Aging and individual differences in rapid two-choice decisions

ROGER RATCLIFF
Ohio State University, Columbus, Ohio

ANJALI THAPAR
Bryn Mawr College, Bryn Mawr, Pennsylvania

and

GAIL McKOON
Ohio State University, Columbus, Ohio

The effects of aging on performance were examined in signal detection, letter discrimination, brightness discrimination, and recognition memory, with each subject tested on all four tasks. Ratcliff's (1978) diffusion model was fit to the data for each subject for each task, and it provided a good account of accuracy and the distributions of correct and error response times. The model's analysis of the components of processing showed that aging had three main effects: The nondecision components of processing were slower and the decision criteria were more conservative for 60- to 74-year-old and 75- to 85-year-old subjects than for college students, but the quality of the evidence on which decisions were based was as good for the older subjects as for college students on some of the tasks. Individual differences among subjects in components of processing tended to be preserved across the tasks, as was shown by strong correlations across the tasks in the parameters of the model that represent the components of processing. For example, if the evidence on which a subject's decisions were based was good in one task, it tended to be good in all four tasks.

Across a wide variety of tasks, research has shown that cognitive processing slows with age. Until recently, the most prominent account of this decline was the generalized-slowing hypothesis, according to which either all cognitive processes slow with age or some general mechanism that contributes to many processes slows with age (see, e.g., Brinley, 1965; Cerella, 1994). Although that hypothesis has been favored by many researchers because it provides a relatively simple and intuitively appealing explanation of age-related decrements across many laboratory tasks and everyday behaviors, it has also been challenged on a number of fronts (Cerella, 1994; Fisher & Glaser, 1996; Fisk & Fisher, 1994; Hertzog, 1992; Myerson, Wagstaff, & Hale, 1994). One central problem is that the hypothesis does not allow the decision components of a task to be separated from the quality of the information upon which the decisions are based. Another is that the hypothesis has been focused on mean response times (RTs) for correct responses. The full set of data that need to be addressed also includes accuracy and correct and error RT distributions (Ratcliff, Spieler, & McKoon, 2000). Recently, it has been shown that Ratcliff's diffusion model (Ratcliff, 1978, 1988, 2002; Ratcliff & Rouder, 1998, 2000; Ratcliff &

Smith, 2004; Ratcliff, Van Zandt, & McKoon, 1999) can handle the full set of data from two-choice tasks and also provides an analysis of the components of processing that underlie performance in a task and how they are differentially affected by aging (for more detailed background, see Ratcliff, Thapar, & McKoon, 2001, 2003, 2004, and Thapar, Ratcliff, & McKoon, 2003, hereafter referred to collectively as "RTM"). For this article, we tested the diffusion model in the aging domain by examining performance of individual subjects across four two-choice tasks, with college-age subjects, 60- to 74-year-olds, and 75- to 85-year-olds. The goal was an analysis of individual differences, and the key question was the extent to which components of processing are consistent for an individual across tasks.

The diffusion model assumes that fast two-choice 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 Figure 1B, where the starting point is labeled z and the boundaries are labeled 0 and a. When one of the boundaries is reached, a response is initiated. The rate of accumulation of information is called the drift rate (v), and it is determined by the quality of the information available from the stimulus. The better the information quality, the larger the drift rate toward the appropriate decision boundary, and the faster and more accurate the response. Within-trials variability in the accumulation of information results in processes with the same mean drift rate terminating at different times (producing
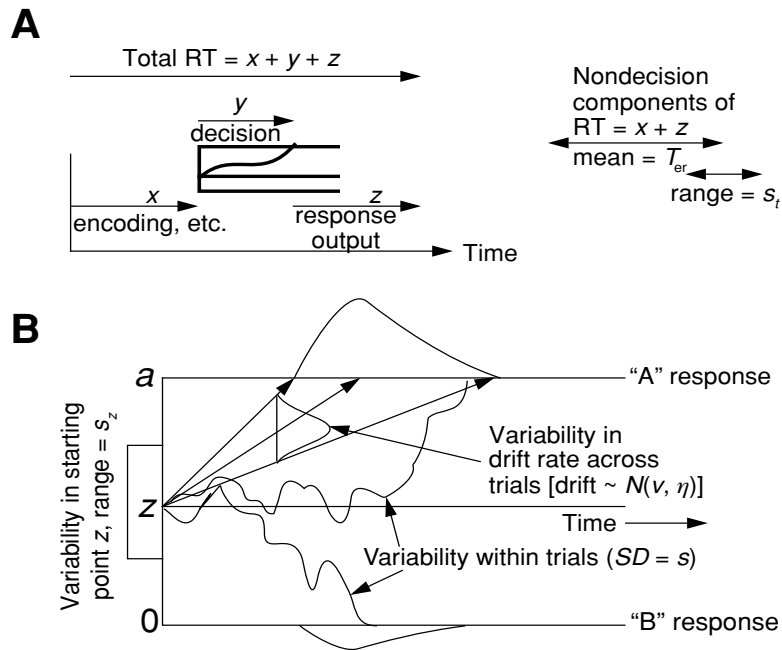
Figure 1. An illustration of the diffusion model. Parameters of the model are a = boundary separation (one value, a_{sp}, for speed instructions and another, a_{ac}, for accuracy instructions); z = starting point; T_{er} = mean time for nondecision components of RT; η = SD in drift across trials; s_z = range of the distribution of starting points across trials; v = drift rate (one for each level of difficulty in an experiment); s_t = range of the distribution of nondecision times across trials; s = SD of variability in drift within trials. s is a scaling parameter, set to 0.1 as in previously published fits of the model.

RT distributions) and sometimes at different boundaries (producing errors). Speed–accuracy trade-offs are modulated by the positions of the boundaries: Moving boundaries closer to the starting point speeds responses and decreases accuracy. RT distributions in two-choice tasks are positively skewed, which comes about naturally in the model by simple geometry: The increase in RTs is larger if a lower value of drift rate is decreased by some amount than if a larger drift rate is decreased by the same amount. Besides the decision process, there are nondecision components of processing, such as encoding and response execution (Figure 1A). These processes are combined in the model, and their contribution to RT has the mean T_{er} (Ratcliff & Tuerlinckx, 2002).

In the diffusion model, the components of processing are assumed to vary from trial to trial. Variability in drift rate across trials (normally distributed with SD η) gives rise to error responses that are relatively slow compared to correct responses, and variability in starting points across trials (uniformly distributed with range s_z) gives rise to relatively fast errors. Whether errors are faster or slower than correct responses for an experimental condition depends on the relative amounts of drift rate and starting point variability, drift rate values, and boundary positions (Ratcliff et al., 1999). Across-trials variability

in T_{er} is uniformly distributed with range s_t (Ratcliff & Tuerlinckx, 2002).

The diffusion model serves to map performance characteristics onto underlying processes. From the probability of a correct response and the RT distributions for correct and error responses for each of the conditions in an experiment, the model extracts estimates of the quality of the stimulus information that enters the decision process for each condition (drift rate), the amount of information that must be accumulated before a decision can be made (boundary positions), the time taken by nondecision components of RT (T_{er}), and the amount of variability across trials in each of the processing components. In its initial applications to aging research, the contribution of the diffusion model has been to identify which components are responsible for performance decrements in two-choice tasks for 60- to 74-year-old subjects relative to college-age subjects. In the five tasks that they examined, RTM found that older subjects were usually slower than college subjects in the nondecision components of processing that are summarized by T_{er}, and that older subjects usually set their response boundaries more conservatively than college subjects. One or both of these factors accounted for most of the difference in speed between the college and older subjects in all four experiments. Across-trials vari-

ability in components of processing was generally not significantly different between the two groups of subjects. The most interesting result was that drift rates did not differ significantly between the college and older subjects for four tasks: a signal detection task, masked brightness discrimination, recognition memory, and lexical decision (RTM; Ratcliff, Thapar, Gomez, & McKoon, 2004). Only for masked letter discrimination were the drift rates for the older subjects lower, in accord with psychophysical research that has found decrements with age for stimuli with higher spatial frequency (Spear, 1993). In all of the experiments just described, different subjects performed each task. In this article, however, the subjects were the same for each task, and they included 75- to 85-year-olds as well as 60- to 74-year-olds. The goals were to test the application of the diffusion model with a broader range of subjects and to examine whether the components of processing identified by the model are significantly correlated from one task to another for individual subjects.

THE EXPERIMENTS

The first experiment used a signal detection task in which subjects were asked to decide whether the number of asterisks displayed on a computer screen was "large" or "small." The other three experiments' tasks were masked letter discrimination, masked brightness discrimination, and recognition memory. Following RTM, we expected response boundaries and T_{er} to vary with age in all four experiments. For the signal detection task, there are no perceptual or memory limits on the information available to the subjects, and so, as in Ratcliff et al. (2001), we expected drift rates not to vary with age. For the other tasks, RTM found for 60- to 74-year-olds that drift rates were lower relative to the college students only in masked letter discrimination. For 75- to 85-year-olds, drift rates might also be lower in masked brightness discrimination and recognition memory.

The 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 (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. There were no significant differences on any of these measures for the three groups of subjects: 10 college students from Northwestern University, 10 subjects from 60 to 74 years old, and 10 subjects from 75 to 85 years old.

Each subject participated in 16 sessions, 4 sessions on each of the tasks, in the order signal detection, letter discrimination, brightness discrimination, and recognition memory. For each experiment, on alternating blocks of trials, instructions stressed that responses be either as accurate as possible or as fast as possible. The subjects were given feedback appropriate to the instructions, either accuracy feedback on each trial or a "too slow" message when an RT was over 700 msec.

Signal Detection

For each trial, a number of asterisks between 1 and 100 was generated from a signal distribution, normal with mean 57.5, or a noise distribution, normal with mean 39.5, each with an *SD* of 14.4 (Ratcliff et al., 1999; Ratcliff et al., 2001, Experiment 1). The asterisks were placed in random positions in a 10×10 array of blank characters on a computer screen. The subjects were asked to decide whether the number of displayed asterisks was "large" or "small." Accuracy feedback was given on all trials: If the number of asterisks was very large or very small, feedback indicated that "large" or "small," respectively, was the correct response. For intermediate numbers of asterisks, feedback was probabilistic, sometimes indicating "large" and sometimes "small" as the correct response (Ratcliff et al., 2001). There were 12 blocks of 96 trials per session. For the data analyses, the numbers of asterisks were grouped into eight experimental conditions so that the mean RTs and accuracy values were about the same for the stimuli within a group.

Letter Discrimination

For each block of trials, there were two target letters continuously displayed in the top left and right corners of the computer screen. On each trial, one of the letters was displayed at the center of the screen for 10, 20, 30, 40, 50, or 60 msec and then masked, as in Thapar et al. (2003). A subject's task was to indicate which letter was presented. There were 12 blocks of 96 trials per session. Performance for the 40-, 50-, and 60-msec durations was near ceiling, so data from these conditions were combined into one condition for data analyses.

Brightness Discrimination

The stimuli were 64×64 squares of black and white pixels displayed on a gray background of 320×200 pixels (Ratcliff et al., 2003). There were six levels of brightness for the squares, achieved with six values of the probability of a pixel being white (.350, .425, .475, .525, .575, and .650). A square was displayed for 50, 100, or 150 msec, followed by a mask made up of four 64×64 checkerboard patterns presented sequentially, and the subjects were asked to decide whether each square was "bright" or "dark." There were eight blocks of 144 trials per session.

Recognition Memory

The stimuli were high-, low-, and very-low-frequency words (Ratcliff, Thapar, & McKoon, 2004) in 20 study–test blocks per session. For each block, the study list consisted of words displayed for 1 sec each, 9 presented once and 9 presented three times (3 high, 3 low, and 3 very low frequency in each case), and the immediately following test list consisted of the 18 studied words plus 18 new

words (6 high, 6 low, and 6 very low frequency). For each session, stimuli were chosen randomly without replacement from the three pools.

RESULTS

All subjects adjusted their performance according to instructions, but the college subjects were more willing to sacrifice accuracy for speed; they made 4%–8% more errors with speed than with accuracy instructions, whereas the older subjects made only 1%–2% more errors (Figure 2). Otherwise, the main results were that accuracy rates for the 60- to 74-year-olds were lower than for the college subjects for letter discrimination but not for the other tasks, whereas accuracy rates for the 75- to 85-year-olds were also lower for brightness discrimination and a little lower for recognition memory.

For all subjects and experiments, RTs were consistently longer, by 100–300 msec, with accuracy than with speed instructions (Figure 2). With speed instructions, RTs increased by 60–150 msec from college subjects to the 60- to 74-year-olds and from the 60- to 74-year-olds to the 75- to 85-year-olds (except in brightness discrimination). With accuracy instructions, 75- to 85-year-olds were about 200 msec slower than 60- to 74-year-olds, but the 60- to 74-year-olds were only about 10–60 msec slower than the college subjects. Age-related increases in RTs like those obtained here have been the basis for the general-slowing hypothesis (Brinley, 1965; Ratcliff et al., 2000).

In the next paragraphs, we show first that the model fit the data well for individual subjects for individual experiments; second, how components of processing change with age; third, that sensible and consistent interpretations

of performance across the three age groups and the four tasks are provided by the model; and fourth, which components of processing are correlated for individuals across the four tasks.

For each subject individually for each experiment, the diffusion model was simultaneously fit to the probability of a correct response and the RT distributions for correct and error responses for all conditions of the experiment. The fits of the model are constrained by several assumptions. First, subjects can adjust their response boundaries according to whether instructions emphasize speed or accuracy, but they cannot change drift rates. Drift rates are affected only by the quality of the information from a stimulus. Second, when subjects adjust response boundaries, they can separately adjust the distances from the starting point to the two boundaries. For the data presented here, the model fit well with the two distances set equal to each other (i.e., $z = a/2$), except for recognition memory. Third, it is assumed that T_{er} is constant across levels of stimulus difficulty and instructions. Although small differences in T_{er} might reasonably be assumed between, for example, speed versus accuracy instructions, little is added to the quality of the fits of the model to the data, and nothing significant would change in interpretations of the data. Finally, for the fits presented here, it was assumed that all the across-trials variability parameters were constant across levels of stimulus difficulty and instructions.

For each experiment, there was one value of drift rate for each experimental condition or group of conditions. For the signal detection experiment, the four conditions for which "large" was the correct response had the same absolute values of drift rates as the four conditions for which "small" was the correct response, except that indi-
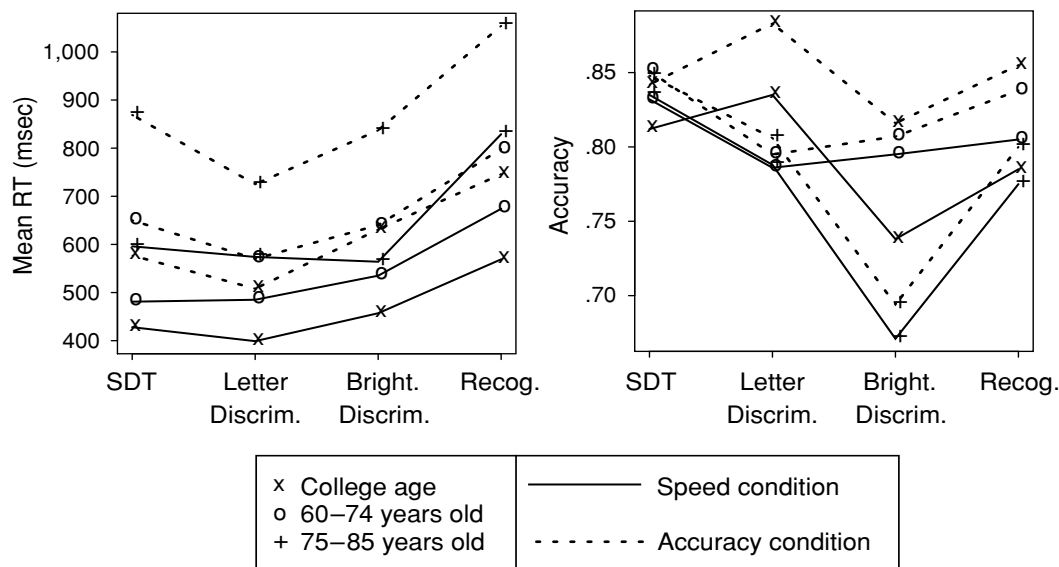

Figure 2. Mean values of correct RT and accuracy, averaged across subjects and experimental conditions for each task and each group of subjects. For both mean RTs and accuracy, the top line of the pair of lines for a group in each panel is for conditions with accuracy instructions, and the bottom line is for conditions with speed instructions.

Table 1
Mean Chi-Square Goodness-of-Fit Values for the Diffusion Model,
Averaged Over Subjects

Task	Young		60–74 Years Old		75–85 Years Old
	Mean	Prev.	Mean	Prev.	Mean
Signal detection	460		413		409
Letter discrimination	172	84	181	119	127
Brightness discrimination	869	680	945	951	919
Recognition memory	435	368	323	431	408

Note—For N experimental conditions, 10 quantile RTs (5 correct and 5 error), and a model with M parameters, the number of degrees of freedom is $df = N(12 - 1) - M$. The degrees of freedom and number of parameters for the signal detection task were 164 and 12; for the letter discrimination task, 77 and 11; for the brightness discrimination task, 376 and 19; and for the recognition memory task, 180 and 18. The numbers in the "Prev." columns are the average chi-square values from previously published experiments (Ratcliff, Thapar, & McKoon, 2003, 2004; Thapar, Ratcliff, & McKoon, 2003).

vidual subjects set their criterion between the two at different points. To represent this bias, a parameter c was added to the model (Ratcliff et al., 2003), such that the value of c was added to each of the four drift rates for "large" stimuli and to the negatives of the four drift rates for "small" stimuli. For letter discrimination, there was one value of drift rate for each of the three shortest stimulus durations and a fourth value for the three longest durations combined. For brightness discrimination, the three brightness conditions for which the correct response was "bright" had the same absolute values of drift rates as the three conditions for which the correct response was "dark," except that, as in the signal detection experiment, there was a bias parameter c. The three values of drift rate and c were different for the three stimulus durations. For the recognition memory experiment, there were nine drift rates, six for studied words (three levels of frequency, one or three presentations) and three for nonstudied words (levels of frequency). Also, the starting point z was an additional parameter because it was not equidistant between the two boundaries. The college subjects set the starting point nearer the "old" boundary by about 5% of total boundary separation, and both groups of older subjects set the starting point nearer the "new" boundary by between 5% and 20%. The same minimization routine as in previous research was used to find the values of all the parameters (Figure 1) that produced the smallest chi-square value for the data for each subject in each experiment.

The full range of data that the model must accommodate and the quality of its fits to the data are usually presented via quantile probability functions (see, e.g., RTM), which display, for each condition in an experiment, the quantiles of the RT distributions for correct and error responses and the probabilities of the responses. Such functions simultaneously show the relative speeds and accuracies of correct and error responses, the shapes of the RT distributions, and how speed and accuracy covary across conditions. However, the number of functions required for the four experiments and three subject groups presented

in this article would be large (40 panels), so instead goodness of fit is summarized statistically in Table 1.

Overall, the model fit the data well, as well as it had in the studies by RTM. To allow comparison of the chi-square values across the experiments reported here, those values were divided by their respective degrees of freedom. The resulting scores were 2.6, 2.1, 2.4, and 2.2 for the signal detection, letter discrimination, brightness discrimination, and recognition memory experiments, respectively. These scores, as well as the chi-square values in Table 1, are similar to those from the RTM experiments (except for the letter discrimination task, for which there were about 50% more observations here than in Thapar et al., 2003). Although the mean chi-square values in Table 1 are larger than the .05 significance level, the numbers of degrees of freedom and numbers of observations are so large that even small deviations between model and data lead to significant chi-square values (see Ratcliff et al., 2004, for further discussion).

INTERPRETING THE DATA THROUGH THE DIFFUSION MODEL

Age Effects

Figure 3 shows parameter values for the three age groups as a function of the four experiments. Table 2 shows which of the most diagnostic parameter values (a_{sp}, T_{er}, v, and η) change with age, and Table 3 shows values of the means and SDs of a_{sp}, T_{er}, and v (subscripts "sp" and "ac" refer to the speed and accuracy conditions, respectively). The tables report only values of a_{sp} because, as described below, subjects were more consistent in setting a_{sp} than a_{ac}.

First, with speed instructions, boundary separation was larger for 60- to 74-year-olds than for college subjects in six out of eight comparisons. With accuracy instructions, boundary separation was larger for the 75- to 85-year-olds than for the other groups of subjects. Second, the nondecision components of processing were consistently longer for 60- to 74-year-olds than for college subjects and for 75- to
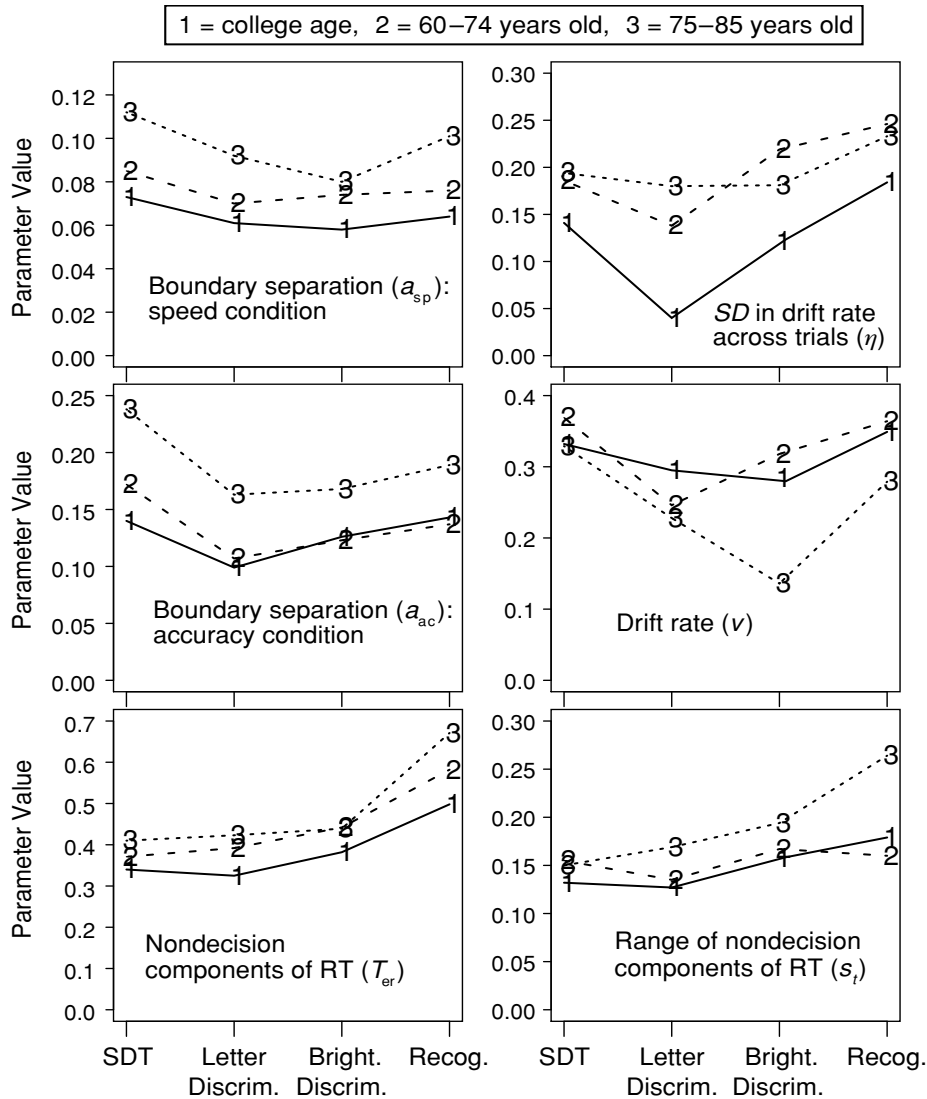
Figure 3. Mean parameter values averaged across subjects and conditions for each task and subject group. Values of s_z were between 0.004 and 0.008 for the letter discrimination experiment and between 0.021 and 0.040 for the other experiments.

85-year-olds than for 60- to 74-year-old subjects (only one exception out of eight comparisons; see Table 2). Third, the patterns of drift rates varied from task to task. In the signal detection task, there was no deficit for 60- to 74-year-olds relative to college subjects, and in recognition memory, there were only small differences. In letter discrimination, there was a drop in drift rate from college to 60- to 74-year-old subjects but no further drop to 75- to 85-year-old subjects, whereas in brightness discrimination, there was no drop from college to 60- to 74-year-old subjects but a significant drop for 75- to 85-year-olds. This matches the psychophysical finding that the ability to discriminate high-spatial-frequency information (which was necessary to discriminate between letters in our task) declines earlier in the aging process than the ability to discriminate lower spatial

frequency information (which was necessary to discriminate between bright and dark stimuli).

These results have generality in two ways. First, the findings generally replicate those of RTM and Ratcliff, Thapar, and McKoon (in press), which were derived with different subjects (and larger numbers, typically 30–40) for each experiment. Second, the conclusions are not specific to the diffusion model; other models that adequately fit the data provide the same conclusions (Ratcliff, Thapar, Smith, & McKoon, 2004). This is partly because in all the models, drift rate (or the rate of accumulation of information in accumulator models) averaged across conditions is relatively highly correlated with average accuracy. Thus, on average, if accuracy does not decline with age, neither does drift rate or accumulation rate. However, we caution that the mapping

Table 2
Effects of Aging on Components of Processing: Parameter Differences for the Diffusion Model

Task	60- to 74-Year-Olds vs. Young Subjects				75- to 85-Year-Olds vs. 60- to 74-Year-Olds			
	a_{sp}	T_{er}	v	η	a_{sp}	T_{er}	v	η
Signal detection	higher	longer	n.s.	n.s.	n.s.	longer	n.s.	n.s.
Letter discrimination	higher*	longer	lower*	higher*	higher	longer*	n.s.	higher
Brightness discrimination	higher	longer	n.s.	higher	n.s.	n.s.	lower	n.s.
Recognition memory	higher	longer	n.s.	n.s.	higher	longer*	lower*	n.s.

Note—Significance of the differences was computed using simple t tests. Asterisks indicate marginally significant effects (.05 < p < .10), and "n.s." indicates nonsignificant effects; all other differences were significant (p < .05). Note that for young versus older subjects, the results for a_{sp} (boundary separation for the speed condition), T_{er} (nondecision components of processing), and v (drift rate) replicated results from RTM (apart from a_{sp} for brightness discrimination).

from drift rate to accuracy is not perfect, because speed can be traded for accuracy by adjusting decision criteria.

Mapping Components of Processing Onto Performance

For each subject in each experiment, we computed correlations between their data—specifically, the mean values of accuracy, correct RT$_{sp}$, and error RT$_{sp}$ averaged across experimental conditions—and the parameter values from the best fits of the model to their data.

Across experiments, subjects set the boundary separation parameter more consistently with speed than with accuracy instructions, as evidenced by the relatively low correlation, .33, between a_{sp} and a_{ac}. This is not unexpected, because though most subjects interpret speed instructions in the same way, as demanding the maximum possible speed, subjects vary in their interpretations of just how much accuracy is required by accuracy instructions; some subjects do not differ much from the speed condition (e.g., 50 msec slower), but others differ a lot (e.g., 300 msec slower).

Table 3
Means and Standard Deviations for Parameter Values Across the 10 Subjects in Each Group

Parameter	Experiment	Young		60–74 Years Old		75–85 Years Old	
		M	SD	M	SD	M	SD
a_{sp}	SDT	.073	.009	.085	.013	.112	.068
	Let	.061	.008	.070	.013	.092	.023
	Brit	.057	.007	.074	.016	.080	.017
	Recg	.064	.009	.076	.010	.101	.015
T_{er}	SDT	.340	.027	.371	.029	.410	.037
	Let	.325	.028	.392	.029	.423	.047
	Brit	.399	.026	.442	.036	.439	.069
	Recg	.498	.044	.582	.038	.671	.123
v	SDT	.332	.110	.369	.103	.329	.125
	Let	.295	.094	.246	.173	.228	.085
	Brit	.269	.103	.318	.127	.136	.097
	Recg	.349	.212	.364	.115	.279	.080

Note—The number of observations in each group is 10, so the SE for a parameter value is SD/$\sqrt{10}$. SDT, signal detection experiment; Let, letter discrimination experiment; Brit, brightness discrimination experiment; Recg, recognition memory experiment.

With speed instructions, the correlation averaged over subjects and experiments between T_{er} and correct RT$_{sp}$ was .78 and that between a_{sp} and correct RT$_{sp}$ was .75, with a critical value for p = .05 of .27.[1] With accuracy instructions, the correlations for T_{er} and a_{ac} with correct RT$_{ac}$ were .39 and .72, respectively, with the lower value for T_{er} arising from the inconsistency of boundary settings across experiments with accuracy instructions. These results mean that, overall, slower subjects tended to have higher values of a_{sp} and a_{ac}, and with speed instructions longer values of T_{er}. Averaged over subjects, experiments, and instructions, the correlation between drift rate and accuracy was .74, indicating that more accurate subjects had higher values of drift rate.

The nonsignificant correlations were correct RT$_{sp}$ and drift, −.29; accuracy and T_{er}, .18; accuracy and a_{sp}, .22; and accuracy and correct RT$_{sp}$, −.05. The corresponding correlations for correct RT$_{ac}$ and accuracy were also not significant (accuracy values were highly consistent across speed and accuracy conditions, so the average was used). The correlations between error RT$_{sp}$ and the parameter values followed those for correct RT$_{sp}$, because these two correlated at .93. For the college and 60- to 74-year-old subjects, this pattern of correlations matches that obtained in earlier studies.

Among the correlations between variability parameters and other parameters, only that between drift rate and η correlated significantly (.66).

Individual Differences Across Experiments

We examined whether the model parameters were consistent across experiments for individual subjects by computing the correlations between each pair of experiments for each parameter. With four experiments, this yields six correlations for each of the three groups of subjects, and these are shown in Figure 4. Averaging across pairs of experiments and subject groups, the correlations were a_{sp}, .32; a_{ac}, .13; T_{er}, .47; η, .27; s_z, −.02; drift, .37; and s_t, .11. The critical value for the correlation coefficient was .14 (computed using a Monte Carlo method that mimics the multiple correlations and averages across subject groups). Thus, the values of a_{sp}, T_{er}, drift rates, and variability in drift across trials were reasonably consistent across ex-
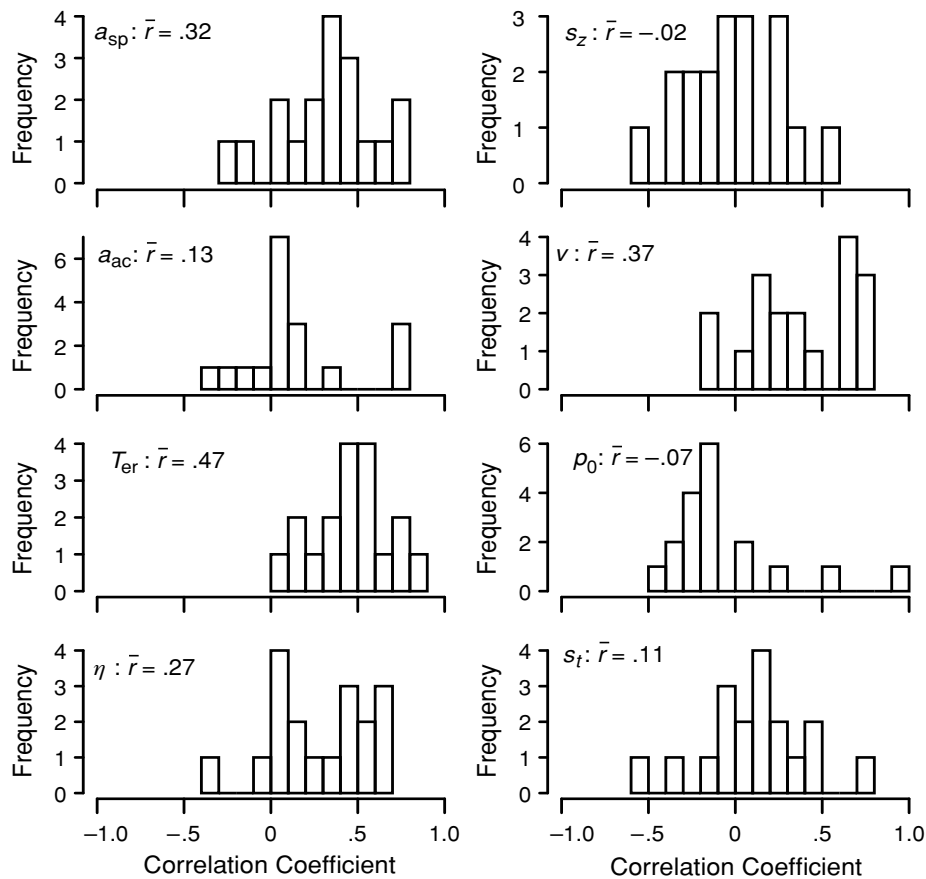
Figure 4. Histograms of correlations for parameter values of the model across all pairs of the four experiments and across the three subject groups. The mean values of the correlations are shown, and the critical value of the mean correlation is .14. p_0 is the proportion of contaminants estimated using the method of Ratcliff and Tuerlinckx (2002).

periments for individuals, whereas the values of the other parameters were not (for the variability parameters, this was partly because of high variability in their estimates [Ratcliff & Tuerlinckx, 2002], and for a_{ac}, we believe this was because subjects tended to interpret instructions differently from task to task, depending on how they perceived task difficulty).

For the simple two-choice tasks we have studied, the finding that there is consistency in the behavior of components of processing across tasks suggests a common decision process across tasks. Most surprising is the correlation in drift rates across tasks. This suggests a task-independent factor, so that if an individual is good at one task (i.e., has high drift rates), they are, on average, good at the other tasks.

GENERAL DISCUSSION

One of the main problems with traditional analyses of the effects of aging on simple two-choice tasks is that the two dependent measures, RT and accuracy, can show different patterns of aging effects. For example, with several

of our tasks, we have found no differences between young and older subjects in accuracy but reasonably large differences in RTs. The relatively large differences in RTs suggest a theoretical view in which aging has a relatively large effect on the rate of information processing. But the accuracy data suggest a view in which aging has relatively little effect on the quality of information extracted from stimuli. These two views are not necessarily contradictory, but they cannot be adequately reconciled without theory.

In the diffusion model framework, the large differences in RTs between the older and young subjects can arise from differences in nondecision components of processing, differences in criteria settings, and/or differences in drift rates. Fitting the model to data allows us to determine which of these or which combination of these is responsible for increases in RT. The different possibilities can be discriminated in fits to data because they have identifiably different effects on the shapes of RT distributions (Ratcliff & Tuerlinckx, 2002).

The diffusion model fit all the data from all four tasks, and it did so at the level of individual subjects who varied across a wide range of performance levels. The model's

interpretations of the components of processing indicate that response accuracy is mainly a function of the quality of the evidence from a stimulus, and that response speed is mainly a function of decision boundaries and, to some degree, of nondecision components of processing. Individual differences in components of processing were significantly correlated across experiments, and this was true for college students, 60- to 74-year-olds, and 75- to 85-year-olds. A subject who was slow in one task because of slow nondecision components of processing or conservative response criteria tended to be slow in the other tasks for the same reasons. A subject who was accurate in one task because of large values of drift rate tended to be accurate in the other tasks for the same reason. The significant correlations across tasks suggest that these components of an individual's processing are not task specific, but instead are significantly consistent across tasks.

For aging research, these results are important in showing that the much slower responses of older adults relative to young adults in cognitive tasks, especially in these simple two-choice tasks, can be due mainly to conservative response criteria, accompanied by relatively small slowdowns in nondecision components of processing. For the tasks studied here, differences in the quality of the evidence available from the stimuli between college and 60- to 74-year-old subjects appeared only for letter discrimination. A large difference in the quality of the evidence between 60- to 74-year-old and 75- to 85-year-old subjects appeared for brightness discrimination, as well as a small difference for recognition memory. This means that for these tasks (as well as for lexical decision; Ratcliff, Thapar, Gomez, & McKoon, 2004), there is little effect of age on the quality of evidence extracted from the stimulus, except with the perceptual tasks and marginally with recognition memory. For more complex tasks, tasks with memory loads, or tasks involving switching of attention, there may well be deficits in evidence in the decision process, but such questions will require a theoretical analysis of the kind promoted here.

It is noteworthy that large changes in RTs usually correspond to relatively small differences in accuracy (Figure 2). Although older subjects seem reluctant to move away from their generally conservative response criteria, altering criteria can produce 100- to 200-msec differences in RTs at a cost of only a 1%–3% decrease in accuracy. Such a trade-off might sometimes be appropriate in practical settings.

As the diffusion model's theoretical analyses are brought to various simple two-choice tasks, the rates of extraction of information from stimuli are decoupled from criterion effects and from nondecision components of processing. Instead of a monolithic account of processing speed in terms of only mean correct RTs, we have instead an account based on all aspects of the data, allowing the quality of information extracted from stimuli to be separated from subject-adjustable decision criteria. Such theoretical approaches might be fruitfully applied to other domains requiring more complex processing and might produce more detailed pictures of the loci of deficits in performance with increasing age. 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 for quantitative modeling approaches in general.

REFERENCES

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: Biological determinants of speed of behavior and its changes with age* (pp. 114-149). Springfield, IL: Thomas.

Cerella, J. (1994). Generalized slowing in Brinley plots. *Journals of Gerontology*, **49B**, P65-P71.

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. *Journals of Gerontology*, **49B**, P81-P89.

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.

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.

Myerson, J., Wagstaff, D., & Hale, S. (1994). Brinley plots, explained variance, and the analysis of age differences in response latencies. *Journals of Gerontology*, **49B**, P72-P80.

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. (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 response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review*, **9**, 278-291.

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

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

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

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

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

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology & 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 & Language*, **50**, 408-424.

Ratcliff, R., Thapar, A., & McKoon, G. (in press). Applying the diffusion model to data from 75–90 year old subjects in five experimental tasks. *Psychology & Aging*.

Ratcliff, R., Thapar, A., Smith, P. L., & McKoon, G. (2005). Ageing and response times: A comparison of sequential sampling models. In J. Duncan, L. Phillips, & P. McLeod (Eds.), *Measuring the mind: Speed, control, and age* (pp. 3-32). Oxford: Oxford University Press.

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

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

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

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

NOTE

1. The critical value of r for $p = .05$ and 10 subjects was .55. With 21 rs tested (all combinations of three data values and four parameters), the critical value was raised to .81, but because the averages over 12 experiments were computed, the critical value was lowered to .27. The critical values were obtained using 10,000 Monte Carlo simulations exactly mimicking the multiple comparisons and averages over experiments.