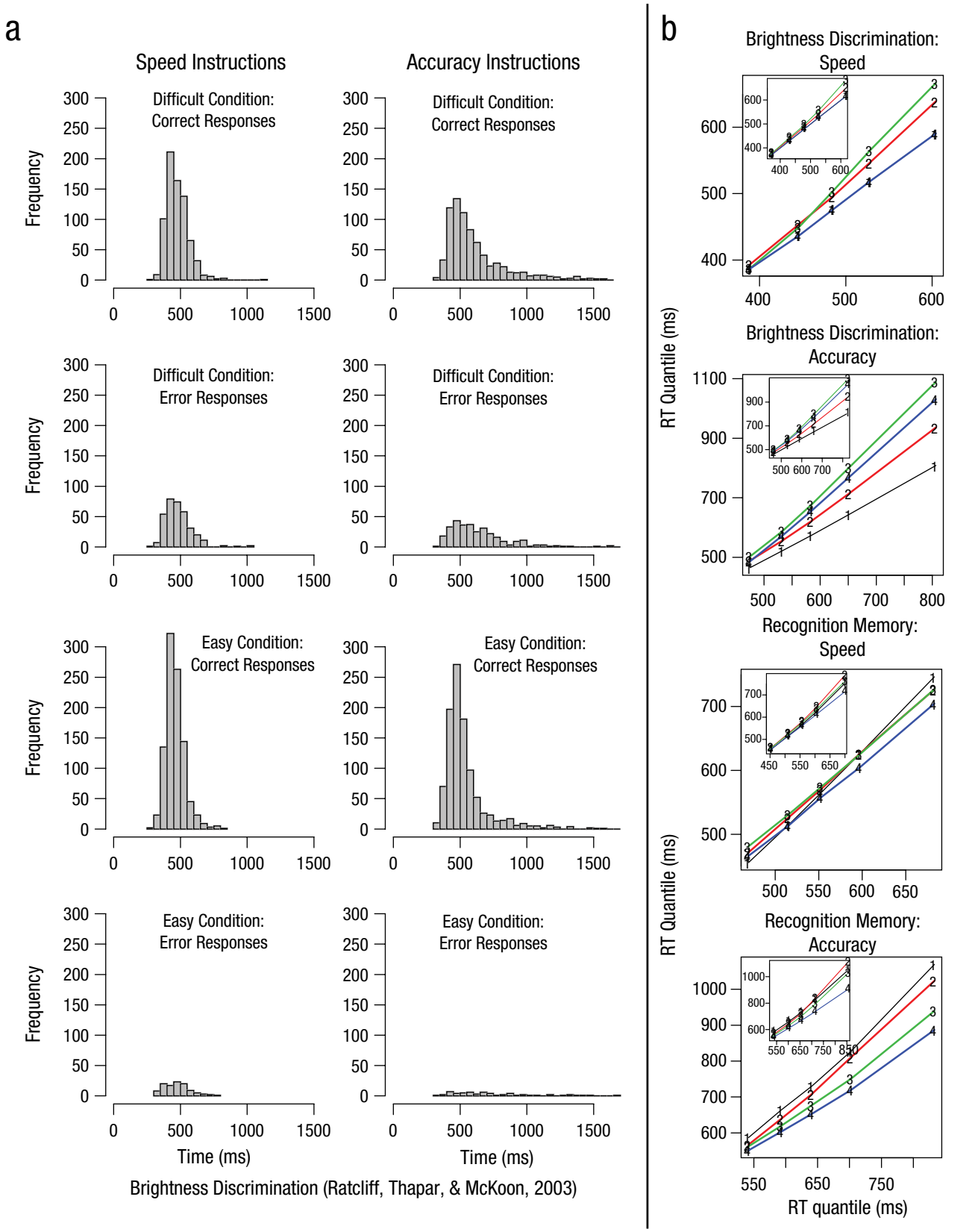
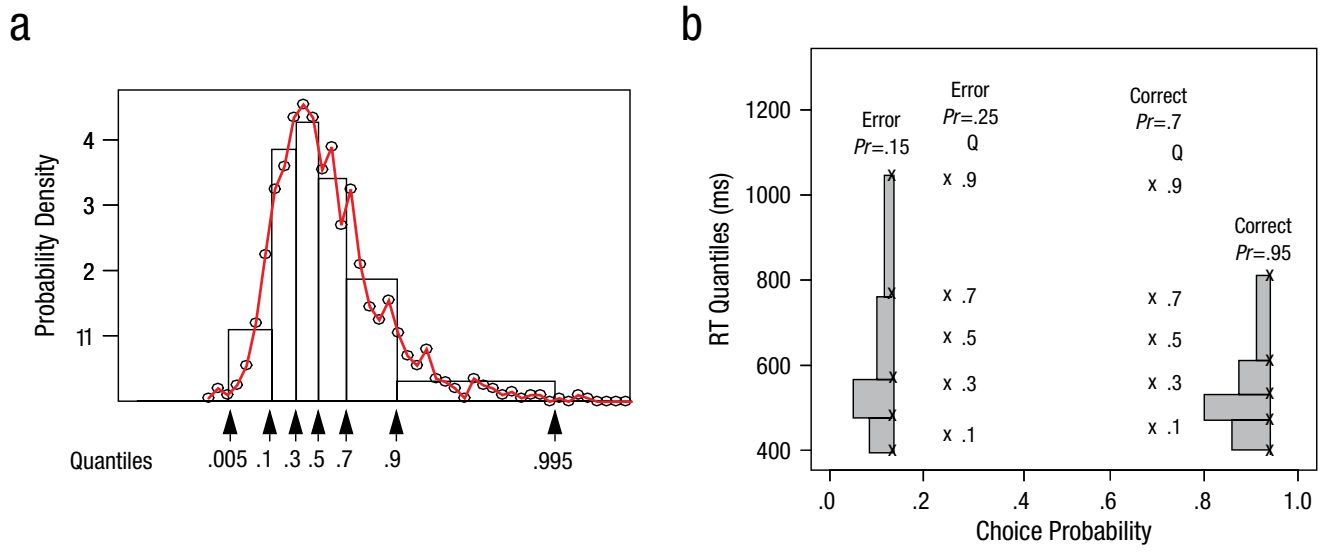
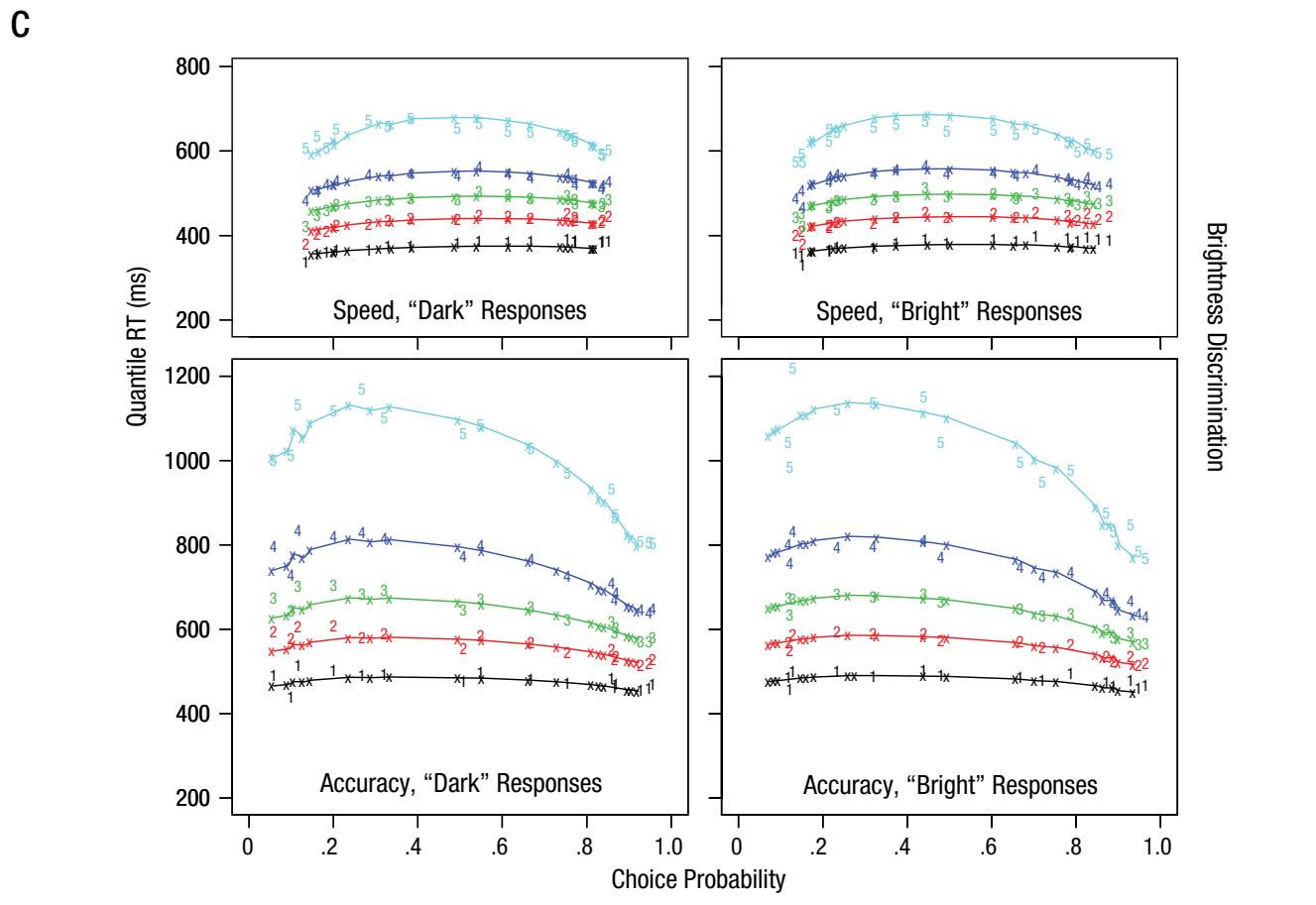


Fig. 3. Examples of response-time (RT) distributions and sample quantile-quantile (Q-Q) plots (Ratcliff, Thapar, & McKoon, 2003). Panel a shows one subject's RT distributions for easy- and difficult-condition correct and error responses under speed and accuracy instructions. Panel b shows Q-Q plots for brightness discrimination and recognition memory with young subjects. One condition was selected and its quantiles computed, then four conditions were selected that spanned the fastest to slowest conditions, and the quantiles of these conditions were plotted against the quantiles for the first condition. The insets are the Q-Q plots for the corresponding diffusion-model fits.





(continued)

Fig. 4. *(continued)*

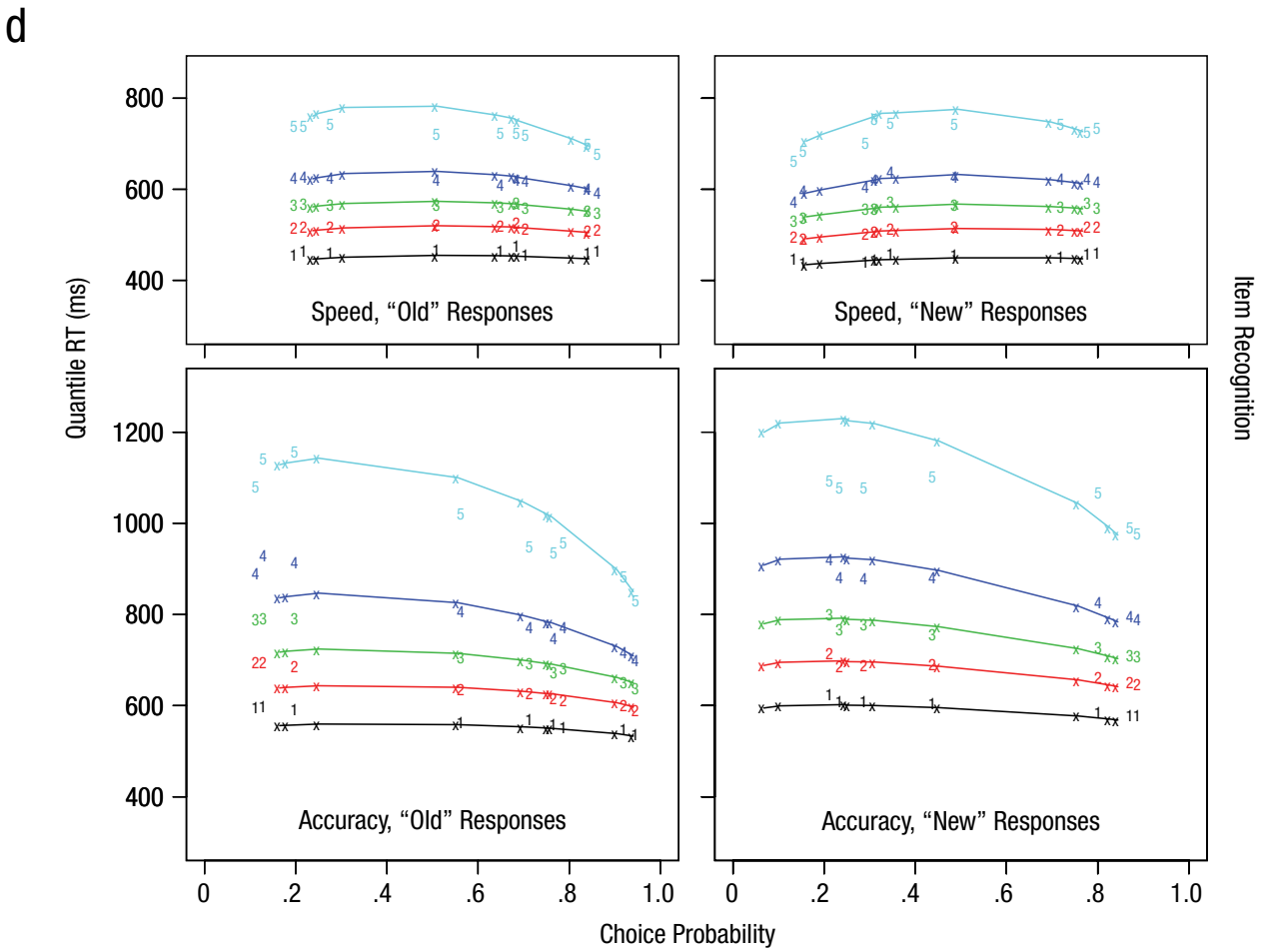


Fig. 4. Examples of how to understand quantiles and sample quantile probability plots. Panel a shows a response-time (RT) distribution overlaid with rectangles of area .2 constructed from the .1, .3, .5, .7, and .9 quantiles (representing the fastest 10% of responses, the fastest 30%, the median RT, the slowest 30%, and the slowest 10%, respectively). Outside the .1 quantile and .9 quantiles are rectangles of area .1 drawn from the .005 to .1 and .9 to .995 quantiles. The quantile rectangles capture the main features of the RT distribution and are therefore a reasonable summary of overall distribution shape. Panel b shows quantile RTs for the .1, .3, .5, .7, and .9 quantiles (stacked vertically) plotted against response proportion for four example conditions. Correct responses are plotted to the right, and error responses to the left. Panels c and d show quantile-probability functions for the brightness-discrimination (Ratcliff, Thapar, & McKoon, 2003) and recognition-memory tasks (Ratcliff, Thapar, & McKoon, 2004) with young subjects. The numbers 1 through 5 represent the .1, .3, .5, .7, and .9, quantiles, and the *x*s and lines are diffusion-model fits.

of correlation is that RT and accuracy may be telling us different things about processing, and hypotheses about individual differences based on one dependent variable could be contradicted by the behavior of the other dependent variable.

Fits of the Diffusion Model

The diffusion model explains the regularities we have discussed. Individually and together, these regularities place strong constraints on the diffusion model and on any competitor model that might be developed. The model explains why latency-probability plots have the regular, inverted-U shapes that they do (see fits of the model in Fig. 4c and 4d). RTs are slower for more difficult conditions because drift rates are lower. Processes with lower drift rates take longer to reach a criterion, and, importantly, this means that error RTs slow with difficulty in the same way that RTs for correct responses do. The model also explains why the plots have the same shapes with instructions that stress speed as with instructions that stress accuracy: Criterion settings are largely independent of drift rate, and it is drift rates that determine the shapes. Latency-probability plots also show whether errors are faster or slower than correct responses, something handled by the diffusion model with different mixtures that arise from across-trial variability in drift rate and starting point (Ratcliff & McKoon, 2008).

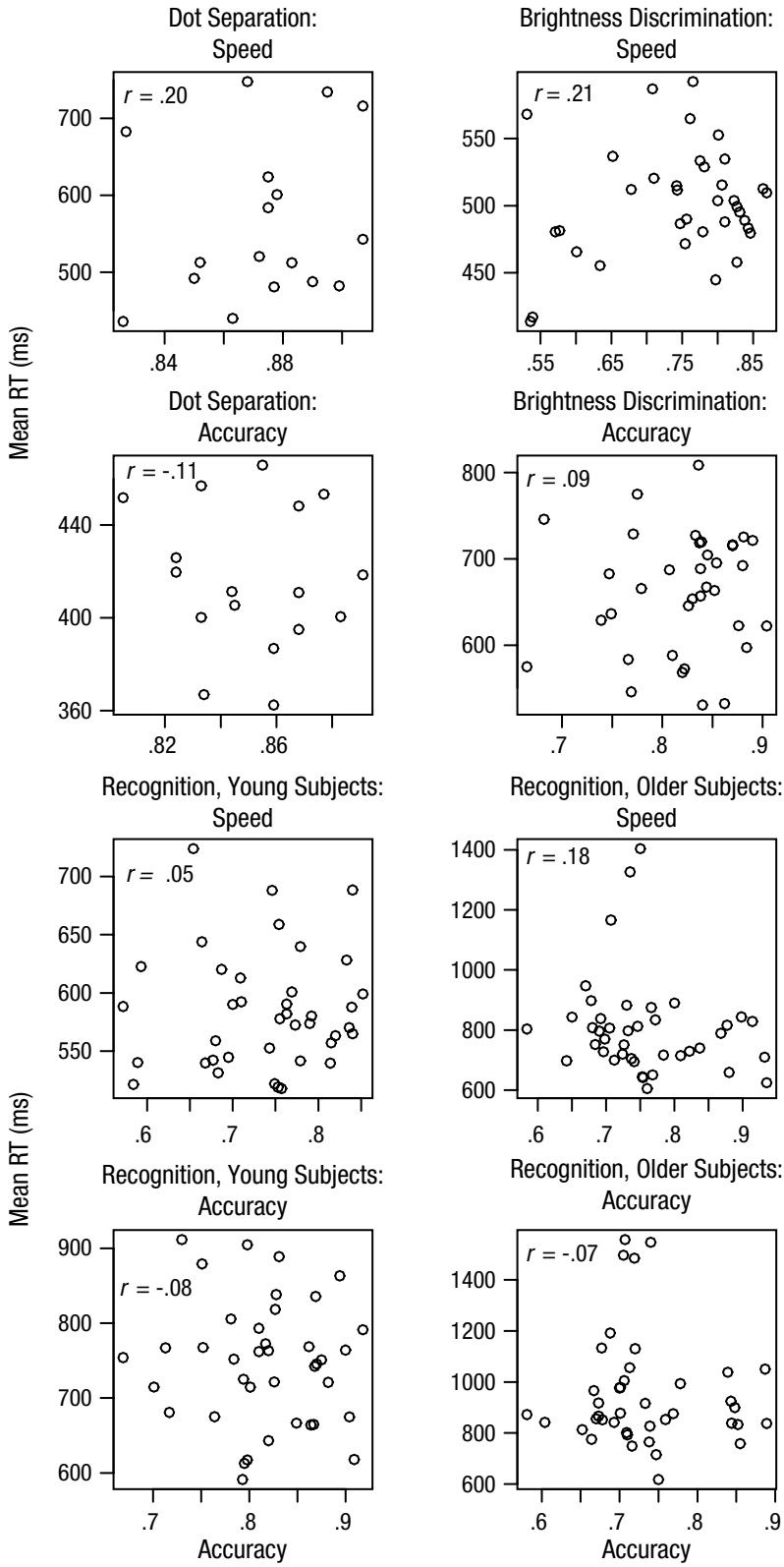


Fig. 5. Mean response time (RT) and accuracy (averaged over conditions) for each subject in a distance-discrimination task (Ratcliff, Thapar, & McKoon, 2001; top left graphs), a brightness-discrimination task (Ratcliff, Thapar, & McKoon, 2003; top right graphs), and a recognition task (Ratcliff, Thapar, & McKoon, 2004), plotted against each other. The correlations are shown as insets.

The model also explains, by simple geometry, why RT distributions have the shapes that they do (see fits of the model for the Q-Q plots in Fig. 3b). Indeed, it has little or no flexibility in predicting anything else, as Ratcliff (2013) showed. This invariance extends to families of distributions from different experiments. If, for example, the quantiles from a perceptual experiment with one group of subjects are plotted against the quantiles from a memory experiment with another group of subjects, then the resulting Q-Q plots will again be nearly linear, much like those in Figure 2b. This invariance is predicted by the diffusion model (there are some exceptions—e.g., plots from accuracy- vs. speed-instructions conditions—but the diffusion model predicts such exceptions). The model also explains why RT distributions change as they do across levels of difficulty—small changes in the leading edge and larger spread in the tail (see fits of the model to quantile-probability functions in Figures 4c and 4d—the only discrepancies in fit being in the .9 quantiles for item recognition).

One of the uses of the diffusion model is as a psychometric tool in individual-differences research. The model cannot predict what differences among individuals will be observed, because performance depends both on the quality of evidence entering the decision process and on the decision criteria adopted by the subjects. However, it can be used to understand individual differences among tasks or conditions and to characterize correlations in model parameters across tasks (Ratcliff et al., 2010; Ratcliff et al., 2015).

The model also accounts for the lack of correlation between accuracy and RTs across subjects, as shown in Figure 5. Variability in drift rates across trials produces large changes in accuracy but only small to moderate changes in RT, especially when the criteria are close together, as occurs under speed-emphasis instructions. Variability in decision criteria produces large changes in RT with only small to moderate changes in accuracy. If drift rates and decision criteria vary independently across the subjects in an experiment (as we typically find), mean RT and accuracy will be relatively uncorrelated.

Recently, Jones and Dzhafarov (2014) criticized decision models, including the diffusion model, arguing that if their assumptions about sources of variability in performance are relaxed, then the models can be made to fit any set of data and thus are unfalsifiable. In response, Smith, Ratcliff, and McKoon (2014) pointed out that the unfalsifiable model constructed by Jones and Dzhafarov was a deterministic or near-deterministic model, in which they had replaced moment-by-moment variability in the evidence entering the decision process with arbitrary distributions of trial-to-trial variability. This construction fundamentally altered the properties of the model and meant it could make arbitrarily flexible predictions. In contrast to a deterministic model, the standard diffusion model is able to account for data not because of the flexibility of its assumptions but because it makes distinctive predictions that closely match the relationships found in empirical data.

The standard diffusion model applies to two-choice tasks, but there are other models that apply to simple RT and vigilance tasks (Ratcliff & Van Dongen, 2011; Smith, 1995). Multichoice decision making and confidence are also active areas of research in both psychology and neuroscience, and diffusion models feature prominently (see Niwa & Ditterich, 2008; Ratcliff & Starns, 2013). Finally, there other sequential-sampling models that represent noisy evidence accumulation as a diffusion process, and in the comparisons made so far, they exhibit much the same behavior as the two-choice diffusion model (Donkin, Brown, Heathcote, & Wagenmakers, 2011; Ratcliff, Thapar, Smith, & McKoon, 2005).

Conclusion

There are significant regularities in RTs, accuracy, and the ways that they covary across tasks, experimental conditions, individual subjects, and subject populations. A priori, there is no reason why RT distributions should have similar shapes in all of these cases (in fact, only a small proportion of the theoretically possible range of experimental outcomes is actually observed in practice). The fact that distributions do have the same approximate shape is highly diagnostic of the processes that underlie decisions, and it imposes strong constraints on any model of these processes. The diffusion model produces exactly the same kinds of highly constrained RT-distribution shapes as are found experimentally. This close correspondence between theory and data across a range of experimental tasks and subjects suggests that the model is capturing something fundamental about decision making in these tasks, which in turn licenses its use as a psychometric tool for interpreting processing in normal and disordered cognition.

Recommended Reading

Ratcliff, R., & Childers, R. (2015). (See References). A source on fitting methods and individual differences, which also points back to articles on model fitting.

Ratcliff, R., & McKoon, G. (2008). (See References). A review of the model and applications.

Ratcliff, R., & Starns, J. J. (2013). (See References). Modeling confidence and multichoice decision making with diffusion class models.

Shadlen, M. N., & Kiani, R. (2013). (See References). A recent review of diffusion-model applications to neuroscience (though we do not agree with all the points made).

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

Preparation of this article was supported by National Institute on Aging Grant R01 AG041176, Department of Education/ Institute of Education Sciences Grant R305A120189, and Australian Research Council Discovery Grant DP140102970.

References

Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E. J. (2011). Diffusion versus linear ballistic accumulation: Different models for response time, same conclusions about psychological mechanisms? *Psychonomic Bulletin & Review*, *55*, 140–151.

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2015). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology*. Advance online publication. doi:10.1146/annurev-psych-122414-033645

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535–574.

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science*, *274*, 427–430.

Jones, M., & Dzhafarov, E. N. (2014). Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. *Psychological Review*, *121*, 1–32.

Mulder, M. J., Bos, D., Weusten, J. M. H., van Belle, J., van Dijk, S. C., Simen, P., . . . Durson, S. (2010). Basic impairments in regulating the speed-accuracy tradeoff predict symptoms of attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *68*, 1114–1119.

Niwa, M., & Ditterich, J. (2008). Perceptual decisions between multiple directions of visual motion. *Journal of Neuroscience*, *28*, 4435–4445.

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

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

Ratcliff, R. (2013). Parameter variability and distributional assumptions in the diffusion model. *Psychological Review*, *120*, 281–292.

Ratcliff, R., Cherian, A., & Segraves, M. (2003). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of simple two-choice decisions. *Journal of Neurophysiology*, *90*, 1392–1407.

Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model. *Decision*, *2*, 237–279.

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

Ratcliff, R., & Starns, J. J. (2013). Modeling response times, choices, and confidence judgments in decision making: Recognition memory and motion discrimination. *Psychological Review*, *120*, 697–719.

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, *16*, 323–341.

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

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. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology*, *60*, 127–157.

Ratcliff, R., Thapar, A., Smith, P. L., & McKoon, G. (2005). Aging and response times: A comparison of sequential sampling models. In J. Duncan, P. McLeod, & L. Phillips (Eds.), *Measuring the mind: Speed, control, and age* (pp. 3–31). Oxford, England: Oxford University Press.

Ratcliff, R., Thompson, C. A., & McKoon, G. (2015). Modeling individual differences in response time and accuracy in numeracy. *Cognition*, *137*, 115–136.

Ratcliff, R., & Van Dongen, H. P. A. (2009). Sleep deprivation affects multiple distinct cognitive processes. *Psychonomic Bulletin & Review*, *16*, 742–751.

Ratcliff, R., & Van Dongen, H. P. A. (2011). A diffusion model for one-choice reaction time tasks and the cognitive effects of sleep deprivation. *Proceedings of the National Academy of Sciences, USA*, *108*, 11285–11290.

Shadlen, M. N., & Kiani, R. (2013). Decision making as a window on cognition. *Neuron*, *80*, 791–806.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review*, *102*, 567–593.

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

Smith, P. L., Ratcliff, R., & McKoon, G. (2014). The diffusion model is not a deterministic growth model: Comment on Jones and Dzhafarov (2014). *Psychological Review*, *121*, 679–688.

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

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.

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, 7:14. Retrieved from http://journal.frontiersin.org/article/10.3389/fninf.2013.00014/full

Zeguers, M. H. T., Snellings, P., Tijms, J., Weeda, W. D., Tamboer, P., Bexkens, A., & Huizenga, H. M. (2011). Specifying theories of developmental dyslexia: A diffusion model analysis of word recognition. *Developmental Science*, *14*, 1340–1354.*Invited Manuscript*