Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 50 (2004) 408–424
Journal of Memory and Language
www.elsevier.com/locate/jml

A diffusion model analysis of the effects of aging on recognition memory☆

Roger Ratcliff,[a,*] Anjali Thapar,[b] and Gail McKoon[a]

[a] *Department of Psychology, The Ohio State University, 1827 Neil Avenue, Columbus, OH 43210, USA*
[b] *Bryn Mawr College, 101 N. Merion Ave., Bryn Mawr, PA 19010-2899, USA*

Received 3 June 2003; revision received 12 November 2003

Abstract

The effects of aging on response time were examined in a recognition memory experiment with young, college age subjects and older, 60–75 year old subjects. The older subjects were slower than the young subjects but almost as accurate. Ratcliff's (1978) diffusion model was fit to the data and it provided a good account of response times, their distributions, and accuracy values. The fits showed a 100 ms slowing of the nondecision components of response time for older subjects relative to young subjects, and roughly equal response criteria settings with accuracy instructions but more conservative settings for the older subjects with speed instructions. In the diffusion model, the decision process is driven by the rate of accumulation of evidence from the stimulus. We found that the rate of accumulation for older subjects was a non-significant 7% lower than the rate for young subjects, indicating that the output from recognition memory entering the decision process was not significantly worse for the older subjects. The results are compared to those obtained from letter discrimination, brightness discrimination, and signal detection-like tasks.

Keywords: Diffusion; Recognition; Aging; Response time; Speed–accuracy

A central finding in the literature on aging is that as people age, their response times increase. Along with the increase in response times, performance sometimes shows a decrease in accuracy. Recently, Ratcliff, Thapar, and McKoon (2001, 2003) and Thapar, Ratcliff, and McKoon (2003) examined the effects of aging on performance in several two-choice tasks: signal detection-like tasks, a brightness discrimination task, and a letter discrimination task. By applying the diffusion model (Ratcliff, 1978, 1981, 1985, 1988; Ratcliff & Rouder, 1998, 2000; Ratcliff, Van Zandt, & McKoon, 1999) to the data, they were able to separate out the effects of aging on several components of processing in the model. The older subjects in their studies adopted more conservative decision criteria than the young subjects and they were also slower in components of processing outside the decision itself (e.g., encoding and response execution). In some tasks, the quality of the stimulus evidence driving the decision process was not significantly lower for the older subjects than the young ones, although in other tasks, it showed a deficit. For the brightness and letter discrimination tasks, the deficits occurred exactly as would be predicted from psychophysical research on the effects of aging on visual discrimination (Coyne, 1981; Fozard, 1990; Owsley, Sekuler, & Siemsen, 1983; Spear, 1993): a deficit occurred with the high spatial frequencies in letters in the letter discrimination task but not with the low spatial frequencies of the stimuli in the brightness discrimination task.

☆ Preparation of this article was supported by NIA Grant R01-AG17083, NIMH Grants R37-MH44640 and K05-MH01891.
* Corresponding author.

0749-596X/$ - see front matter
doi:10.1016/j.jml.2003.11.002

Not only does the diffusion model allow the response time and accuracy data from a two-choice task to be analyzed in terms of the components of processing required by the task, it also has an advantage over other approaches because it deals with all the aspects of the data: correct and error response times and their relative speeds, the shapes of response time distributions for both correct and error responses, and accuracy values. The model has been successful in accounting for the data from a variety of two-choice response time tasks and a variety of task manipulations (Ratcliff, 1978, 1981, 1988, 2002; Ratcliff, Gomez, & McKoon, in press; Ratcliff & Rouder, 1998, 2000; Ratcliff et al., 2001, 2003; Ratcliff et al., 1999; Thapar et al., 2003). In this article, the diffusion model is applied to examine the effects of aging on recognition memory.

Previously, the main conclusion in the literature has been that aging has little effect on recognition memory; data show no effect at all or only a small effect (Balota, Dolan, & Duchek, 2000; Bowles & Poon, 1982; Craik, 1994; Craik & Jennings, 1992; Erber, 1974; Gordon & Clark, 1974; Kausler, 1994; Naveh-Benjamin, 2000; Neath, 1998, Chap. 16; Rabinowitz, 1984; Schonfield & Robertson, 1966). However, the studies from which these conclusions are drawn have measured only recognition accuracy, not response time. As mentioned above, older adults are typically slower, often much slower, than young adults on cognitive tasks. This presents a puzzle: slowing for older adults has often been interpreted as a deficit such that, for example, cognitive operations are not fully completed in the available time and the products of earlier operations are not fully available for later operations (e.g., Salthouse, 1996). In this context, the findings that older adults show no deficit in recognition memory relative to young subjects are surprising.

The experiment presented here measured both accuracy and response time. Simultaneously accounting for both measures, the diffusion model allows the separation of components of processing. In particular, the aim was to determine whether older adults are slower than young adults and if so, to determine whether the slowing reflects a deficit in the quality of information available from memory or increased processing time for some other, non-memory, components of processing.

In the experiment, the number of repetitions of the words to be learned and word frequency were manipulated. With these two variables, accuracy can be varied from high to moderately low. Sweeping out response times over a wide range of accuracy values provides maximal constraints on fitting the diffusion model to data (Ratcliff & Tuerlinckx, 2002). We also manipulated speed–accuracy criteria to provide additional constraints on the model. For some trials, subjects were instructed to respond as quickly as possible, and on other trials, as accurately as possible. The overall aim was to allow the effects of aging on response time and accuracy to be examined in a unified framework, separating out the quality of the memory information entering the decision process from the speed–accuracy decision criteria adopted by subjects and from non-decision components of processing.

The diffusion model

The diffusion model is a model of the cognitive processes involved in making simple two-choice decisions. It separates the quality of evidence entering the decision from the decision criteria and from other, non-decision processes such as encoding the stimulus and response execution. The model applies only to relatively fast two-choice decisions (mean response times less than about 1000–1500 ms) and only to decisions that are a single-stage decision process (as opposed to the multiple-stage processes that might be involved in, for example, reasoning tasks or card sorting tasks). Other models in the class of diffusion models have been applied to decision making (Busemeyer & Townsend, 1993; Roe, Busemeyer, & Townsend, 2001) and simple reaction time (Smith, 1995). For a detailed comparison of the sequential sampling models, see Ratcliff and Smith (in press).

The diffusion model assumes that decisions are made by a noisy process that accumulates information over time from a starting point toward one of two response criteria or boundaries, as in Fig. 1, where the starting point is labeled z and the boundaries are labeled a and 0. When one of the boundaries is reached, a response is initiated. The rate of accumulation of information is

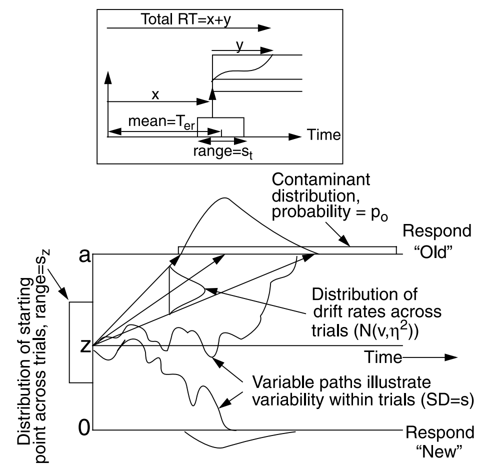

Fig. 1. An illustration of the diffusion model. Parameters of the model are: a, boundary separation; z, starting point; T_{er}, mean value of the non-decision component of response time; η, *SD* in drift across trials; s_z, range of the distribution of starting point (z) across trials; v, drift rate; p_0, proportion of contaminants; s_t, range of the distribution of non-decision times across trials; and s, *SD* in variability in drift within trials.

called the drift rate (v), and it is determined by the quality of the information extracted from the stimulus. In the case of recognition memory, this would be the quality of the match between the test word and memory. For example, a very low frequency word presented for study three times in an experiment would have a high degree of match and therefore a high drift rate, whereas a high frequency word presented once would have a low degree of match and therefore a low drift rate. There is noise (variability) in the process of accumulating information from the starting point toward the boundaries so that processes with the same mean drift rate do not always terminate at the same time (producing response time distributions) and do not always terminate at the same boundary (producing errors). This source of variability is called "within trial" variability. Empirical response time distributions are positively skewed and in the diffusion model, this is naturally predicted by simple geometry: the three arrows from the starting point to the upper boundary in Fig. 1 have equal size differences in drift but produce unequal differences in response time at the upper boundary (see also Fig. 1, Ratcliff & Rouder, 1998).

Components of processing are assumed to be variable across trials and the assumption of such variability allows the model to account for differences in response times between correct and error responses (see Luce, 1986). Variability in drift rate across trials leads to slow errors and variability in starting point leads to fast errors (see Ratcliff & Rouder, 1998; Ratcliff et al., 1999). Drift rate is assumed to be normally distributed with standard deviation η and starting point is assumed to be uniformly distributed with range s_z.

In the experiment presented in this article, subjects are sometimes instructed to respond as quickly as possible and sometimes to respond as accurately as possible. Speed–accuracy tradeoffs are modeled by altering the boundaries of the decision process—wider boundaries require more information before a decision can be made and this leads to more accurate and slower responses.

The non-decision components of processing such as encoding and response execution are combined in the diffusion model into one component with mean T_{er} (see the inset in Fig. 1). Like drift rate and starting point, the non-decision component of processing is assumed to have variability across trials, uniformly distributed with range s_t. The effect of this source of variability depends on the mean value of drift rate (Ratcliff & Tuerlinckx, 2002). With a large value of mean drift rate, variability in the non-decision component acts to shift the leading edge of the response time distribution shorter than it would otherwise be (by as much as 10% of s_t). With smaller values of drift rate, the effect is smaller. The estimated standard deviation in the distribution of the non-decision component of processing is typically less than one quarter the standard deviation in the decision process, and so the combination of the two (the convolution) has little effect on the shape of the combination or on the standard deviation of the distribution. For example, if $s_t = 100$ ms ($SD = 28.9$ ms) and the SD in the decision process is 100 ms, the combination (square root of the sum of squares) is 104 ms. With variability in the non-decision component of processing, Ratcliff and Tuerlinckx showed that the diffusion model could fit data with considerable variability in the .1 quantiles of the response time distributions across experimental conditions.

In sum, the parameters of the diffusion model correspond to the components of the decision process as follows: z is the starting point of the accumulation of evidence, a is the upper boundary and the lower boundary is set to 0; η is the standard deviation in mean drift rate across trials; s_z is the range of the starting point across trials; T_{er} is the mean time taken up by the non-decision components of processing, and s_t is the range of the values of T_{er} across trials. For each different stimulus condition in an experiment, it is assumed that the rate of accumulation of evidence is different and so each has a different value of drift, v. We assume that the standard deviation in drift rate (η) across trials is the same for all stimulus conditions even though it might be expected to vary. However, η is not well estimated; quite large changes in this parameter have small changes on predictions except for error response times which are usually highly variable (see Ratcliff & Tuerlinckx, 2002). Allowing the parameter to vary across conditions would result in minimal improvement to the overall fit to the data at a cost of increasing the number of parameters. Within-trial variability in drift rate (s) is a scaling parameter for the diffusion process (i.e., if it were doubled, other parameters could be multiplied or divided by two to produce exactly the same fits of the model to data). s is set to 0.1 in fits to the data as it has been in other applications of the model to data.

The diffusion model provides a meeting point between the decision process and memory, namely the values of drift rate extracted from the data by the model. These represent the degree of match between the test items in each experimental condition and memory, and so serve as familiarity values similar to the familiarity values of the global memory models (e.g., Gillund & Shiffrin, 1984; Hintzman, 1986; McClelland & Chappell, 1998; Murdock, 1982; Shiffrin & Steyvers, 1997). For a complete theoretical account of memory, the memory models could be used to generate values of familiarity for the different conditions of an experiment and these could be used as values of drift rate in the diffusion model to produce accuracy and response time values.

Experiment

The experiment used a standard recognition memory procedure with 20 study-test lists per session. For each list, high, low, and very low frequency words were presented for study once or three times each. The very low frequency words were known to subjects (they were not "rare" unfamiliar words like those used by Wixted, 1992). A test list immediately followed each study list. Subjects were instructed to respond to each word of the test list according to whether it had or had not appeared on the study list (responding "old" or "new"). On alternate lists, subjects were given speed or accuracy instructions. For speed lists, they were asked to respond as quickly as possible and for accuracy lists, they were asked to respond as accurately as possible.

Method

Subjects

Thirty-nine young adults (16 men and 23 women) and 41 older adults (11 men and 30 women) participated in the experiment. The young adults were college students who participated for course credit in an introductory psychology course at Northwestern University. The older adults were healthy, active, community-dwelling individuals, aged 60–75 years old, living in Evanston, IL or the nearby suburbs and they were paid for their participation. The older subjects had to meet the following inclusion criteria to participate in the study: a score of 26 or above on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975); a score of 15 or less on the Center for Epidemiological Studies-Depression Scale (CES-D; Radloff, 1977); and no evidence of disturbances in consciousness, medical or neurological disease causing cognitive impairment, head injury with loss of consciousness, or current psychiatric disorder. The means and standard deviations for standard background characteristics are presented in Table 1. The characteristics for the young subjects were collected from a large pool of subjects from the same population as those tested in the experiment (although the experimental group was not a subset of the larger group). Subjects were tested individually for 2, 3, or 4 sessions, the number of sessions required to produce two sessions of stable data (i.e., responses were not becoming significantly faster from session to session).

Table 1
Subject characteristics

Measure	Older adults		Young adults	
	M	SD	M	SD
M age	70.0	4.2	19.6	0.5
Years of education	16.4	2.1	12.6	0.9
MMSE	29.1	0.9	29.0	1.1
WAIS-R vocabulary	15.0	2.2	14.4	1.9
WAIS-R picture completion	14.5	2.5	10.8	2.5
CESD: Total	7.4	6.4	8.4	4.6

Note. MMSE = Mini-Mental State Examination; WAIS-R = Wechsler Adult Intelligence Scale-Revised; CESD = Center for Epidemiological Studies–Depression Scale. The subjects background characteristics are for a group of 133 subjects from the same pool as those tested here.

Stimuli

The stimuli were high, low, and very low frequency words. There were 800 high frequency words with frequencies from 78 to 10,600 per million (mean = 325, $SD = 645$, Kucera & Francis, 1967); 800 low frequency words with frequencies of 4 and 5 per million (mean = 4.41, $SD = 0.19$); and 741 very low frequency words, with frequencies of 1 per million or no occurrence in the Kucera and Francis' corpus (mean = .365; $SD = .48$). All the very low frequency words occurred in the Webster's Ninth Collegiate Dictionary (1990), and they were screened by three Northwestern undergraduate students and any words they did not know were eliminated. Stimuli were chosen randomly without replacement from the pools.

Apparatus

Stimuli were presented on a Pentium II class machine and responses were collected on the keyboard.

Procedure

The experiment consisted of 20 study-test trials per session. Each study list consisted of 36 words, 3 high, 3 low, and 3 very low frequency words presented once, and 3 high, 3 low, and 3 very low frequency words presented 3 times. Each word was displayed for 1 s. The repeated words were presented with at least 2 other words intervening between presentations. Each test list consisted of the 18 studied words and 18 new words, the latter consisting of 6 high, 6 low, and 6 very low frequency words.

Subjects received alternating speed and accuracy blocks of trials. In accuracy blocks, if a response was correct, there was a 250 ms pause and then the next test word; if a word was incorrect, the word "ERROR" was displayed for 300 ms, then erased, then there was a 250 ms pause and then the next test word was presented. In speed blocks, there was no accuracy feedback, and if a response was longer than 800 ms for young subjects and 900 ms for older subjects, "TOO SLOW" was displayed following a 300 ms delay from the response. Then there was a 300 ms pause before the next test word. In both kinds of trials, if a response was shorter than 250 ms, a message "TOO FAST" was displayed for 1500 ms (in order to discourage fast guessing).

In accuracy blocks, subjects were instructed to respond accurately. In the speed blocks, subjects were instructed to respond quickly, using the "TOO SLOW" message as a guide to indicate when they were responding too slowly.

In each session, there were 10 blocks of accuracy trials and 10 blocks of speed trials. Therefore, in each session, there was a total of 30 observations for each studied word condition (each level of word frequency and number of repetitions) and 60 observations for each new word condition (each level of word frequency). A minimum of two sessions per subject were used in data analyses.

Older subjects received one to two sessions of practice. Altogether, they participated in four sessions and the sessions for which data were included in analyses were those for which accuracy and response time had stabilized. Young subjects were stable after only a few practice trials. For both older and young subjects, the data from the first two study-test blocks in the first session used for the data analyses were eliminated.

Results

Response times smaller than 300 ms and greater than 3000 ms were eliminated for young subjects (less than 0.8% of the data) and response times smaller than 300 ms and greater than 4000 ms were eliminated for older subjects (less than 0.3% of the data). The different upper cutoffs were required because response times for older subjects were longer than those for young subjects. Further discussion of outliers and contaminants is presented in the section on fitting the diffusion model to the data.

A summary of the results is shown in Table 2 with values of response time and accuracy collapsed over number of repetitions of study words and word frequency. With speed instructions, accuracy is about the same for the young and older subjects whereas with accuracy instructions, accuracy is about 5% higher for the young subjects. Response times show larger differences. With speed instructions, young subjects' correct responses average about 160 ms shorter than older subjects', whereas with accuracy instructions, young subjects' correct responses are about 100 ms shorter. Error response times are generally longer than correct response times except for the young subjects with speed instructions where there is little difference between correct and error response times.

The speed–accuracy manipulation produced about a 7% change in accuracy and a 170 ms change in correct response times for the young subjects, but only a 3% change in accuracy and a 100 ms change in correct response times for the older subjects. This means that, with speed instructions, the young subjects were more willing to sacrifice accuracy for speed than the older subjects.

For "old" responses to studied words, very low frequency words have higher accuracy values than low frequency words which in turn have higher accuracy values than high frequency words. Also, words presented three times have higher accuracy values than words presented once. For "new" responses to nonstudied words, responses to very low frequency words were more accurate than responses to low frequency words which in turn were more accurate than responses to high frequency words (i.e., a mirror effect, Glanzer & Adams, 1985). These effects can be read from Figs. 3 and 4 on the x-axis for probability and on the y-axis for median response time. The main effects of word frequency and speed–accuracy condition were all significant for older and young subjects for old and new items for both accuracy and response time except that the effect of word frequency was not significant for response time for older subjects. The effect of repetition was significant for older items for both old and young subjects for both accuracy and response time. There were a number of significant two-way interactions with speed–accuracy condition that showed a reduction in the size of the effects going from the accuracy to the speed conditions (e.g., both older and young subjects showed interactions for response time for speed–accuracy condition with both repetition and word frequency).

Table 2
Mean response time and accuracy values for the experiment collapsed over repetitions and word frequency

Subjects	Item type and speed–accuracy condition	"Old" response time	"Old" response probability	"New" response time	"New" response probability
Young	New speed	581	.228	587	.772
Young	Old speed	582	.708	577	.292
Young	New accuracy	848	.147	775	.853
Young	Old accuracy	723	.781	829	.219
Older	New speed	783	.231	763	.769
Older	Old speed	737	.712	783	.288
Older	New accuracy	933	.209	875	.791
Older	Old accuracy	820	.742	933	.258

Note. The number of observations per row for young subjects is about 13,200 and for older subjects about 18,100 (the different numbers of observations arose because we used data from three sessions for some older subjects and data from only two sessions for young subjects).

Brinley plots

A procedure that has been standard for examining response times in aging research is to plot older subjects' response times for each experimental condition against young subjects' response times for the same conditions to produce what is called a Brinley plot (Brinley, 1965). The main result from this plotting has been that the function is usually a straight line with a slope greater than 1. This result has been taken as evidence for a decrement in the speed of processing with age, a decrement such that older subjects' response times are lengthened relative to young subjects' response times by a factor that is the slope of the function. This interpretation has been widely accepted (e.g., Birren, 1965; Cerella, 1985, 1990; Cerella, Poon, & Williams, 1980; Fisk & Warr, 1996; Salthouse, 1985, 1996; Salthouse, Kausler, & Saults, 1988) although it has engendered much criticism and debate (Allen, Ashcraft, & Weber, 1992; Allen, Madden, Weber, & Groth, 1993; Cerella, 1994; Fisher & Glaser, 1996; Fisk & Fisher, 1994; Hartley, 1992; Hertzog, 1992; Lima, Hale, & Myerson, 1991; Madden, 1989; Madden, Pierce, & Allen, 1992; Myerson, Ferraro, Hale, & Lima, 1992; Myerson, Wagstaff, & Hale, 1994; Perfect, 1994).

Ratcliff, Spieler, and McKoon (2000) argued against this interpretation from a theoretical perspective by showing that, in the framework of explicit models of processing, there are multiple ways that slowing in older subjects' mean response times can come about, such as changes in the rate of accumulation of evidence or changes in decision criteria settings. The fits of the diffusion model to the data from four experiments (Ratcliff et al., 2001, 2003; Thapar et al., 2003) showed that sometimes differences in decision criterion settings are responsible for most of the increase in the older subjects' response times, along with a small increment in the non-decision components of processing, and sometimes, in addition to these two factors, a decrement in the quality of evidence encoded from the stimuli (i.e., a smaller value of drift rate in the diffusion model) is responsible.

Although in models like the diffusion model, Brinley plots can be produced from any of several different mechanisms, and therefore are not theoretically constraining, we present them here for the data from our experiment to show that our results give the linear functions that have been obtained in previous studies. Fig. 2 shows three fitted straight lines, one for the data from speed blocks, one for the data from accuracy blocks, and one for all the data combined. In each case, the mean response times for correct responses for older subjects are plotted against the mean response times for young subjects for each experimental condition. The points on each function are the points for correct responses for 9 experimental conditions: studied words for 3 and 1 presentations crossed with 3 values of word frequency, plus 3 values of word frequency for new words. For the speed blocks, the slope was 2.00 (intercept −420 ms); for the accuracy blocks, the slope was 0.98 (intercept 115 ms); and for the combined data, the slope was 0.63 (intercept 375 ms). The fact that the slope varies according to whether all or part of the data are plotted illustrates one of the problems with the slowing hypotheses derived from Brinley plot analyses: it would not be expected that the amount of cognitive slowing for older subjects relative to young ones would depend on what portion of the experimental data was considered. What is most surprising in these plots is the fact that the slope is close to 1.0 for the accuracy condition and much less than 1.0 for the combined data. A slope of less than 1.0 would mean, under the slowing interpretation of Brinley slopes, that older subjects' cognitive processing was faster than young subjects'.

Ratcliff, Spieler, and McKoon (in press) have argued that in order to assess the goodness of fit of a straight line to Brinley plots, standard error bars should be presented and these are shown in Fig. 2. If ellipses are drawn around the data points to represent confidence regions, then for the straight lines for the speed and accuracy conditions, all the data points except two or three lie within 2 standard errors of the line, and the exceptions are within 2.5 standard errors. This means that the straight lines provide good fits to the Brinley plots. However, the straight line fit to the combined data

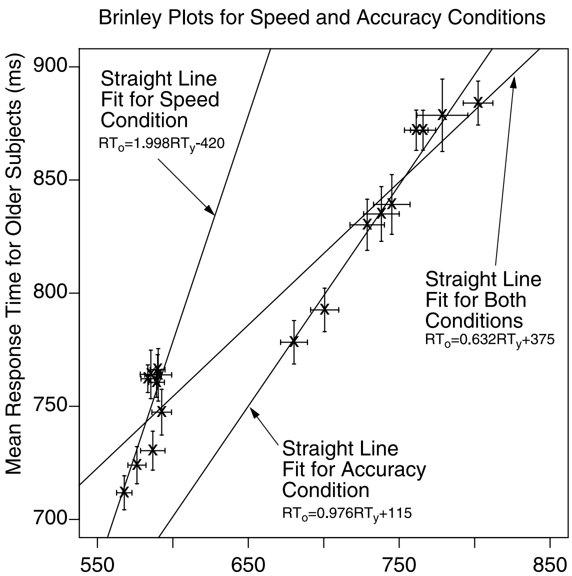

Fig. 2. Brinley plots for mean response times for correct responses. The points on the graph represent the same conditions for older and young subjects (three values of word frequency crossed with items presented three times or one time, and new items). Straight lines are fitted for speed and accuracy conditions separately and for the conditions combined. Error bars are 2 standard errors in mean response time.

shows poor fits; all the data points except one in the speed condition and four in the accuracy condition significantly miss the straight line.

Quantile probability functions

One of the main arguments for the use of quantitative models is that, in order to provide a complete explanation of processing in a task, it is necessary to account for all aspects of experimental data. A model that deals with data only qualitatively is unlikely to be able to make correct quantitative predictions, and a model that deals only with mean response times for correct responses is incomplete and almost guaranteed to make incorrect predictions about other aspects of the data. The diffusion model produces predictions for correct and error response times, their distributions, and accuracy, and it is tested by applying it simultaneously to all these aspects of the data. Plotting all of these aspects of the data separately would make their relative behaviors difficult to grasp, so we display the data in quantile probability functions (Figs. 3 and 4).

On quantile probability functions, the probability of the response determines position on the x-axis and the quantile response times are plotted above each other on the y-axis. For example, if accuracy for low frequency words studied three times was 98% correct, then the quantiles of the response time distribution in this condition for correct responses would be plotted above the .98 point on the x-axis. Specifically, in Figs. 3 and 4, the .1, .3, .5 (median), .7, and .9 quantiles are plotted for each of the 9 experimental conditions: three values of

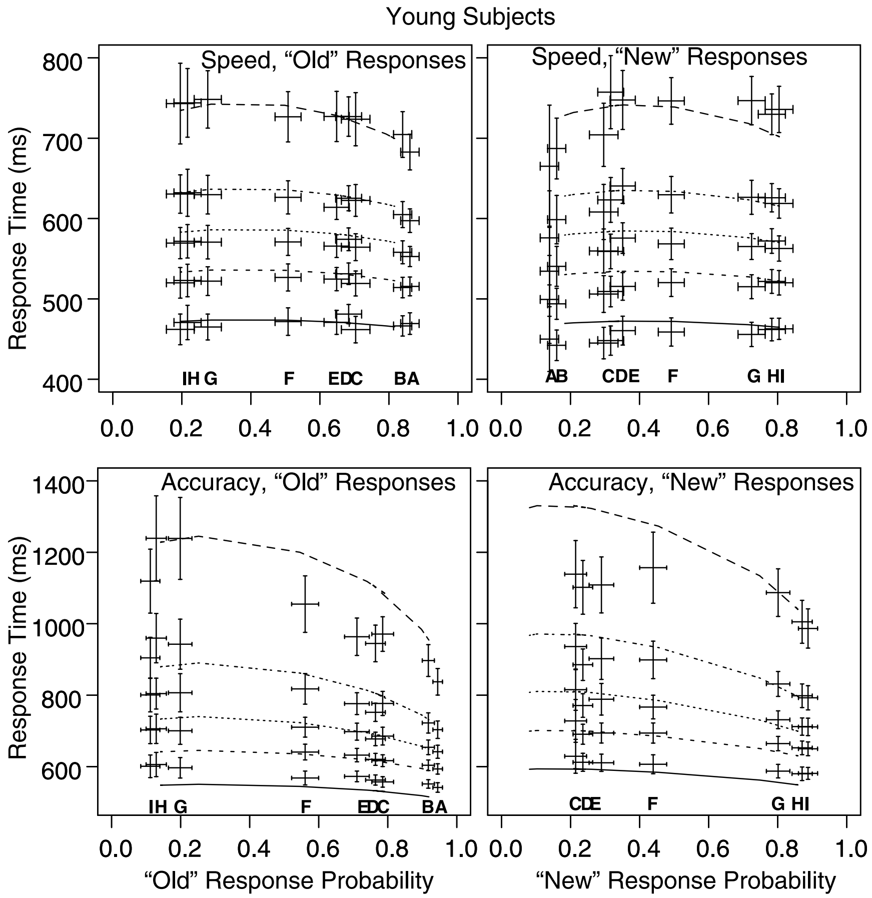

Fig. 3. Quantile-probability plots for young subjects. The lines represent the theoretical fits of the diffusion model and the centers of the confidence intervals are the empirical quantile response times. The quantile response times in order from the bottom to the top are the .1, .3, .5, .7, and .9 quantiles. The confidence intervals are plus or minus 2 *SE*s in the quantile response times and accuracy values across subjects. For "new" responses in the accuracy conditions, a moderate proportion of subjects do not have enough responses to allow computation of quantiles and so no quantiles are presented. The symbols A–I below the columns of quantiles label the experimental conditions. A, 3V; B, 3L; C, 3H; D, 1V; E, 1L; F, 1H; G, NH; H, NL; and I, NV; where H, high frequency words; L, low frequency words; V, very low frequency words; and 3, three presentations; 1, one presentation; and N, new words.

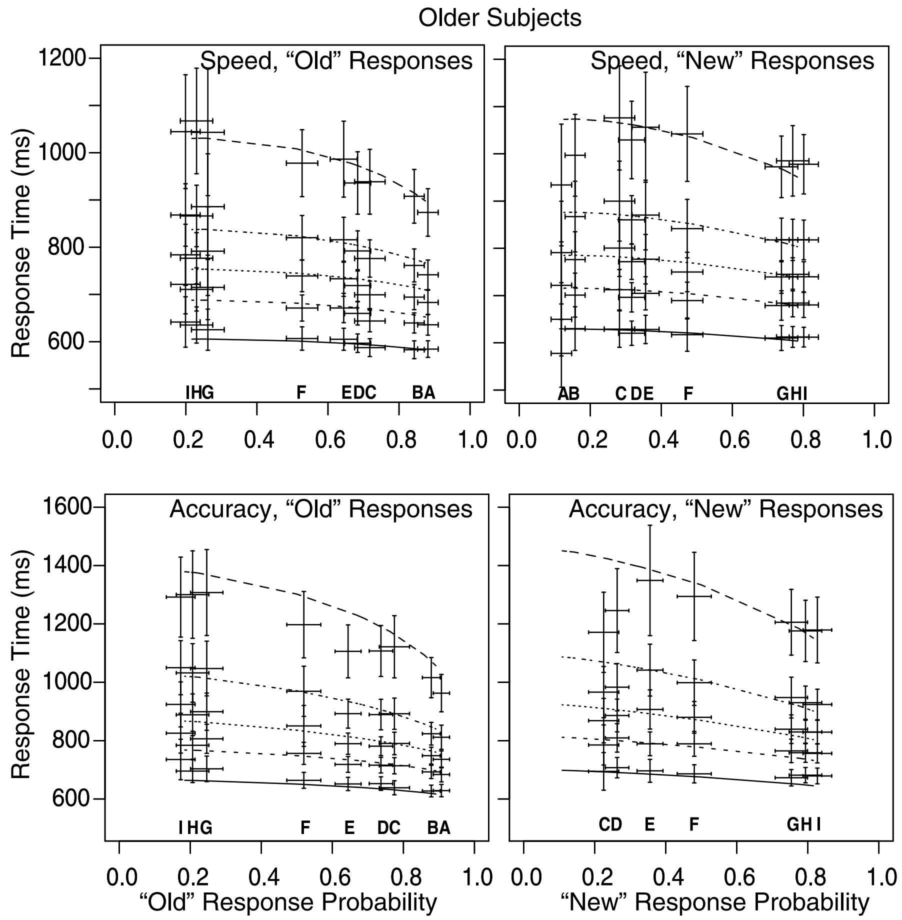

Fig. 4. Quantile probability plots for older subjects. The symbols mean the same as in Fig. 3.

word frequency for unstudied (new) words, and three values of word frequency for words studied once and for words studied three times. The left hand panels show data for test items for which the subjects responded ‘‘old,’’ and the right hand panels show responses for which the subjects responded ‘‘new.’’ The top panels show data for the conditions with speed instructions and the bottom panels show data for the conditions with accuracy instructions. Correct responses fall on the right hand sides of the functions and error responses on the left.

We computed standard errors on the quantile response times and on the probabilities of ‘‘old’’ and ‘‘new’’ responses across subjects and we used these to provide plus or minus 2 standard error bars around the data points. In Figs. 3 and 4, the data points are at the centers of the cross hairs formed by the error bars. The lines are the best fitting predicted functions from the diffusion model, which will be discussed later. Extreme errors, error responses with probability less than .1, for ‘‘new’’ responses for low and very low frequency words in the conditions with accuracy instructions had too few responses for many of the subjects to allow quantiles to be computed (at least 5 data points are needed per subject to produce the 5 response time quantiles), so they are not shown on the figures.

For the data for young subjects (Fig. 3), for speed instructions, median response times change by about 30 ms across the various values of accuracy. With accuracy instructions, the change in median response times across conditions is about 80 ms. Older subjects (Fig. 4) show similar trends with about a 50 ms change in median response time across conditions with speed instructions and about an 80 ms change across conditions with accuracy instructions.

Quantile probability functions provide a summary picture of the shapes of the response time distributions. The .1 quantile response times for correct responses change by less than 50 ms across the various levels of accuracy for both young and older subjects with both speed and accuracy instructions. In contrast, the .9

quantile response times change by up to several hundred milliseconds. With speed instructions, for both young and older subjects, the .9 quantile response times change by a little less than 100 ms across conditions. With accuracy instructions, the change in the .9 quantile response times for both young and older subjects is about 240 ms.

Error responses are generally slower than correct responses, except for correct "new" responses with speed instructions for young subjects where they are a little faster. Across all conditions, the change in median error response times is mainly reflected in the response time distribution spreading rather than shifting.

In sum, the results show the expected patterns for the effects of difficulty from word frequency and repetition manipulations, speed versus accuracy instructions, and age on performance. The question for the diffusion model is what components of processing are responsible for these effects.

The diffusion model

In the experiment presented in this article, subjects were instructed to respond as quickly as possible in some blocks of trials and to respond as accurately as possible in other blocks. Speed–accuracy tradeoffs are modeled by altering the boundaries of the decision process—wider boundaries require more information before a decision can be made and this leads to more accurate and slower responses. The manipulations of word frequency and number of repetitions were designed to vary accuracy from relatively high to relatively low; the effects of these variables are modeled by differences in drift rate, a different drift rate for each word frequency/repetition condition.

The diffusion model was fit to the data by minimizing a χ^2 value with a general SIMPLEX minimization routine that adjusts the parameters of the model to find the parameters that give the minimum χ^2 value (see Ratcliff & Tuerlinckx, 2002, for a full description of the methods). The data entered into the minimization routine for each experimental condition were the response times for each of the five quantiles for correct and error responses and the accuracy values. The quantile response times and the diffusion model were used to generate the predicted cumulative probability of a response by that quantile response time. Subtracting the cumulative probabilities for each successive quantile from the next higher quantile gives the proportion of responses between adjacent quantiles. For the χ^2 computation, these are the expected values, to be compared to the observed proportions of responses between the quantiles (multiplied by the number of observations). The observed proportions of responses between adjacent quantiles are the proportions of the distribution between adjacent quantiles (i.e., the proportions between 0, .1, .3, .5, .7, .9, and 1.0 are .1, .2, .2, .2, .2, and .1) and are multiplied by the number of observations. Summing over (observed − expected)^2/expected for all conditions gives a single χ^2 value to be minimized.

Research on fitting the diffusion model to data (Ratcliff & Tuerlinckx, 2002) has found that when long or short outlier response times are added to simulated data, the χ^2 method cannot accurately recover the parameter values that were used to generate the data. To address this problem, as noted earlier, short outliers were identified by examining the time at which accuracy begins to rise above chance (e.g., Swensson, 1972) and they were trimmed out (response times shorter than 300 ms were eliminated for both young and older subjects for the data reported here). Long outliers (responses longer than 3000 ms for young subjects and 4000 ms for older subjects) were also eliminated from the analyses. Ratcliff and Tuerlinckx showed that any remaining contaminant response times can be explicitly modeled. A parameter (p_0) is added to represent the probability of a contaminant in each condition of the experiment. The contaminant is assumed to come from a uniform distribution that has maximum and minimum values corresponding to the maximum and minimum response times in the condition. The psychological assumption behind this choice is that subjects are delayed by a random amount of time on some small proportion of trials, i.e., a momentary lapse of attention. For the data reported here, the value of the probability parameter, p_0, was the same across all experimental conditions (speed and accuracy and the different levels of difficulty) for each subject group. There may be better ways of estimating the range or distribution of contaminants, but the small proportion usually estimated (less than 2% here), the ease of implementation, the fact that this adds only one parameter to the model, and the ability to recover parameter values better than without this assumption (Ratcliff & Tuerlinckx, 2002) all indicate that it was a reasonable assumption.

For the fits of the model presented here, five parameters were held constant across the nine conditions and speed versus accuracy instructions: the mean duration of non-decision components of processing (T_{er}), across trial variability in non-decision components of processing (s_t), across trial variability in the range of the starting point of the diffusion process across trials (s_z), across trial variability in drift rate (η), and the probability of contaminant response times (p_0). Holding these five parameters constant reflects the assumption that neither speed versus accuracy instructions nor the quality of the information from the stimulus (frequency or number of repetitions) affects any of these components of the decision process. The speed–accuracy manipulation is assumed to affect only the separation of the boundaries in the diffusion process, and the drift rates are assumed to be constant across the speed and accuracy conditions, varying only with word frequency and number of repetitions. Changes in drift rate move points along the quantile probability function, but cannot alter the function in any other way. With these combinations of parameters, the model must account for accuracy rates, the relative speeds of correct and error responses, and the shapes of the response time distributions for correct and error responses. Specifically, with only boundary separation varying, the model must account for the small changes in accuracy and the large changes in response time between the speed and accuracy conditions. With only drift rate varying, the model must account for the changes in accuracy and distribution shape for both error and correct responses for both the speed and accuracy conditions as a function of word frequency and number of repetitions.

We fit the diffusion model to the data in two ways. First, each subject's data were fit individually and the parameter values averaged across subjects. The means for each of the parameters are shown in Tables 3 and 4 along with their standard deviations. Standard errors in the parameter values (the basis for significance tests additional to those presented below) can be found by dividing the standard deviations by the square root of the number of subjects (39 young subjects and 41 older subjects). The second fits of the model were to the data averaged across all the older subjects and averaged across all the young subjects. Specifically, accuracy values were averaged across subjects for each condition and each quantile response time was averaged across subjects for each condition. The latter fits were used as the basis for the predictions displayed in Figs. 3 and 4 (the lines). Group data have often been used in fitting models and the assumption (usually implicit) is that the fits and parameter values to the group data will turn out to be the same as the averages from the fits for the individual subjects. We provide both for comparison (see also, Ratcliff et al., 2003; Thapar et al., 2003). Note that the parameter values obtained from the group data and the average parameter values across individuals are either within 2 *SE*s or close to within 2 *SE*s of each other

Table 3
Parameter values and standard deviations from fits of the diffusion model

Group	a_s	z_s	a_a	z_a	T_{er}	η	s_z	p_0	s_t
Older (average parameters over subjects)	0.100	0.044	0.138	0.059	576	0.193	0.012	0.011	218
Older (fit to average data)	0.100	0.045	0.138	0.063	594	0.218	0.006	0.002	232
Young (average parameters over subjects)	0.074	0.037	0.138	0.062	471	0.168	0.013	0.008	219
Young (fit to average data)	0.069	0.035	0.132	0.058	486	0.164	0.011	0.016	203
Older standard deviations (over subjects)	0.027	0.010	0.043	0.013	45	0.057	0.006	0.007	66
Young standard deviations (over subjects)	0.014	0.009	0.029	0.013	34	0.054	0.008	0.005	56

Note. a_s, boundary separation for speed condition; a_a, boundary separation for accuracy condition; z_s, starting point for speed condition; z_a, starting point for accuracy condition; T_{er}, mean non-decision component of response time; η, standard deviation in drift across trials; s_z, range of the distribution of starting point (z); v, drift rates; p_0, proportion of contaminants; and s_t, range of the distribution of non-decision times.

Table 4
Drift rate parameter values and standard deviations from fits of the diffusion model

Group	v_{NH}	v_{NL}	v_{NV}	v_{1H}	v_{1L}	v_{1V}	v_{3H}	v_{3L}	v_{3V}
Older (average parameters over subjects)	−0.186	−0.224	−0.220	0.041	0.081	0.113	0.152	0.259	0.290
Older (fit to average data)	−0.206	−0.215	−0.243	−0.015	0.095	0.109	0.166	0.256	0.297
Young (average parameters over subjects)	−0.177	−0.229	−0.249	0.050	0.104	0.141	0.187	0.253	0.289
Young (fit to average data)	−0.163	−0.252	−0.246	0.003	0.114	0.150	0.125	0.246	0.274
Older standard deviations (over subjects)	0.102	0.129	0.103	0.073	0.050	0.082	0.109	0.093	0.101
Young standard deviations (over subjects)	0.067	0.074	0.090	0.050	0.034	0.045	0.094	0.088	0.094

Note. The subscripts of the drift rates (v) are: N, new words; 1, words presented once; 3, words presented three times; H, high frequency words; L, low frequency words; and V, very low frequency words.

(see Tables 3 and 4; standard errors and z-scores can be computed from the standard deviations and the numbers of subjects in the two groups). Also, the parameter values are in the range of parameter values from other experiments (Ratcliff, 2002; Ratcliff & Rouder, 1998, 2000; Ratcliff et al., 2001, 2003; Ratcliff et al., 1999; Thapar et al., 2003). The standard deviations across subjects are larger for older subjects than for young subjects. In part, this is because several of the parameter values are larger for older versus young subjects and this leads to increased variability in the estimates of all parameters (see Ratcliff & Tuerlinckx, 2002).

Goodness of fit

In general, the model captures the changes in response time distributions for both correct and error responses and accuracy as a function of word frequency and number of repetitions with only drift rate changing. The response time quantiles lie on the same functions for word frequency and repetitions, the function lines representing the predictions drawn in Figs. 3 and 4. This means that the effects of word frequency and repetition can be modeled by changes in a single variable, drift rate. A pattern of data that could not be handled by changes in drift rate alone would be one in which there are, for example, two separate sets of quantile functions, one for items presented once as a function of word frequency and one for items presented three times as a function of word frequency (e.g, the quantile function for items presented once might be vertically displaced relative to the quantile-probability function for items presented three times).

The model also captures the effects of speed and accuracy instructions with only boundary separation changing. The only systematic misses are in some of the .1 and .9 quantile response times with accuracy instructions for young subjects, and there are only two or three significant misses for older subjects across all conditions. For the .9 quantiles for young subjects, the response times are quite large and it may be that subjects are terminating many of the longest processes early, i.e., not allowing them to run to completion. For the .1 quantile response times, within the framework of the diffusion model, the misses possibly indicate that the value of the non-decision component of response time (T_{er}) is changing by 20 ms between speed and accuracy conditions. However, the misses are modest and only about 26 out of 340 predictions fall outside the confidence intervals (there are 340 data points shown instead of 360 because two extreme error conditions for "new" responses to old stimuli with accuracy instructions did not have enough observations per subject to produce quantiles).

Fits for 15 out of the 41 older subjects and 11 out of the 39 young subjects produced significant χ^2 values. The χ^2 statistic had a critical value of 424 with 378 degrees of freedom. This number of misses means that there are more significant misses than would be expected by chance. But the power of the χ^2 statistic grows with the number of observations (because it is based on frequencies; $(O-E)^2/E$ grows as N grows) so that even a small difference in the proportions of observed and expected frequencies (e.g., .01) will become significant as N grows. There are a number of reasons having to do with the stability of data that suggest that we should not expect perfect fits. For example, if subjects changed some aspect of their performance systematically across trials, such as altering decision criteria or changing guessing strategies, or they became systematically more fatigued across trials, then variability would be introduced into the data beyond that accounted for by the model. The presented fits show that the model captures most of the trends in the data without highly systematic misses. If other models were equally capable of explaining the data, then the next step would be competitive model testing (see Ratcliff & Smith, in press; Van Zandt, Colonius, & Proctor, 2000).

Differences in parameter values across groups

Analysis of the parameter estimates using t tests and ANOVA with a .05 significance level showed that older subjects differed from the young subjects in several ways. First, the value of T_{er} was larger for older subjects than young subjects by about 100 ms ($t(75.80) = 12.00$; all t tests used the Welch correction for degrees of freedom). This result replicates results in Ratcliff et al. (2001, 2003) and Thapar et al. (2003), though with a larger sized difference between young and older subjects. Second, the older subjects did not have significantly lower drift rates than the young subjects. Drift rates were submitted to a two-way ANOVA, with drift rates for new items, which had the opposite sign to those for old items, converted to the same sign. Drift rates differed significantly among conditions, as would be expected ($F(8, 693) = 86.34$, $MSE = 0.0208$), but although they were slightly lower for older subjects, the difference between young and old was not significant ($F(1, 78) = 0.94$, $MSE = 0.0060$). Third, boundary separation was greater for older subjects than young subjects with speed instructions ($t(62.76) = 5.38$) but not with accuracy instructions ($t(72.19) = -0.102$), replicating previous findings (Ratcliff et al., 2001, 2003; Thapar et al., 2003) that older subjects adopt more conservative decision criteria. Finally, the standard deviation in drift rate was just significantly larger for older subjects than young subjects ($t(80) = 2.07$). There was no difference in the range of the distributions of starting points, ($t(71.13) = -0.051$), the range of the distributions of non-decision components of processing ($t(79.00) = -.092$), or the proportions of contaminants ($t(69.43) = 1.90$).

Although accuracy was higher by about 5% for young than older subjects with accuracy instructions (averaged over old and new test words), the drift rates were not significantly different nor was boundary separation. The 5% difference in accuracy comes about because of the difference between old and young in the standard deviations of the drift rates across trials. A change in η of .05 (keeping the other parameters fixed at the values in Tables 3 and 4) gives rise to about a 4% change in accuracy. The non-significant .01 difference in drift rate gives rise to a further 1% difference in accuracy. So although parameter values are very similar for older and young subjects in the accuracy conditions, the observed differences in accuracy are accounted for by the relatively small differences in drift rates and variability in drift rates across trials.

Individual differences

Figs. 5 and 6 display histograms of the distributions of the parameter values from the individual subject fits for the young and older subjects, like the ones in Ratcliff et al. (2001) and Thapar et al. (2003). Only a few of the histograms show serious deviations from symmetry and these are the only ones we discuss. First, the distribution of the values of s_z, the range of starting points, is right skewed with mostly small but a few large values. The few large values appear to occur because they are associated with large values of boundary separation. Second, the distribution of s_t values (the range of the non-decision component) is relatively flat with a couple of small values. Third, for older subjects, boundary separation a (and hence starting point z because z and a are highly correlated) is right skewed because some of the subjects adopted extremely conservative decision criteria. This was not observed for the young subjects. However, there is considerable overlap in the range of boundary separations for older and young subjects with a few older subjects adopting much more conservative decision criteria than the young subjects.

The main result from Figs. 5 and 6 is that there is nothing surprising about the distributions of parameter values across individuals; most of the distributions are quite symmetric. The most straightforward interpretation of this is that each of the components of processing represented by the parameter values (apart from boundary separation for older subjects as noted above) is selected from a roughly symmetric distribution.

Correlations among parameter values

Relationships among the overall levels of the dependent variables (accuracy, correct response times, and error response times) and the main parameter values for interpreting aging effects, namely, drift rate, boundary separation, and the non-decision components of processing, are shown in Figs. 7 and 8. To obtain values for mean response time and accuracy, means were computed

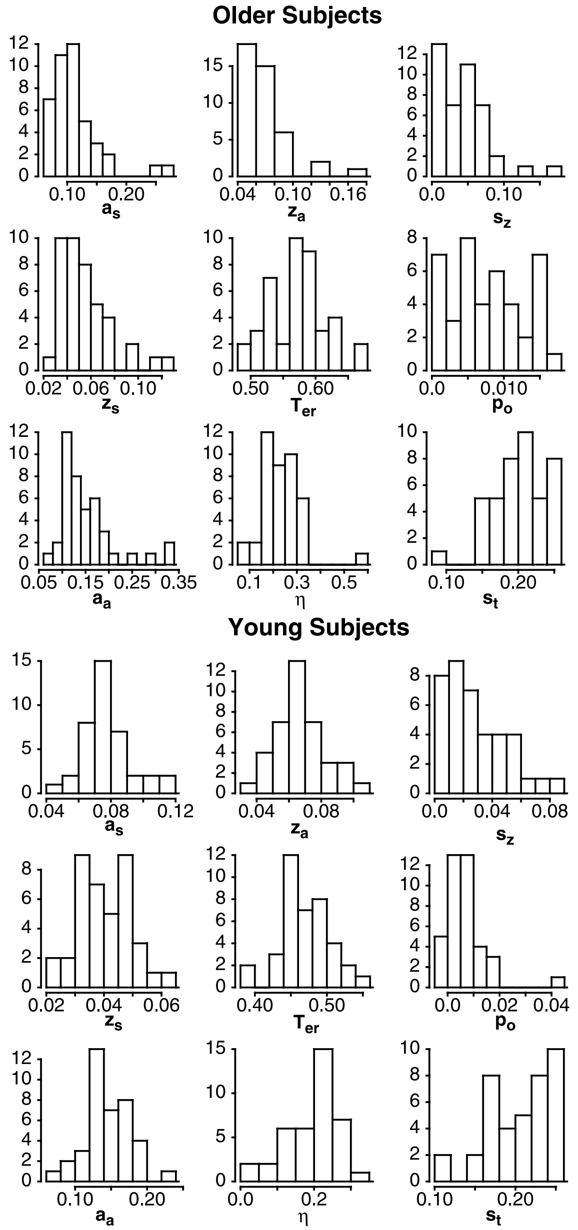

Fig. 5. Histograms for the parameter values for fits of the diffusion model to individual subjects for the experiment: a_s, boundary separation for speed conditions; a_a, boundary separation for accuracy conditions; z_s, starting point for speed conditions; z_a, starting point for accuracy conditions; T_{er}, non-decision component of response times; η, standard deviation in drift across trials; s_z, range of the distribution of starting point (z) across trials; p_0, proportion of contaminants; and s_t, range of the distribution of non-decision times across trials.

over all old and new responses, for all the conditions for number of repetitions, word frequency, and speed–accuracy instructions. The values of boundary separation were

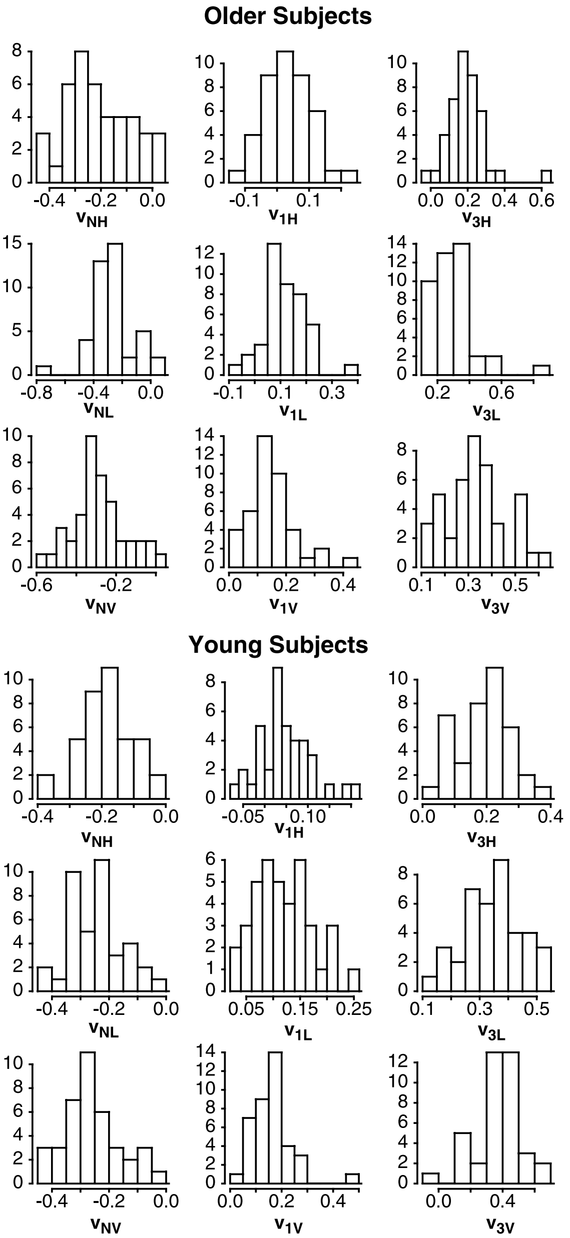

Fig. 6. Histograms for the drift rate (v) parameter values across older and young subjects for fits of the diffusion model for the experiment. The subscripts of the drift rates are: N, new words; 1, words presented once; 3, words presented three times; H, high frequency words; L, low frequency words; and V, very low frequency words.

the average of the boundary separations for speed and accuracy instructions and the drift rates were the averages over the nine word conditions (where the drift rates for new items had their signs reversed). Figs. 7 and 8 show scatter plots of these quantities against each other along with the values of their correlations.

The key results from Figs. 7 and 8 are: First, the overall levels of accuracy and mean response time are uncorrelated. Second, correct and error response times are strongly correlated. Third, boundary separation and drift rate are uncorrelated. Fourth, boundary separation and mean response time (both for correct and error responses) are strongly correlated. Fifth, accuracy and drift rate are strongly correlated. Sixth, the parameter for the non-decision components of processing, T_{er}, is not strongly correlated with any of the other quantities. These results indicate that the overall level of response time was determined by the boundary separation that subjects adopted. More conservative subjects produced overall longer response times and less conservative subjects produced overall shorter response times. Overall response time was not determined by drift rate; that is, faster subjects did not have higher drift rates than slower subjects. However, the overall level of accuracy was determined by drift rate, that is, by whether recognition memory was good or bad. The overall level of accuracy was not a function of boundary separation. That is, the degree of conservativeness did not affect overall accuracy because changes in boundary separation had only modest effects on accuracy (see Table 2). In addition, older and young subjects showed very similar trends. Other correlations not shown (e.g., speed and accuracy boundary separation, drift rates, and so on) are similar to those reported in Thapar et al. (2003).

It is important to note that these correlations concern individual differences in the overall levels of performance. Even though overall accuracy and overall response time are not correlated across subjects, within a subject and across conditions, changes in drift rate, for example, have strong and reliable effects on both accuracy and response time, as can be seen in Figs. 3 and 4.

General discussion

The data from the experiment presented here illustrate the puzzle identified in the introduction: the older subjects showed a large deficit in response times but only a small deficit in accuracy. The older subjects were slower than the young subjects by about 100 ms with accuracy instructions and about 160 ms with speed instructions, but they were only about 6% less accurate with accuracy instructions and equally accurate with speed instructions. The response time data suggest a theoretical view in which aging has a relatively large effect on recognition memory, but the accuracy data suggest a view in which it does not (e.g., Balota et al., 2000; Salthouse, 1996).

The diffusion model allows parameter values for the model's components of processing to be extracted from the joint accuracy and response time data. Application of the model identified both similarities and differences between the older and young subjects. First, the quality

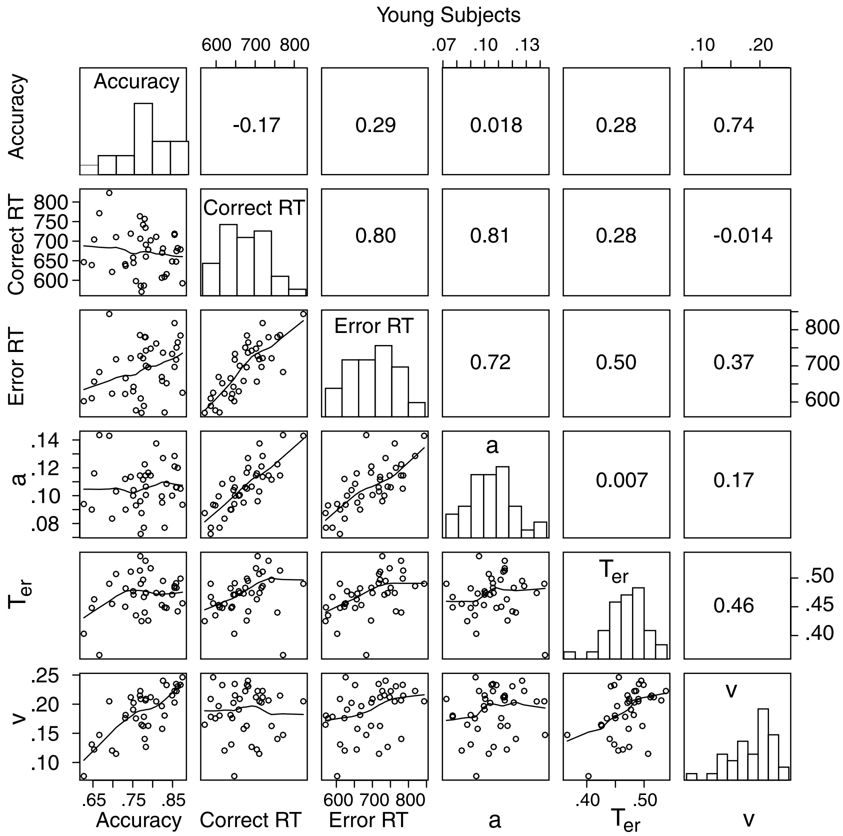

Fig. 7. Scatter plots for young subjects for accuracy, mean response time for correct responses, mean error response time, average boundary separation (a), non-decision component of processing (T_{er}), and average drift rate (v) are shown in the bottom left panels. Correlations corresponding to the scatter plots are in the top right hand half of the figure. Histograms of the quantities are presented on the diagonals. The lines are smoothed lines based on medians (Becker, Chambers, & Wilks, 1988).

of the information from memory (drift rate in the model) was not significantly different between the two groups of subjects. The older subjects' memory for studied items was not significantly worse than the young subjects'. The increases in response time for the older subjects came from two sources: the older subjects were slower in the non-decision components of processing by about 100 ms, and they adopted more conservative decision criteria (with speed but not accuracy instructions).

Thus, the findings reported here are consistent with previous conclusions based on accuracy measures. Craik and McDowd (1987), for example, found no difference in recognition performance, and reviews of the literature (Balota et al., 2000; Craik, 1994; Neath, 1998, Chap. 16) have concluded that the effects of aging on recognition are modest or absent.

In the diffusion model framework, the large differences in response times between the older and young subjects arise from differences in non-decision components of processing and differences in criteria settings. Previously, the most common way of addressing the effects of aging on speed of processing has been to plot older subjects' mean response times for each condition in an experiment against young subjects' means (a Brinley plot) and use the slope of the resulting function as a measure of the factor by which the older subjects are slowed relative to the young subjects. In the recognition memory experiment presented here, if this plot were constructed for the data with speed instructions, the conclusion would be that the older subjects were slowed by a factor of about 2; if the plot were constructed for the data with accuracy instructions, the conclusion would be that the older subjects were not slowed with respect to the young subjects; if the speed and accuracy data were combined (ignoring the misses in Fig. 2 to the joint straight line), the conclusion would be that the young subjects had a deficit relative to the older subjects. The advantage of analyzing the data with the diffusion model is that the separate contributions to overall response time of the decision criteria, the non-decision components of processing, and the rates of accumulation of evidence from the stimuli can be extracted from accuracy values and response time distributions for

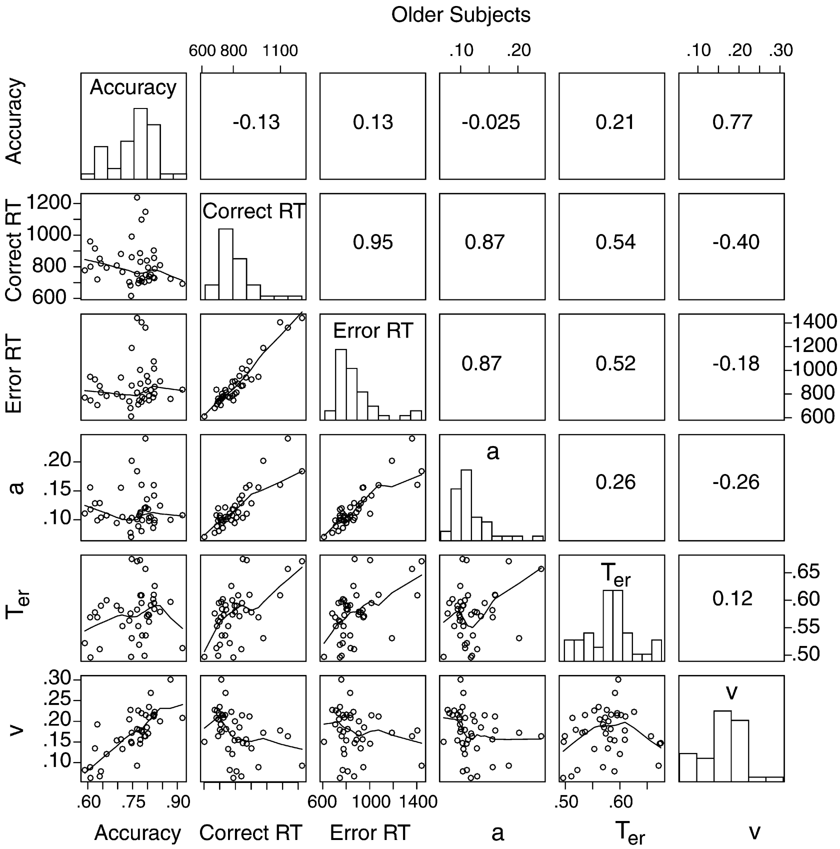

Fig. 8. Scatter plots for older subjects for the same information as in Fig. 7.

correct and error responses. Once the decision criteria and non-decision components have been factored out, the rate of accumulation of evidence can be used as a measure of the quality of information entering the decision process from memory.

The results obtained in the experiment reported here are similar to those obtained in a signal detection task (Ratcliff et al., 2001), in a brightness discrimination task (Ratcliff et al., 2003), and in a letter discrimination task with masking (Thapar et al., 2003). Older subjects were slower than young subjects by about 40–80 ms in the non-decision components of response time. With speed instructions, they adopted more conservative decision criteria in all the experiments. With accuracy instructions, they adopted more conservative decision criteria in the signal detection and letter discrimination experiments, but they were about equal in the brightness discrimination experiment, as they were in the experiment reported here. In all of the experiments except masked letter discrimination, older subjects had about the same drift rates as young subjects. The difference between the two perceptual tasks, letter discrimination and brightness discrimination, is consistent with the literature on letter identification using accuracy and threshold measures (Spear, 1993). In that literature, there is a deficit as a function of age for high spatial frequency stimuli (e.g., letters), but no deficit with low spatial frequency stimuli (e.g., brightness patches).

As the diffusion model's theoretical analyses are brought to various tasks, the rates of extraction of information from stimuli are decoupled from criterion effects and from non-decision components of processing. Instead of a monolithic account of processing speed in terms of only mean correct response times, we have instead an account based on all aspects of the data that allows the quality of information extracted from stimuli to be separated from subject-adjustable decision criteria. The data and analyses from the study reported here and from similar studies provide a growing body of support for the diffusion model in particular and quantitative modeling approaches in general.

References

Allen, P. A., Ashcraft, M. H., & Weber, T. A. (1992). On mental multiplication and age. *Psychology and Aging, 7*, 536–545.

Allen, P. A., Madden, D. J., Weber, T. A., & Groth, K. E. (1993). Influence of age and processing stage on visual word recognition. *Psychology and Aging, 8*, 274–282.

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

Becker, R. A., Chambers, J. M., & Wilks, A. R. (1988). The new S language. Pacific Grove, CA: Wadsworth & Brooks/Cole: Pacific Grove.

Birren, J. E. (1965). Age-changes in speed of behavior: Its central nature and physiological correlates. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging, and the nervous system* (pp. 191–216). Springfield, IL: Thomas.

Bowles, N. L., & Poon, L. W. (1982). An analysis of the effect of aging on recognition memory. *Journal of Gerontology, 37*, 212–219.

Brinley, J. F. (1965). Cognitive sets, speed and accuracy of performance in the elderly. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging and the nervous system* (pp. 114–149). Springfield, IL: Thomas.

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

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin, 98*, 67–83.

Cerella, J. (1990). Aging and information-processing rate. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 201–221). San Diego, CA: Academic Press.

Cerella, J. (1994). Generalized slowing in Brinley plots. *Journal of Gerontology: Psychological Sciences, 49*, 65–71.

Cerella, J., Poon, L. W., & Williams, D. M. (1980). Age and the complexity hypothesis. In L. W. Poon (Ed.), *Aging in the 1980s: Psychological issues* (pp. 332–340). Washington, DC: American Psychological Association.

Coyne, A. C. (1981). Age difference and practice in forward visual masking. *Journal of Gerontology, 36*, 730–732.

Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science, 3*, 155–158.

Craik, F. I. M., & Jennings, J. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51–110). Hillsdale, NJ: Lawrence Erlbaum Associates.

Craik, F. I. M., & McDowd, J. M. (1987). Recall, recognition, and aging. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 474–479.

Erber, J. T. (1974). Age differences in recognition memory. *Journal of Gerontology, 29*, 177–181.

Fisher, D. L., & Glaser, R. A. (1996). Molar and latent models of cognitive slowing: Implications for aging, dementia, depression, development, and intelligence. *Psychonomic Bulletin & Review, 3*, 458–480.

Fisk, A. D., & Fisher, D. L. (1994). Brinley plots and theories of aging: The explicit, muddled, and implicit debates. *Journal of Gerontology: Psychological Sciences, 49*, 81–89.

Fisk, J. E., & Warr, P. (1996). Age and working memory: The role of perceptual speed, the central executive, and the phonological loop. *Psychology and Aging, 11*, 316–323.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198.

Fozard, J. L. (1990). Vision and hearing in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 150–170). San Diego, CA: Academic Press.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1–67.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition, 13*, 8–20.

Gordon, S. K., & Clark, W. C. (1974). Adult age differences in word and nonsense syllable recognition memory and response criterion. *Journal of Gerontology, 29*, 659–665.

Hartley, A. A. (1992). Attention. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 3–49). Hillsdale, NJ: Erlbaum.

Hertzog, C. (1992). Aging, information processing, and intelligence. In K. W. Schaie (Ed.), *Annual review of gerontology and geriatrics* (Vol. 11, pp. 55–79). New York: Springer.

Hintzman, D. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*, 411–428.

Kausler, D. H. (1994). *Learning and memory in normal aging*. San Diego, CA: Academic Press.

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

Lima, S. D., Hale, S., & Myerson, J. (1991). How general is general slowing? Evidence from the lexical domain. *Psychology and Aging, 6*, 416–425.

Luce, R. D. (1986). *Response times*. New York: Wiley.

Madden, D. J. (1989). Visual word identification and age-related slowing. *Cognitive Development, 4*, 1–29.

Madden, D. J., Pierce, T. W., & Allen, P. A. (1992). Adult age differences in attentional allocation during memory search. *Psychology and Aging, 7*, 594–601.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A Bayesian approach to the effects of experience in recognition memory. *Psychological Review, 105*, 724–760.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89*, 609–626.

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

Myerson, J., Wagstaff, D., & Hale, S. (1994). Brinley plots, explained variance, and the analysis of age differences in response latencies. *Journal of Gerontology: Psychological Sciences, 49*, 72–80.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1170–1187.

Neath, I. (1998). Developmental changes in memory. In *Human memory: An introduction to research, data, and theory* (pp. 350–375). Pacific Grove, CA: Brooks/Cole Publishing Company.

Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity through adulthood. *Vision Research, 23*, 689–699.

Perfect, T. J. (1994). What can Brinley plots tell us about cognitive aging? *Journal of Gerontology: Psychological Sciences, 49*, 60–64.

Rabinowitz, J. C. (1984). Aging and recognition failure. *Journal of Gerontology, 39*, 65–71.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.

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

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

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

Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling the accumulation of partial information. *Psychological Review, 95*, 238–255.

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 and Review, 9*, 278–291.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). Diffusion model account of lexical decision. *Psychological Review* (in press).

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

Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 127–140.

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

Ratcliff, R., Spieler, D., & McKoon, G. (2000). Explicitly modeling the effects of aging on response time. *Psychonomic Bulletin and Review, 7*, 1–25.

Ratcliff, R., Spieler, D., & McKoon, G. (2004). Analysis of group differences in processing speed: Where are the models of processing? *Psychonomic Bulletin and Review* (in press).

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 and Psychophysics, 65*, 523–535.

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

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

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

Salthouse, T. A. (1985). *A theory of cognitive aging*. Amsterdam: North-Holland.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403–428.

Salthouse, T. A., Kausler, D. H., & Saults, J. S. (1988). Utilization of path-analytic procedures to investigate the role of processing resources in cognitive aging. *Psychology and Aging, 3*, 158–166.

Schonfield, A. E. D., & Robertson, B. A. (1966). Memory storage and aging. *Canadian Journal of Psychology, 20*, 228–236.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM: Retrieving effectively from memory. *Psychonomic Bulletin and Review, 4*, 145–166.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review, 102*, 567–591.

Spear, P. D. (1993). Neural bases of visual deficits during aging. *Vision Research, 33*, 2589–2609.

Swensson, R. G. (1972). The elusive tradeoff: Speed versus accuracy in visual discrimination tasks. *Perception and Psychophysics, 12*, 16–32.

Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging, 18*, 415–429.

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

Wixted, J. T. (1992). Subjective memorability and the mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 681–690.