

# Effects of Aging and IQ on Item and Associative Memory

Roger Ratcliff  
The Ohio State University

Anjali Thapar  
Bryn Mawr College

Gail McKoon  
The Ohio State University

The effects of aging and IQ on performance were examined in 4 memory tasks: item recognition, associative recognition, cued recall, and free recall. For item and associative recognition, accuracy and the response time (RT) distributions for correct and error responses were explained by Ratcliff's (1978) diffusion model at the level of individual participants. The values of the components of processing identified by the model for the recognition tasks, as well as accuracy for cued and free recall, were compared across levels of IQ (ranging from 85 to 140) and age (college age, 60–74 years old, and 75–90 years old). IQ had large effects on drift rate in recognition and recall performance, except for the oldest participants with some measures near floor. Drift rates in the recognition tasks, accuracy in recall, and IQ all correlated strongly. However, there was a small decline in drift rates for item recognition and a large decline for associative recognition and cued recall accuracy (70%). In contrast, there were large effects of age on boundary separation and nondecision time (which correlated across tasks) but small effects of IQ. The implications of these results for single- and dual-process models of item recognition are discussed, and it is concluded that models that deal with both RTs and accuracy are subject to many more constraints than are models that deal with only one of these measures. Overall, the results of the study show a complicated but interpretable pattern of interactions that present important targets for modeling.

*Keywords:* reaction time, diffusion model, memory, item and associative recognition, IQ and age

In memory research, there is general agreement on a distinction between item and associative information. In tasks that tap item information, participants are asked to decide whether a test item was presented earlier in an experiment. In tasks that tap associative information, participants are asked to decide whether two items of a pair were presented earlier in the same pair or in a different pair. This distinction has a long history in experimental psychology. Murdock (1974), for instance, made the distinction a centerpiece of his approach to memory, and he reviewed much of the earlier work separating these two forms of memory. More recently, a number of studies have provided compelling evidence for the distinction (Clark & Shiffrin, 1992; Hockley, 1991, 1994; Hockley & Cristi, 1996; Humphreys, 1976, 1978; Malmberg & Xu, 2007; Murdock, 1974, 1992).

In this article, we set global memory models as the context for our research on item and associative information (Dennis & Humphreys, 2001; Gillund & Shiffrin, 1984; Humphreys, Bain, & Pike, 1989; McClelland & Chappell, 1998; Murdock, 1982; Shiffrin &

Steyvers, 1997). The item–associative distinction has been a fundamental component of these models since their beginnings. In all of these models, item and associative information are stored together in a single memory store. Whether items are stored as vectors (Murdock, 1982) or cue–target associations (Gillund & Shiffrin, 1984), there are not two separate stores. Instead, item and associative recognition correspond to different ways of retrieving information from a single memory store. In contrast to global memory models, there are dual-process models (see e.g., Kelley & Wixted, 2001; Yonelinas, 1997). For these models, item recognition and associative information can depend on separate sources of information. Item recognition is said to depend mostly on “familiarity,” whereas associative recognition is said to depend mostly on “recollection.” At the end of this article, we discuss dual-process theories in detail.

The experiment presented in this article provides new results that address the item–associative distinction. In the experiment, item recognition was tested with participants studying lists of single words, each followed by a test list of single words to which they responded “old” or “new.” Associative recognition was tested with participants studying lists of pairs of words, each followed by a test list made up of pairs of words. Participants responded according to whether the words of a pair had appeared in the same pair in the study list (“intact”) or different pairs in the study list (“rearranged”).

The first innovation reported here is that we applied the well-established diffusion model for response times (RTs) and accuracy to the data from associative recognition. Most previous studies of

---

This article was published Online First June 27, 2011.

Roger Ratcliff and Gail McKoon, Department of Psychology, The Ohio State University; Anjali Thapar, Department of Psychology, Bryn Mawr College.

Preparation of this article was supported by National Institute on Aging Grant R01-AG17083.

Correspondence concerning this article should be addressed to Roger Ratcliff, Department of Psychology, The Ohio State University, 1835 Neil Avenue, Columbus, OH 43210. E-mail: ratcliff.22@osu.edu

associative recognition have measured only accuracy. For this study, we measured both accuracy and RTs. It is important to stress that these two measures can behave differently as a function of age. For item recognition, previous research has found that older adults can be nearly as accurate as young adults, but they are much slower. For associative recognition, one might find a similar result for speed—older adults much slower—but one might also find that they are less accurate. The fact that the two variables can behave differently means that a full explanation of item and associative recognition must accommodate both accuracy and RT measures.

In previous articles, we and others have shown that the diffusion model can account for changes in RTs and accuracy across conditions in a range of experimental tasks, including animacy categorization, brightness discrimination, color discrimination, item recognition, letter discrimination, lexical decision, numerosity discrimination, recognition memory, and visual search (Ratcliff, 1978, 1981, 2002; Ratcliff & Rouder, 2000; Ratcliff, Van Zandt, & McKoon, 1999; Voss, Rothermund, & Voss, 2004). In research in domains such as aging, child development, sleep deprivation, depression, anxiety, aphasia, hypoglycemia, and vigilance, diffusion model analyses have led to new and different interpretations of performance, in particular by taking into account differences in speed–accuracy tradeoff settings across participants and experimental conditions (Geddes et al., 2010; Ratcliff, Love, Thompson, & Opfer, in press; Ratcliff, Perea, Coleangelo, & Buchanan, 2004; Ratcliff, Schmiedek, & McKoon, 2008; Ratcliff, Thapar, & McKoon, 2001, 2003, 2004, 2006a, 2006b, 2007, 2010; Ratcliff & Van Dongen, 2009; Schmiedek, Oberauer, Wilhelm, Suß, & Wittmann, 2007; Spaniol, Madden, & Voss, 2006; Wagenmakers, Van Der Maas, & Grassman, 2007; White, Ratcliff, Vasey, & McKoon, 2009, 2010a, 2010b).

In the present experiment we used an individual differences approach to ask questions about item and associative memory. The diffusion model was used to describe and compare the components of processing that underlie performance at the level of individual participants. The model explains differences in performance among individuals in terms of their speed–accuracy trade-off settings and the quality of the information upon which their decisions are based, as well as processes such as encoding and response execution.

The second innovation for the research reported here is that we examined the effects of both IQ and age on item and associative recognition. We used differential effects of IQ and age to separate the retrieval of item information from the retrieval of associative information. In essence, we show that item and associative recognition decline differentially with IQ and age.

Our approach here can be distinguished from previous studies of IQ. Much of the work using IQ has focused on an “abilities” approach in which measures from a range of different tasks are combined to represent a single construct. We know of little research on the relationship between IQ and item or associative recognition (Kaufman, DeYoung, Gray, Brown, & Mackintosh, 2009). Furthermore, what research has been done for IQ and long-term memory has been framed in terms of, for example, how working memory and speed-of-processing measures can be used to account for individual differences in long-term memory (see e.g., Unsworth, 2010).

For the effects of aging, the situation is different, with considerable research on the effects of aging on item and associative memory. Data have suggested that associative memory declines with age more than item memory (see e.g., Buchler & Reder, 2007; Craik, 1983, 1986; Craik & McDowd, 1987; Healy, Light, & Chung, 2005; Kausler, 1994; Naveh-Benjamin, 2000; Schonfield & Robertson, 1966; Wahlin, Backman, & Winblad, 1995). For item recognition, the decrements with age have been relatively small (Balota, Dolan, & Duchek, 2000; Bowles & Poon, 1982; Craik, 1994; Craik & Jennings, 1992; Erber, 1974; Gordon & Clark, 1974; Kausler, 1994; Neath, 1998, Chap. 16; Rabinowitz, 1984; Schonfield & Robertson, 1966). For associative recognition, Old and Naveh-Benjamin (2008) conducted a meta-analysis of data from 90 studies and found larger age-related deficits for associative recognition than for item recognition, under a wide variety of experimental manipulations.

In summary, there were three specific goals for the experiment described in this article. The first was to use Ratcliff’s diffusion model (Ratcliff, 1978; Ratcliff & McKoon, 2008) to extract, from RT and accuracy data, measures of the components of processing involved in item and associative recognition. The aim was that differences among individuals in their speed–accuracy criterion settings could be separated out of their data so as to allow pure comparisons of memory ability across the two tasks.

The second goal was to examine the effects of age and IQ on item and associative recognition. We expected that item and associative recognition would decline in memory at different rates as a function of age but perhaps interacting with IQ.

The third goal was to examine individual differences within age groups. In the experiment, there were three groups of participants: college-age individuals, 60- to 74-year-olds, and 75- to 90-year-olds. IQs ranged from 85 to 140. We expected performance to be worse for lower IQ participants, but perhaps how much worse would depend on age. For example, high-IQ participants might perform better than low-IQ participants for college-age participants but less so for 75- to 90-year-olds.

Our experiment was focused on item and associative recognition. However, to (modestly) connect with research on recall, we included tests of cued and free recall, both of which require associative information. For cued recall, participants studied pairs of words and were then given the first word of each studied pair and asked to recall the second. For free recall, they studied lists of single words and were then asked to recall as many of them as possible.

### The Diffusion Model, Aging, and IQ

The diffusion model applies to two-choice tasks for which mean RTs are short, typically less than 1.5 s. In the model, evidence about a stimulus is accumulated over time from a starting point ( $z$ ) to one or the other of two criterial amounts, or boundaries, one for each choice. The better the information from a stimulus, the faster evidence is accumulated. The rate of accumulation of evidence is called drift rate ( $v$ ). A response is executed when the amount of accumulated evidence reaches a criterion, either 0 for a negative response or  $a$  for a positive response. The processes outside the decision process (e.g., encoding, memory access, and response execution) are combined into a single parameter of the model with mean duration  $T_{cr}$ . Within-trial variability (noise) in the accumu-

lation of information from the starting point to the criteria results in processes with the same mean drift rate terminating at different times (producing RT distributions) and sometimes terminating at the wrong boundary (producing errors).

The values of drift rate, the criteria, and the nondecision component vary from trial to trial. This assumption is required if participants cannot accurately set these parameters to have the same values from trial to trial (see e.g., Laming, 1968; Ratcliff, 1978). Across-trial variability in drift rate is assumed to be normally distributed with standard deviation  $\eta$ , and across-trial variability in the nondecision component is assumed to be uniformly distributed with range  $s_1$ . Across-trial variability in the distance between the two criteria is equivalent to across-trial variability in the starting point (if such variability is not too large), which is assumed to be uniformly distributed with range  $s_2$ . Across-trial variability in drift rate and starting point are necessary for the model to account for the relative speeds of correct versus error RTs (Ratcliff et al., 1999). The model fits both correct and error RT distributions (an example is given later).

Performance includes, in addition to across-trial variability, “contaminant” responses—responses that are spurious in that they do not come from the decision process of interest (e.g., distraction, lack of attention). To accommodate these responses, on some proportion of trials ( $p_o$ ), a random delay is added to the decision RT. The across-trial variability in  $p_o$  is uniform between the maximum and minimum RTs for each experimental condition (the assumption of a uniform distribution is not critical; recovery of diffusion model parameters is robust to the form of the distribution; Ratcliff, 2008).

The model is designed to explain all aspects of the data: accuracy, mean correct and error RTs, RT distributions, and the relative speeds of correct and error responses. With only a single 45-min experimental session, the model can successfully fit data for individual participants, with standard deviations in the parameter estimates for boundary separation, nondecision time, and drift rate typically 3–5 times smaller than the standard deviations across participants.

The diffusion model is tightly constrained. The most powerful constraint comes from the requirement that the model fit the right-skewed shape of RT distributions (Ratcliff, 1978, 2002; Ratcliff & McKoon, 2008; Ratcliff et al., 1999). In addition, across experimental conditions that vary in difficulty (and are randomly intermixed at test), changes in accuracy, quantile RTs, and the relative speeds of correct and error responses are all captured by changes in only one parameter of the model, drift rate. The response criteria cannot be adjusted as a function of difficulty because it would be necessary for the system to know which level of difficulty was being tested before the accumulation of evidence began. It is also usually assumed that the processes that make up the nondecision component of the model do not vary with difficulty.

The diffusion model has been applied to a range of experimental tasks with younger and older adults as participants (Ratcliff, Thapar, Gomez, & McKoon, 2004; Ratcliff et al., 2001, 2003; Ratcliff, Thapar, & McKoon, 2004; Ratcliff et al., 2006a, 2006b, 2007, 2010; Spaniol et al., 2006; Thapar, Ratcliff, & McKoon, 2003). In particular, item recognition data have shown large increases in RTs with age coupled with small changes in accuracy or no changes in accuracy at all (Ratcliff, Thapar, & McKoon, 2004;

Ratcliff et al., 2006a, 2007, 2010). The RT data have suggested large decrements in information in memory with age, whereas the accuracy data have suggested only small decrements. The diffusion model reconciles these seemingly inconsistent results by mapping the two dependent variables onto the same underlying decision process. We have found that large increases in RTs with age are due mainly to increases in criteria settings and the duration of the nondecision processes and that small or nonexistent deficits in accuracy are due to small or nonexistent decreases in drift rates. From these findings we have concluded that, for item recognition, drift rates change little with age.

With IQ, item recognition data have shown a different pattern. Ratcliff et al. (2010) found that accuracy increases with IQ but that IQ has only small effects on RT. The model handles this with drift rates. Drift rates increase with IQ, as would be expected, but the changes in the criteria and the nondecision component are small.

## Experiment

In the present experiment we used the four previously described tasks: item recognition, associative recognition, cued recall, and free recall. For the two recognition tasks, we collected sufficient data to allow estimation of the components of processing identified by the diffusion model at the level of individual participants. Also, for all four tasks, we collected sufficient data to allow calculations of meaningful correlations. We calculated correlations among all the performance measures and, for the two recognition tasks, all the components of the diffusion model.

The experiment was designed to answer several questions. One was whether the diffusion model could be extended to associative recognition. The second was whether there would be significant correlations across participants between item recognition and the associative recognition and recall tasks; if so, the data would suggest that the tasks depend on a common representation in memory.

Two further questions were, first, whether there were significant effects of age and IQ on item recognition, associative recognition, cued recall, and free recall and, second, whether the effects of age were modulated by IQ. For example, it might be that performance declined with age for low-IQ participants more than for high-IQ participants, and perhaps this might be true for associative but not item recognition.

In the experiment, each participant participated in three sessions, one to test item recognition, one to test associative recognition, and one to test cued and free recall. Some of the demographic measures described later were tested at the end of the first session, some at the end of the second, and some at the end of the third. All four tasks consisted of a series of study–test blocks. For the item recognition task, each block was made up of 12 single words to study plus one buffer word at the end of the study list, followed by 26 test words, with the first two test words being either the buffer word or a negative filler. For each test word, participants were asked to respond “old” or “new” according to whether the word had appeared in the immediately preceding study list. For the associative recognition task, each block contained eight pairs of words to study, each presented twice in random order, with one buffer pair at the beginning of each block and a different buffer pair at the end, followed by eight test pairs, with the first pair testing either an intact buffer pair or a rearranged pair

constructed from the two buffer pairs. For each test pair, participants were asked to respond "intact" if the two words had been studied in the same pair or "rearranged" if they had been studied in different pairs. For the cued recall task, each block contained 16 pairs of words to study, each presented twice, followed by the first words of each pair presented as cues for recall. For the free recall task, each block contained eight single words to study, each presented twice, followed by instructions to recall the eight words.

For the associative recognition task, the words of all the test pairs were presented for study, which means that only associative information could be used to make a decision. Also, participants were strongly encouraged to respond quickly. The aim was that only associative information that was available immediately at test could be used to make a decision, so that it would be unlikely that participants would engage in strategic processes such as several different recall attempts.

## Method

**Participants.** In the experiment there were 46 college-age participants, 45 participants 60–74 years old, and 43 participants 75–90 years old. The college-age participants were recruited at Bryn Mawr College and Ohio State University and in surrounding areas. The older adults were community-dwelling volunteers from the Bryn Mawr, Pennsylvania, and Columbus, Ohio, areas. All participants were paid for their participation—\$15 per session for the older adults, \$12 for college-age participants who did not drive to the lab, and \$15 for those who did. All had a score of 26 or above on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975) and no evidence of disturbances in consciousness, medical or neurological disease causing cognitive impairment, history of head injury with loss of consciousness, or current psychiatric disorder. They also completed the Center for Epidemiological Studies—Depression scale (Radloff, 1977), for which there were no significant differences among the three age groups. To measure IQ, we asked participants to complete the Vocabulary and Matrix Reasoning subtests of the Wechsler Adult Intelligence Scale—Third Edition (WAIS-III; Wechsler, 1997). Means and standard deviations for these participant characteristics are shown in Table 1. With the exception of the raw IQ scores, there were no significant differences on any of the measures shown in Table 1.

It is important to note that even though our 75- to 90-year-old participants matched the younger ones on all the measures just described, it is likely that, relatively speaking, they are higher functioning. This is because they self-select into the experiment, which means that they represent only older adults who are healthy and mobile. Because of this self-selection, our data may actually underestimate declines in performance with age relative to the population from which our younger participants are drawn.

**Stimuli.** There were three pools of words: 800 high-frequency words with frequencies from 78 to 10,600 per million ( $M = 325$ ,  $SD = 645$ ; Kucera & Francis, 1967), 800 low-frequency words with frequencies of 4 and 5 per million ( $M = 4.41$ ,  $SD = 0.19$ ), and 681 very low-frequency words with frequencies of 1 per million or no occurrence in the Kucera and Francis (1967) corpus ( $M = 0.365$ ,  $SD = 0.48$ ). All words occurred in the *Merriam-Webster Ninth Collegiate Dictionary* (Merriam-Webster, 1990). The words were screened by three Northwestern University undergraduate students, and any words

Table 1  
*Participant Characteristics*

Measure	College age		60–74 years old		75–90 years old	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Mean age (in years)	20.4	2.2	68.3	4.4	82.0	4.1
Education (in years)	12.8	1.1	15.0	2.7	14.5	2.8
MMSE	28.7	1.4	28.3	1.5	28.0	1.1
WAIS-III Vocabulary (scaled)	12.6	2.9	12.9	3.0	13.1	2.9
WAIS-III Vocabulary (raw)	46.3	10.1	51.2	10.7	49.3	10.9
WAIS-III Matrix Reasoning (scaled)	12.0	2.8	12.3	3.2	12.4	3.6
WAIS-III Matrix Reasoning (raw)	19.3	4.0	15.0	5.8	11.8	5.3
WAIS-III IQ	113.0	14.6	114.7	16.2	115.6	14.5
CES-D	10.8	7.7	10.0	8.9	10.5	6.4

*Note.* MMSE = Mini-Mental State Examination; WAIS-III = Wechsler Adult Intelligence Scale—Third Edition; CES-D = Center for Epidemiological Studies—Depression scale.

that they did not know were eliminated. For all four tasks, stimuli were chosen randomly without replacement from these pools.

**Procedure.** For all four tasks, stimuli were presented on the screen of a PC. For item recognition and associative recognition, responses were made on the PC's keyboard. For cued and free recall, participants' responses were given verbally with an experimenter recording them.

**Item recognition.** There were 50 study–test blocks. For each block, the study list consisted of six high- and six low-frequency words (in random order), each displayed for 1,300 ms and followed by a 200-ms blank screen and then the next word. An additional filler word, a very low-frequency word, was placed at the end of the study list as a buffer item. The manipulation of word frequency was included for both this and the associative recognition task in order to provide more conditions and therefore greater constraints on fitting the diffusion model to the data.

The test list immediately followed the study list. The first two test words were fillers—either two new, very low-frequency words or one new, very low-frequency word and the last, buffer, word of the study list (the first words of a test list are typically slower than the others, so these two test words were eliminated from data analyses). The remaining 24 test words, presented in random order, were the 12 studied words plus 12 new words, six high- and six low-frequency. Participants were asked to press the *?* key on the keyboard if the test word had been presented in the immediately preceding study list and the *Z* key if not. Participants were encouraged to respond as quickly as possible while still maintaining high accuracy. If the RT for a word was greater than 800 ms for college-age participants or greater than 900 ms for 60- to 90-year-old participants, a message saying *too slow* was displayed for 300 ms. If the RT was shorter than 280 ms, a message saying *too fast* was displayed for 1,500 ms. In all cases, there was a 500-ms blank screen immediately prior to the next test word. Participants were given no feedback on accuracy.

**Associative recognition.** There were 40 study–test blocks. For each block, the study list consisted of four high- and four low-frequency word pairs, each presented twice. The pairs were

presented in random order except that there was at least one other pair between repetitions of a pair. Each pair was displayed for 1,800 ms, followed by a blank screen for 200 ms before the next pair was displayed. Two additional pairs of words from the very low-frequency pool served as buffer pairs, one placed at the beginning and one placed at the end of the study list.

The test list immediately followed the study list. All of the words in the test pairs had appeared in a study pair. The first test pair was either the first or last buffer pair (intact) or one word from each of the two pairs (rearranged). The remainder of the test list consisted of two intact high-frequency pairs, two intact low-frequency pairs, two rearranged high-frequency pairs, and two rearranged low-frequency pairs, presented in random order. Participants were asked to press the *?* key if the two words of a test pair had occurred in the same pair and the *Z* key if the words had occurred in different pairs. The words in the test pairs always occupied the same place as in the study list: If a word was the first of a pair in the study list, it was the first of a test pair, whether the pair was intact or rearranged. Participants were given the same instructions as for the item recognition task: to respond as quickly as possible while maintaining high accuracy.

For each test pair, the first word was displayed for 300 ms and then the second word was presented immediately below the first. The first word was displayed first in order to reduce the variability in RTs that could result from reading times that included both words. Both words remained on the screen until a response was made. If a response was slower than 1,000 ms for college-age participants or 1,100 ms for 60- to 90-year-old participants, *too slow* was displayed for 300 ms. If the response was faster than 280 ms, *too fast* was displayed for 1,500 ms. There was a blank screen of 500 ms immediately before the next test word. No feedback on accuracy was given.

The *too slow* RT feedback was set relatively quickly—at 800–900 ms for the item recognition task and 1,000–1,100 ms for the associative recognition task—in order to encourage participants to go with their first impression (i.e., the first information that was available to them). For each task, the values were above the median RT. With enough time, participants could engage in recall processes that produced information about, for example, where an item appeared in a study list, what other items were before or after it, and so on. The *too slow* messages in our experiments were intended to discourage such slow deliberative processes.

**Cued recall.** The number of blocks varied from participant to participant as determined by the time available in a 30-min session. The mean numbers of blocks were 10.6, 9.8, and 10.1, for the college age, 60- to 74-year-old, and 75- to 90-year-old participants, respectively. Ninety percent of the college-age participants completed 8–12 blocks, 90% of the 60- to 74-year-olds completed 7–12 blocks, and 90% of the 75- to 90-year-olds completed 8–12 blocks.

Each study list was composed of 16 pairs of words, half high-frequency pairs and half low-frequency pairs. The pairs were presented twice, in the same order. At test, the first word of each pair was displayed on the PC screen, and participants were asked to respond with the second word. They made their responses verbally, recorded by an experimenter, with the participant deciding when to go on to the next test cue. The test cues were presented in the same order as the study pairs to equate lag effects (though

it is difficult to see how randomizing the order would change any of the main trends in the results).

**Free recall.** Participants were tested on between 16 and 26 lists, as determined by the time available in a 30-min session. The number of lists per participant averaged 23.7, 21.4, and 22.5, for the college-age, 60- to 74-year-old, and 75- to 90-year-old participants, respectively. Ninety percent of the college-age participants completed 19–26 lists; 90% of the 60- to 74-year-old participants, 15–26 lists; and 90% of the 75- to 90-year-old participants, 14–26 lists.

Each study list consisted of four high- and four low-frequency words presented twice in random order (with the restriction that two other words intervene between the two presentations of a word) plus one very low-frequency word presented as the last word in the study list. Each study word was displayed for 1,300 ms, followed by a 200-ms clear screen before the next study word. Participants recalled a list verbally, with an experimenter recording their responses and with the participants deciding when to go on to the next list.

The study lists were short, only nine words, because we found in piloting the experiment that the oldest participants were near floor, some recalling none of the words. So we opted to provide an easy free recall task in order to keep morale up. With only nine words to study, this task is probably not comparable to the other tasks, because recall could come from short-term as well as long-term memory. In the other tasks, there was enough separation between study and test that only long-term memory was being tested. Therefore, although the free recall data can be used in correlational analyses, to see whether good performance on free recall is associated with good performance on the other tasks, declines in free recall with age cannot be compared directly with declines in the other tasks.

## Empirical Data: Accuracy and RTs as a Function of Age, IQ, and Task

### Summary

1. For accuracy in item recognition, our finding is consistent with previous research (see e.g., Ratcliff et al., 2001, 2003; Ratcliff, Thapar, & McKoon, 2004; Ratcliff et al., 2006a, 2006b, 2007, 2010): There was little decline with age.

2. For accuracy in associative recognition, our finding is also consistent with previous research (see e.g., Naveh-Benjamin, 2000). Accuracy declined considerably with age for all three associative tasks: associative recognition, cued recall, and free recall.

3. The results for IQ were different for the college-age and 60- to 74-year-old participants than for the 75- to 90-year-old participants. For college-age participants and 60- to 74-year-olds, accuracy on both of the recognition tests increased significantly with IQ. But for the 75- to 90-year-olds, it did not; performance for most of the 75- to 90-year-olds was close to floor. For all three groups, there were modest increases with IQ for cued and free recall.

4. For accuracy in associative recognition, the interaction of age and IQ was striking. For the highest IQ participants, the drop in performance with age was precipitous—from about 75% correct to about 65% to close to 50% (floor), for college-age, 60- to 74-year-

old, and 75- to 90-year-old participants, respectively. For the lowest IQ participants, performance was near floor for all three age groups.

5. RTs increased with age for the two recognition tasks. RTs decreased as IQ increased for the recognition tasks, except among the college-age participants (cf. Ratcliff et al., 2010).

In the next sections, we provide details and analyses of variance for these results. Then we use the diffusion model to explain the data in terms of components of processing that underlie performance, and then we examine individual differences.

### Accuracy, Median RTs, and Age for Item and Associative Recognition

For college-age participants, for both item recognition and associative recognition, responses below 300 ms and above 1,500 ms were excluded. This was 0.8% of the data for item recognition and 1.8% for associative recognition. For 60- to 90-year-olds, responses for both tasks were excluded if they were below 300 ms or above 3,500 ms. For the 60- to 74-year-olds, this was 0.3% of the data for both tasks. For the 75- to 90-year-olds, 0.5% of the data were excluded for item recognition and 0.6% of the data for associative recognition. For both tasks, if there was a *too slow* message, the data were included in analyses unless they fell outside of the cutoff values just listed.

The left panel of Figure 1 shows accuracy and median RTs for “old” and “new” responses as a function of age for high- and low-frequency and old and new words. The right panel shows accuracy and median RTs for “intact” and “rearranged” responses as a function of age for pairs made up of two high-frequency words or two low-frequency words.

Comparing associative recognition with item recognition, accuracy decreased with age for associative recognition but not for item recognition,  $F(2, 125) = 24.7$  and  $F(2, 125) = 1.1$ , respectively. RTs increased with age for both tasks,  $F(2, 125) = 49.8$  and  $F(2, 125) = 48.0$ , respectively. In the next paragraphs, we give further details. Although these details heavily constrain the diffusion model in fitting the data, they are not central to the general conclusions about aging and individual differences that we draw.

For item recognition, for all three groups of participants there was a mirror effect in accuracy, with hit rates for low-frequency words higher than for high-frequency words and false alarm rates lower for low-frequency words than for high-frequency words. The  $d'$  values for college-age participants, 60- to 74-year-olds, and 75- to 90-year-olds for high-frequency words for item recognition were 1.23, 1.19, and 1.03, respectively. For low-frequency words, the  $d'$  values were 1.77, 1.77, and 1.60, respectively. For comparison, the same  $d'$  values for associative recognition were 0.96, 0.51, and 0.29 for low-frequency pairs and 0.93, 0.50, and 0.35 for high-frequency pairs for the three age groups, respectively. Thus, for item recognition, we see a mirror effect (see Figure 1) with  $d'$  for low-frequency words higher than for high-frequency words, but there was no effect of word frequency on associative recognition.

The RT difference between high- and low-frequency words changed relatively little with age, probably because of emphasis on speed in the instructions and because of the RT feedback. For the two older age groups, errors were slower than correct responses (“old” responses to not-studied items were slower than “new”

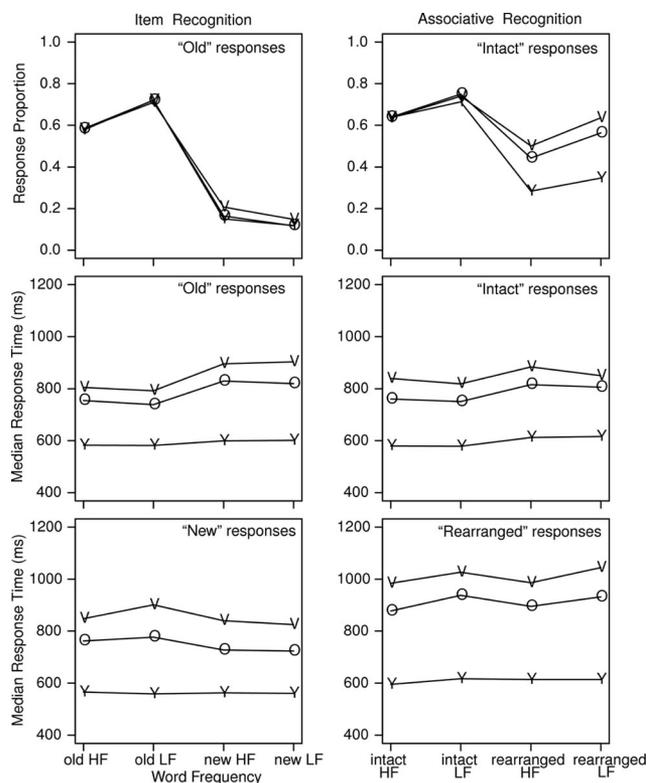


Figure 1. Accuracy and median response times (RTs) for old and new responses for item recognition and “intact” and “rearranged” responses for associative recognition. Note that RTs are not much longer for associative recognition than for item recognition because the first word of a test pair in associative recognition was presented for 250 ms prior to the second word and RT was measured from presentation of the second word. There is less than a 2% difference between HF and LF words in cued and free recall. Y = college-age participants; O = 60- to 74-year-old participants; V = 75- to 90-year-old participants; HF = high-frequency words; LF = low-frequency words.

responses to not-studied items, and “new” responses to studied items were slower than “old” responses to studied items). The college-age participants showed little difference for these comparisons.

For associative recognition, the decrease in accuracy with age was the result of a decrease in accuracy for rearranged, not intact, test pairs (cf. Chalfonte & Johnson, 1996; Light, Patterson, Chung, & Healy, 2004; Mitchell, Johnson, Raye, Mather, & D’Esposito, 2000). There was an overall bias such that “intact” responses were more likely for low- than high-frequency pairs, both for correct responses and errors.

“Rearranged” responses slowed more with age than did “intact” responses. For intact pairs, RTs for correct responses were shorter than for error responses, but for rearranged pairs, the opposite was true. This reflects a general bias such that “intact” responses were faster than “rearranged” responses.

### Accuracy, Median RTs, and IQ for Item and Associative Recognition

Figure 2 plots accuracy and median RTs for correct responses as a function of IQ and age, averaged over high- and low-frequency

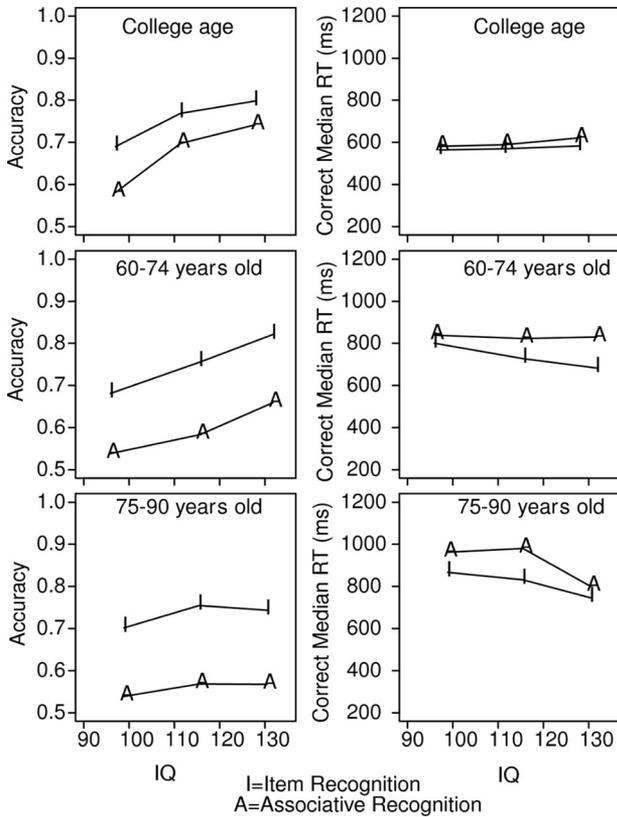


Figure 2. Accuracy and median correct response times (RTs) averaged over word frequency and response type for item and associative recognition as a function of IQ.

words, old/new items for item recognition, and intact/rearranged pairs for associative recognition. The numbers of college-age, 60- to 74-year-old, and 75- to 90-year-old participants, respectively, in the IQ groups were as follows: low-IQ groups: 15, 15, and 14; middle-IQ groups: 15, 15, and 15; and high-IQ groups: 16, 15, and 14.

For item recognition, accuracy was above chance for the lowest IQ participants for all three age groups. The increase in accuracy with IQ was significant,  $F(2, 125) = 18.8$ . The decrease in RTs with IQ was also significant,  $F(2, 125) = 3.2$ , although this effect occurred only for the 60- to 74-year-old and 75- to 90-year-old groups.

For associative recognition, the interaction between age and IQ was one of the central findings of this experiment. Accuracy was near chance for the lowest IQ participants for all three age groups. For the college-age and 60- to 74-year-old participants, accuracy increased as IQ increased,  $F(2, 125) = 21.1$ . This was not true for the 75- to 90-year-old participants, whose performance was close to chance for all levels of IQ; the interaction between age and IQ was significant,  $F(4, 125) = 3.38$ . Median RTs, on the other hand, varied little with IQ,  $F(2, 125) = 0.8$ .

**Accuracy, Age, and IQ for Cued and Free Recall**

Accuracy for high- and low-frequency words for the two recall tasks averaged over levels of IQ is plotted in the top panel of

Figure 3. Accuracy decreased with age,  $F(2, 125) = 26.8$  and  $F(2, 125) = 16.3$ , for cued and free recall, respectively, more for cued than for free recall, probably due to the short lists for free recall (only nine words), compared with the 16 pairs of words for cued recall. For neither task were there significant effects of word frequency. Significant effects of word frequency on recall performance have often been found in the literature, but nonsignificant effects have sometimes been obtained when high- and low-frequency words were mixed within a list (see e.g., Gillund & Shiffrin, 1984), as they were in this experiment.

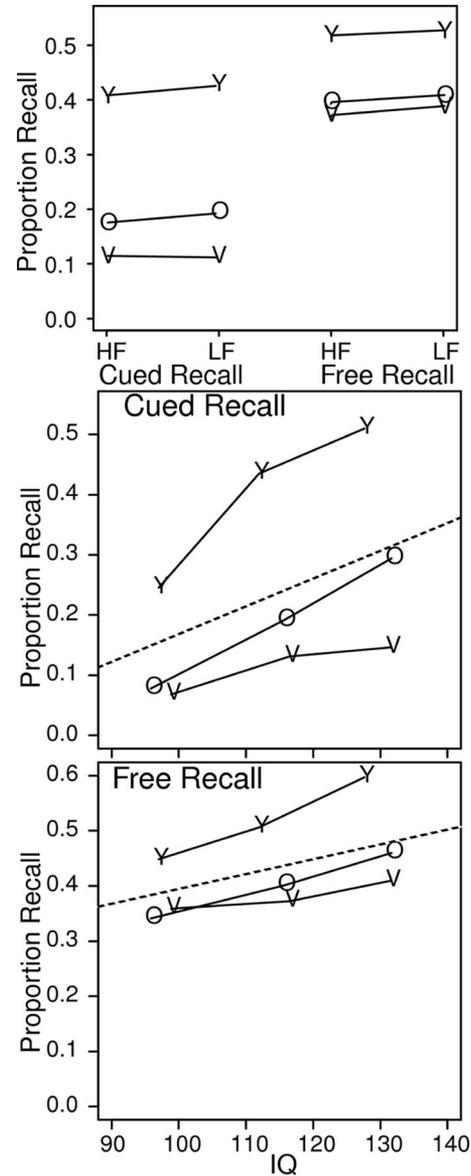


Figure 3. Panel A: Accuracy in cued and free recall as a function of word frequency and age. For cued recall, both words in the study pair are high frequency (HF) or low frequency (LF). Panels B and C: Overall accuracy as a function of IQ for the three participant groups in cued and free recall. Y = college-age participants; O = 60- to 74-year-old participants; V = 75- to 90-year-old participants.

Accuracy increased with IQ for both cued and free recall for all three participant groups,  $F(2, 125) = 66.9$  and  $F(2, 125) = 27.6$  for cued and free recall, respectively. Accuracy as a function of IQ is plotted in the bottom two panels of Figure 3. In cued recall, for the 75- to 90-year-olds, the advantage of higher IQ was less than for the 60- to 74-year-olds and college-age participants, with an interaction,  $F(4, 125) = 3.6$ . But for free recall, the effect was not significant,  $F(4, 125) = 2.0$ . For cued recall, just as for associative recognition, accuracy for the higher IQ participants dropped considerably when comparing college-age participants (about 50% accuracy) with 75- to 90-year-olds (a little less than 15% accuracy).

One possible reason the effects were smaller for free recall than for cued recall is that in free recall, some items might have been retrieved from short-term memory. To test this hypothesis, we examined the data with the first two items that were recalled eliminated from the analyses. If short-term memory were a factor, then with those items eliminated the change in accuracy would be larger as a function of age and IQ. However, with the first two recalled items excluded, the size and pattern of the results did not change (and the interaction was still not significant). This suggests that short-term memory was not a major factor responsible for the smaller decline in performance for free recall relative to cued recall.

### Diffusion Model Analyses

The diffusion model was fit to the data for each task for each participant by minimizing a chi-square value with a general simplex minimization routine (Nelder & Mead, 1965). In this procedure, the values of all the parameters, including the variability parameters, are estimated simultaneously, fitting the model to all the data from all the conditions of an experiment. The minimization routine adjusts the parameters of the model until it finds the parameter estimates that give the minimum chi-square value (see Ratcliff & Tuerlinckx, 2002, for a full description of the method). The data entered into the minimization routine for each experimental condition are the .1, .3, .5, .7, and .9 quantile RTs for correct responses and error responses and the corresponding accuracy values. The diffusion model is used to generate the predicted cumulative probability of a response by each quantile RT. Subtracting the cumulative probabilities for each successive quantile from the next higher quantile gives the proportion of responses between adjacent quantiles. For the chi-square computation, these are the expected values, which are to be compared with the observed proportions of responses between the quantiles (i.e., the proportions between 0, .1, .3, .5, .7, .9, and 1.0, which are .1, .2, .2, .2, .2, and .1) multiplied by the number of observations. Summing over  $(\text{Observed} - \text{Expected})^2 / \text{Expected}$  for all conditions gives a single chi-square value to be minimized.

The model can successfully fit data from single participants if there are around 300–1,000 total observations per participant, which can be collected in 45 min or less for tasks such as those considered in this article. Although such numbers of observations lead to variability in parameter estimates from their true values, the variability is much less than the differences among individual participants. In consequence, correlations of parameter values can be meaningfully used to examine individual differences.

In the analyses that follow, we use the model's parameters in two ways. One is to examine correlations among model parameters across participants in order to examine individual differences on the four tasks. The other is to average over participants to examine the overall effects of age and IQ on the components of processing measured by the model parameters.

### Goodness of Fit

Before we move to the explanations of the item and associative recognition data offered by the model, and comparisons with the cued recall and free recall data, it is important to demonstrate that the model's predictions match the data. In other words, when the best fitting parameter values are determined (by the fitting method described earlier), the predictions from the model using those parameter values should closely match the data. As we detail in the next paragraphs, the match between predictions and data was good.

The first result to note is that the model fitted the associative recognition data well. This is the first time the model has been applied to associative recognition, and so this extends the domain of application of the diffusion model.

Goodness of fit is illustrated in Figure 4 with one condition for the item recognition task (studied, high-frequency words) and one for the associative recognition task (intact, high-frequency pairs). These conditions had a relatively broad range of accuracy values across participants and showed the largest individual differences. (The total number of figures would be too large to show all conditions). The figure plots predictions from the model against the data for each participant. The accuracy values show deviations between predictions and data of no more than about 8%. For the RT quantiles for correct responses, there are only a few misses of more than 100 ms, except for the .7 and .9 quantiles, for which variability was larger (see Ratcliff & Tuerlinckx, 2002). For error responses, there are more misses because there were fewer numbers of observations and variability was large.

For each of the model's parameters, Table 2 shows the mean (the average over participants) of the best fitting value and its standard deviation. Table 2 also shows chi-square goodness of fit values averaged over participants. The critical value for 33 degrees of freedom is 50.7. The mean values are only a little higher than this, showing, along with the predictions in Figure 4, that the model fits the data well (see Ratcliff, Thapar, Gomez, & McKoon, 2004, for discussion of the quality of fits of the diffusion model to data).

### Diffusion Model Parameters, Age, and IQ

As we pointed out earlier, comparisons of performance for younger and older adults have frequently rested on accuracy data alone, setting aside RT data. Interpretations of age and IQ effects from accuracy alone (or from RTs alone) might be appropriate if RTs and accuracy were highly correlated across individuals. Young participants would be both faster and more accurate than older participants. High-IQ participants would be both faster and more accurate than lower IQ participants. But the effects of age and IQ on RTs and accuracy are not always the same. In the data described earlier for item recognition, the older participants differed from the younger participants in RTs but not accuracy. The

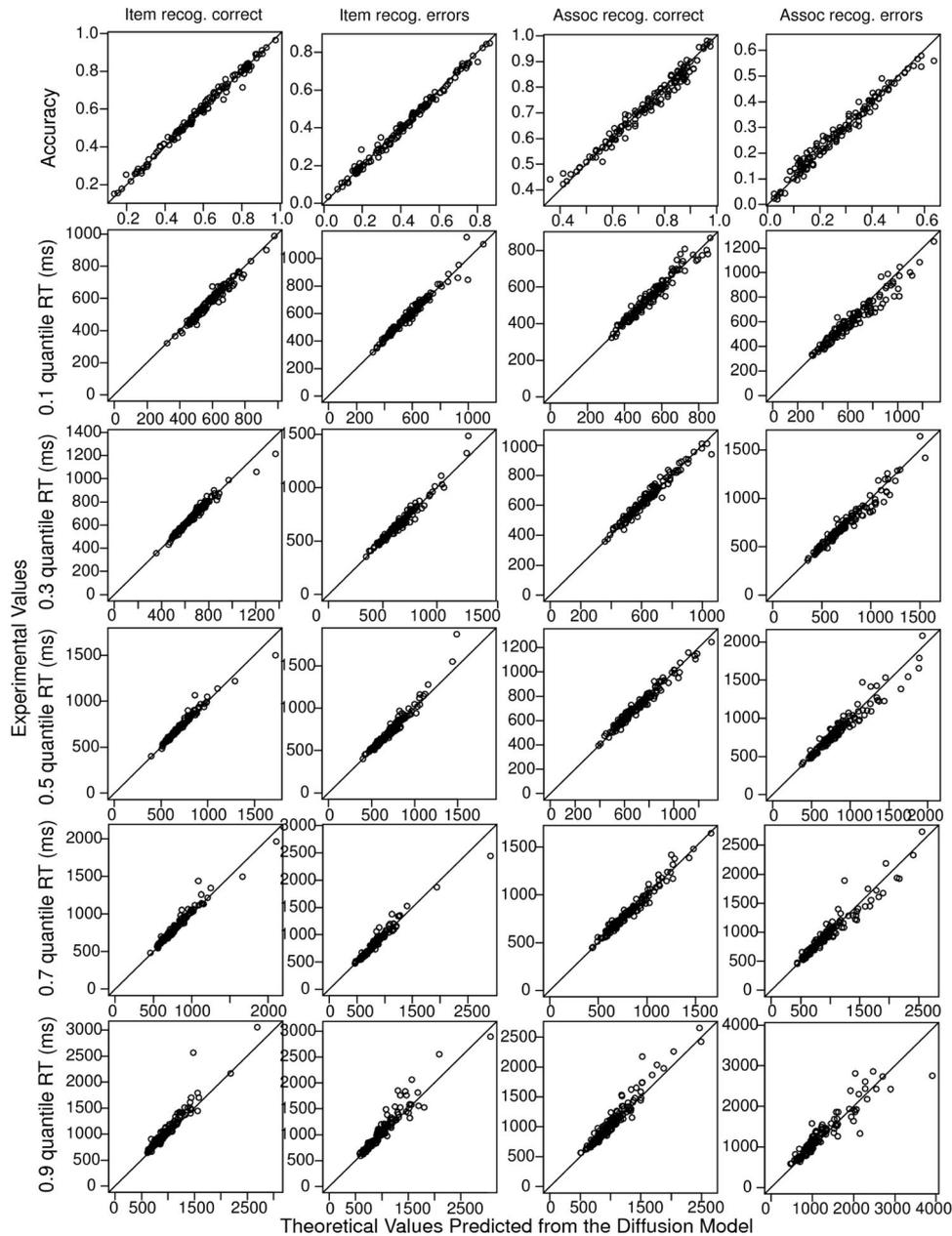


Figure 4. Plots of accuracy and the .1, .3, .5 (median), .7, and .9 response-time (RT) quantiles for data (y-axis) and predicted values from fits of the diffusion model (x-axis) for correct and error responses for a single condition for item recognition (Item recog.) and a single condition for associative recognition (Assoc recog.) for all participants in the three age groups. For Item recog., the data are for high-frequency words presented twice, and for Assoc recog., the data are for high-frequency, intact pairs.

differences in performance for lower compared with higher IQ participants were larger in accuracy than in RTs. For associative recognition, the older participants differed from the younger participants in both RTs and accuracy.

To preview the aging results, drift rates decreased with age by only a small amount for item recognition but by a much larger amount for associative recognition. Both boundary separation and the duration of the nondecision component increased with age, for both item and associative recognition.

To preview the IQ results, drift rates increased with IQ except for associative information for the 75- to 90-year-old participants. Boundary separation and the duration of the nondecision component were not significantly affected by IQ.

**Drift rates.** The top panels of Figure 5 show drift rates for item and associative recognition as a function of age and IQ, and for comparison, Figure 3 shows accuracy for cued and free recall. The drift rates shown in this and subsequent figures are the averages over experimental conditions. For item recognition and

Table 2  
Means in Parameter Values for Participant Groups and Item and Associative Recognition

Task, parameter, and participant group	$a$	$z$	$T_{er}$	$\eta$	$s_z$	$p_o$	$s_t$	$v_1$	$v_2$	$v_3$	$v_4$	$\chi^2$
Item mean												
College age	0.083	0.047	0.476	0.136	0.039	0.004	0.176	0.119	0.249	-0.262	-0.314	78.3
60-74 years old	0.116	0.061	0.575	0.180	0.041	0.001	0.187	0.108	0.217	-0.264	-0.325	65.9
75-90 years old	0.130	0.058	0.623	0.202	0.044	0.001	0.222	0.052	0.164	-0.242	-0.292	71.5
Associative mean												
College age	0.093	0.043	0.461	0.150	0.029	0.000	0.231	0.087	0.156	-0.161	-0.124	45.5
60-74 years old	0.130	0.049	0.569	0.134	0.019	0.000	0.292	0.035	0.107	-0.075	0.005	53.4
75-90 years old	0.138	0.051	0.600	0.116	0.023	0.000	0.310	0.014	0.080	-0.050	0.022	53.5
Item SD												
College age	0.014	0.011	0.048	0.092	0.023	0.009	0.062	0.093	0.117	0.132	0.163	25.8
60-74 years old	0.028	0.019	0.067	0.101	0.032	0.005	0.065	0.137	0.142	0.155	0.179	19.9
75- to 90-year-olds	0.037	0.020	0.062	0.100	0.042	0.003	0.075	0.141	0.133	0.132	0.142	29.0
Associative SD												
College age	0.017	0.010	0.069	0.141	0.029	0.001	0.069	0.125	0.133	0.137	0.153	12.7
60-74 years old	0.036	0.013	0.089	0.108	0.023	0.001	0.066	0.097	0.108	0.093	0.098	20.7
75-90 years old	0.044	0.019	0.087	0.083	0.028	0.000	0.061	0.068	0.071	0.071	0.059	16.5

Note.  $a$  = boundary separation;  $z$  = starting point;  $T_{er}$  = nondecision component of response time;  $\eta$  = standard deviation in drift across trials;  $s_z$  = range of the distribution of starting point  $z$ ;  $p_o$  = proportion of contaminants;  $s_t$  = range of the distribution of nondecision times;  $\chi^2$  is the chi-square goodness of fit measure (with 33 degrees of freedom, the critical value is 50.7). For item recognition,  $v_1$  is for old high-frequency words,  $v_2$  is for old low-frequency words,  $v_3$  is for new high-frequency words, and  $v_4$  is for new low-frequency words. For associative recognition,  $v_1$  is for intact high-frequency pairs,  $v_2$  is for intact low-frequency pairs,  $v_3$  is for rearranged high-frequency pairs, and  $v_4$  is for rearranged low-frequency pairs.

associative recognition, drift rates for “new” or “rearranged” conditions were negative. These were changed to positive values and averaged with the “old” or “intact” drift rates.

Drift rates for item recognition decreased only minimally with age, although the effect was significant,  $F(2, 125) = 3.2$ . For associative recognition, the decrease in drift rates with age was much larger,  $F(2, 125) = 29.1$ .

Drift rates increased significantly as IQ increased,  $F(2, 125) = 19.0$  and  $F(2, 125) = 17.9$ , for the item and associative recognition tasks, respectively. These effects were qualified by interactions: For associative recognition, the increase in drift rate with IQ was significantly smaller for the 60- to 74-year-olds than for the college-age participants,  $F(4, 125) = 5.1$ , and the 75- to 90-year-old participants were near floor for all levels of IQ. For item recognition, the increase in drift rate with IQ was larger for the college-age participants and the 60- to 74-year-olds than for the 75- to 90-year-olds (a marginally significant interaction),  $F(4, 125) = 2.3$ ,  $p = .06$ .

It should be emphasized how bad performance was for the 75- to 90-year-old participants, worse than might be expected. Their drift rates for associative recognition and their accuracy in cued recall were substantially lower than for the 60- to 74-year-old participants, and their drift rates for item recognition were also somewhat lower, suggesting an overall memory deficit.

Moreover, their drift rates and cued recall accuracy increased only minimally with IQ. For associative recognition, there was no significant increase at all with IQ (near-zero drift rates for all levels of IQ). The only instance for which 75- to 90-year-olds' performance was not significantly worse than the 60- to 74-year-olds' was free recall, but that is likely a function of the way the free recall task was structured (with only nine items; see earlier discussion).

**Criteria settings and the nondecision component.** Figure 5 shows the effects of age and IQ on boundary separation and the

nondecision component. As a function of age, the results replicate those of earlier studies (Ratcliff et al., 2001, 2003; Ratcliff, Thapar, & McKoon, 2004; Ratcliff et al., 2006a, 2006b, 2007, 2010). Older participants were slower than younger participants because they set wider criteria and their nondecision component was longer. In associative recognition, for the older participant groups, “intact” responses were faster than “rearranged” responses, and in the diffusion model fits, this was explained by a bias in the starting point  $z$  toward the “intact” decision boundary.

The effects of age on boundary separation and the nondecision component were significant; for associative and item recognition, respectively, for boundary separation,  $F(2, 125) = 22.7$  and  $F(2, 125) = 32.4$ , and for the nondecision component,  $F(2, 125) = 37.1$  and  $F(2, 125) = 72.4$ .

In contrast, the effect of IQ was not significant for either parameter ( $F_s < 1.3$ ). The lack of effect on boundary separation suggests that the speed/accuracy trade-off settings that individuals adopt are independent of IQ.

**Diffusion model variability parameters.** In this and other research on aging using diffusion model analyses, the parameters representing variability are usually not significantly different across conditions. This is partly because the differences are small but also because they have larger variability than do the other parameters (Ratcliff & Tuerlinckx, 2002). However, in order to explain the relative speeds of correct and error responses, the variability parameters are required (Ratcliff & McKoon, 2008).

In terms of statistical tests, variability in the nondecision component (its range,  $s_t$ ) increased as a function of age for item recognition and associative recognition,  $F(2, 125) = 17.6$  and  $F(2, 125) = 5.9$ , and it increased as function of IQ for item recognition,  $F(2, 125) = 4.5$ . The standard deviation in drift across trials (normally distributed with standard deviation  $\eta$ ) increased with both age and IQ for item recognition,  $F(2, 125) = 5.7$  and  $F(2, 125) = 6.0$ , respectively, but not associative recognition. The only

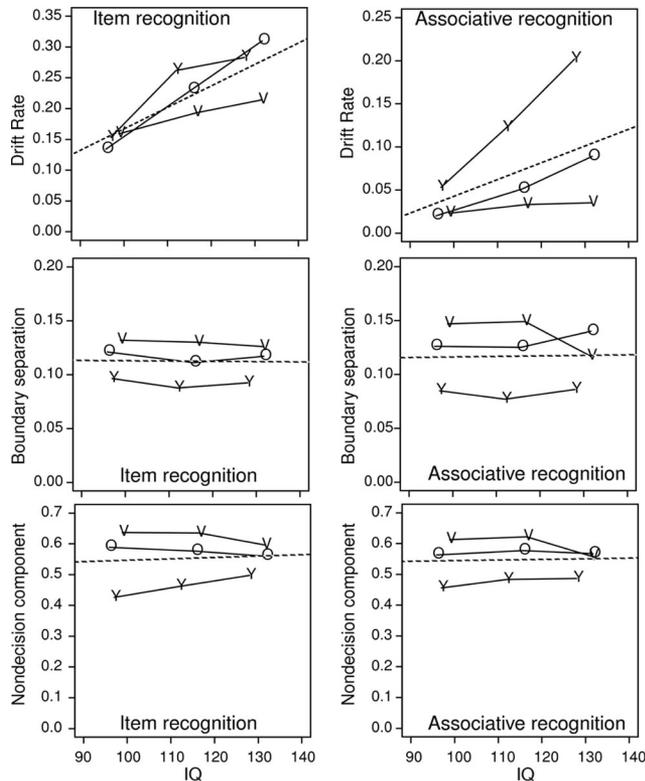


Figure 5. Values of drift rate, boundary separation, and the nondecision component as a function of IQ and age of the three groups of participants. The dotted lines are the averages over the three groups. Y = college-age participants; O = 60- to 74-year-old participants; V = 75- to 90-year-old participants.

other significant effect was a decrease in the range of the starting point ( $s_2$ ) for associative recognition as a function of IQ,  $F(2, 125) = 3.9$ . None of these effects are particularly large, so we did not attempt to interpret them further.

### Individual Differences

As discussed earlier, the values of the diffusion model parameters were averaged across participants, and the means and the standard deviations in the means are shown in Table 2. To calculate the standard deviation in a parameter value for a single participant, we relied on the Monte Carlo simulations done by Ratcliff and Tuerlinckx (2002). They generated sets of simulated data from the diffusion model for several conditions of an experiment, where the conditions differed in drift rate. For each set, they fitted the model back to the simulated data. Because of variability in the simulated data from one set to another, the parameter estimates obtained from fitting the model vary across the data sets. This variability ( $SD$ ) represents the variability in parameter estimates for a single participant. The  $SD$  values for each parameter of the model are shown in Table 4 in Ratcliff and Tuerlinckx's article for 250 observations per condition. For example, in row E, when the true value of  $a$  was 0.16, Ratcliff and Tuerlinckx found an  $SD$  of 0.012 across repeatedly fitting different simulated random samples of data generated with the same parameter values.

As the number of data points per condition increases (or decreases), the  $SD$  of single participants' parameter values decreases (or increases) as a function of the square root of the number of data points. Ratcliff and Tuerlinckx (2002) used 250 observations per condition. So to translate their estimates to our experiment, we multiplied their  $SD$  estimates by  $\sqrt{(250/N)}$ , where  $N$  was the number of observations in the conditions of our experiment. In other words, we scaled Ratcliff and Tuerlinckx's estimates of the  $SD$ s in parameter values against the numbers of observations per condition in our experiment.

In the associative recognition experiment there were about 80 observations per condition, and in the item recognition experiment there were about 190 observations per condition. Thus, using Ratcliff and Tuerlinckx's (2002) Table 4, we found  $SD$ s in individuals' parameter values, with the following results: The  $SD$  in  $a$  (for a single participant) was roughly 2.5–4 times smaller than the  $SD$  across participants (the  $SD$ s that are shown in Table 2); the  $SD$  in  $T_{er}$  was roughly 3–5 times smaller than the  $SD$  across individuals, and the  $SD$  in  $v$  was roughly 3 times smaller than the  $SD$  across individuals (except for the older participants in associative recognition, for whom drift rates approached a floor of zero).

These results mean that one can examine individual differences. For each parameter of the model, the  $SD$  in its value for a single participant is much smaller than the  $SD$  in the parameter's value when that value is obtained by averaging over the means of all the participants in an experiment.

The relatively small  $SD$ s for individual participants are especially noteworthy because there was only a single, 45-min session of data collection. The data from a single session produce parameter estimates that have small enough variability relative to differences among individuals (i.e., enough power) to allow meaningful individual difference analyses to be carried out on the main components of the model: boundary separation, nondecision time, and drift rates.

### Correlations Among Model Parameters and Data

We conducted correlational analyses among accuracy, RTs, IQ, and the main parameters of the diffusion model—drift rate, criterion setting, and the nondecision component. To obtain accuracy and median RT values for item and associative recognition, we averaged the accuracy values and median RTs for correct responses over conditions to give a single value of accuracy and a single value of median RT for each participant. The correlations are shown in Tables 3 through 6. A correlation above .29 is significant at the .05 level for a group of 45 participants (the numbers of participants in the three groups were 46, 45, and 43 for college age, 60- to 75- to year-olds, and 75- to 90-year-olds, respectively). Generally, correlations for the 75- to 90-year-old group were lower than for the other groups.

The results are easy to summarize. They fell into two sets of intercorrelations, with few exceptions. First, accuracy values for the four tasks, drift rates for the item and associative recognition tasks, and IQ were all significantly correlated with each other. Second, for the item and associative recognition tasks, RTs, criteria settings, and the nondecision component were all significantly correlated between the two tasks. And third, there were no significant correlations between the measures of these two groups. In the next paragraphs, we discuss these results in more detail.

Table 3 shows correlations between IQ, accuracy of cued recall, and accuracy of free recall, on the one hand, and accuracy and

Table 3  
*Correlations Between Cued and Free Recall, IQ, and Accuracy and Median RT for Item and Associative Recognition*

Group and measure	Accuracy		Median RT	
	Item recognition	Associative recognition	Item recognition	Associative recognition
College age				
IQ	.656	.533	.245	.313
Cued recall	.775	.616	.080	.129
Free recall	.592	.454	.290	.185
Accuracy item recognition		.687	.340	.293
Accuracy associative recognition			.300	.430
Median RT item recognition				.596
60–74 years old				
IQ	.540	.657	–.092	–.471
Cued recall	.767	.747	–.003	–.413
Free recall	.620	.721	–.073	–.532
Accuracy item recognition		.745	.023	–.442
Accuracy associative recognition			.136	–.379
Median RT item recognition				.718
75–90 years old				
IQ	.333	.331	–.281	–.387
Cued recall	.388	.325	–.152	–.074
Free recall	.400	.575	–.180	–.275
Accuracy item recognition		.597	.231	–.210
Accuracy associative recognition			.038	–.307
Median RT item recognition				.639

Note. RT = response time.

median RT in item and associative recognition, on the other hand. For the college-age and 60- to 74-year-old groups, the mean correlations between IQ and accuracy values for the four tasks were all positive and high. The correlations between the median RTs for the two recognition tasks were greater than .59. The correlations between median RTs and IQ, and median RTs and accuracy, were different for the different participant groups. There were positive correlations between the accuracy measures and the RT measures for the college-age participants. But this flipped to largely negative values for the older groups. This suggests that the better performing and higher ability young participants took longer to make decisions (adopting more conservative decision criteria). But for the older participants, who adopted more conservative decision criteria (Starns & Ratcliff, 2010), better accuracy and ability produced shorter RTs.

Table 4 shows correlations of IQ, drift rates for item and associative recognition, and accuracy values for cued and free recall for the three participant groups. Scatter plots, histograms of the values, and the correlations are also shown in Figure 6 for the 60- to 75- to year-old participant group. The plots are presented to illustrate the correlations. The main result is that for the college-age participants and the 60- to 74-year-olds, all these measures correlated significantly with each other, with correlations greater than .36.

For the 75- to 90-year-old participants, the correlations were lower, but this was because of floor effects in associative recognition and cued recall. The correlations between the two IQ measures for the 75- to 90-year-olds were also lower, partly because the range of the Matrix Reasoning scores was much smaller than for the other two groups (see Table 1), and so the range was compressed relative to the *SD*.

We might not have expected to see declines in the correlations for the 75- to 90-year-olds relative to the other groups between item recognition drift rate, IQ vocabulary score, and free recall because these quantities did not decline with age. The correlations between the IQ vocabulary score, and both the item recognition drift rate and free recall, are smaller than for the younger groups.

Table 4  
*Correlations Between Raw IQ Matrix Reasoning and Vocabulary Scores and Drift Rates for Item and Associative Recognition and Cued and Free Recall*

Group and measure	IQ vocabulary	Item drift	Associative drift	Cued recall	Free recall
College age					
IQ matrix	.602	.362	.524	.420	.459
IQ vocabulary		.604	.675	.683	.648
Item drift			.628	.595	.405
Associative drift				.724	.574
Cued recall					.614
60–74 years old					
IQ matrix	.690	.666	.497	.572	.613
IQ vocabulary		.581	.473	.678	.488
Item drift			.718	.745	.719
Associative drift				.729	.519
Cued recall					.652
75–90 years old					
IQ matrix	.200	.267	.176	.225	.247
IQ vocabulary		.343	.282	.390	.308
Item drift			.483	.208	.580
Associative drift				.279	.376
Cued recall					.450

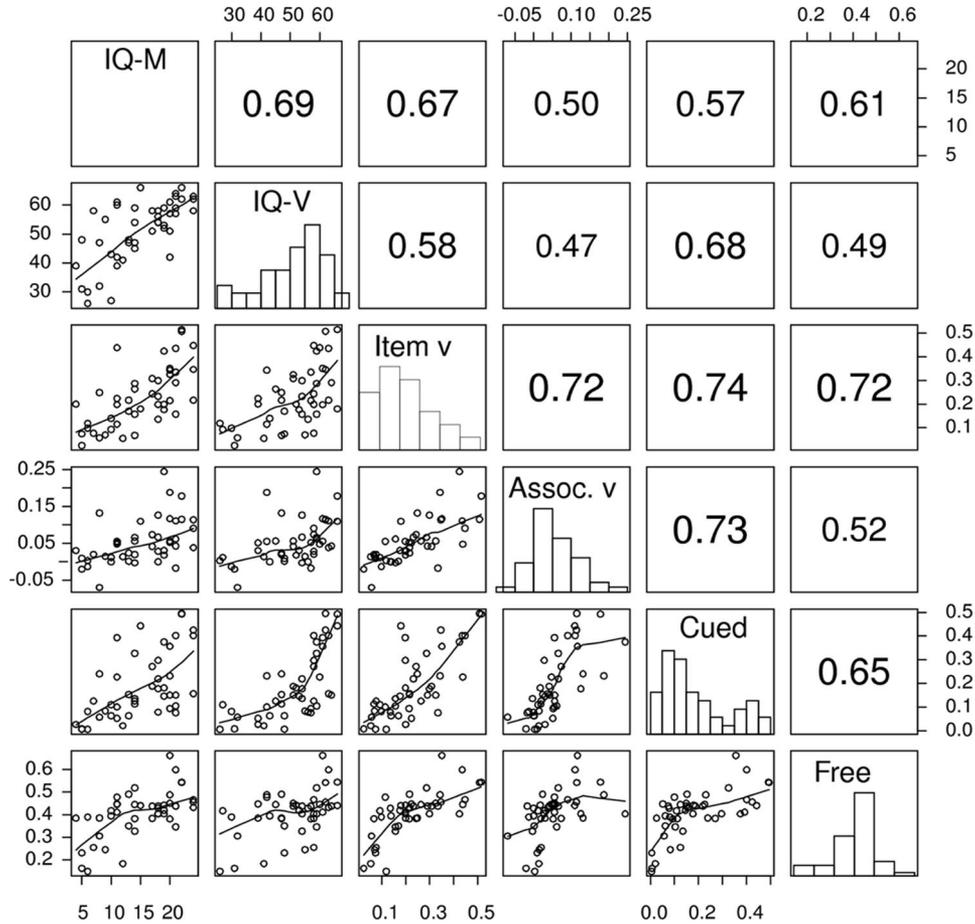


Figure 6. Scatter plots, histograms, and correlations for the Matrix Reasoning (IQ-M) and Vocabulary (IQ-V) subtests of the Wechsler Adult Intelligence Scale—Third Edition (Wechsler, 1997), as well as item and associative recognition drift rates (Item v and Assoc. v, respectively) and cued and free recall accuracy (Cued and Free, respectively).

But the correlation between item recognition drift rate and free recall is high. This suggests that the reduction in the size of the correlations in Table 4 for the 75- to 90-year-old group relative to the other groups is due to two things: floor effects in both associative recognition and cued recall and a reduction in the reliability of the IQ vocabulary score. The correlation between item recognition drift rate and free recall is higher than that for college-age participants, which means that there is no general decline in the correlations across every pair of tasks.

For item recognition (see Table 5), there were significant correlations between accuracy and drift rate and between median RT and both boundary separation and the duration of the nondecision component. The correlations between IQ and boundary separation and between IQ and the nondecision component were not consistent across participant groups (though some were just significant). For associative recognition, the same pattern was obtained, but the correlations were usually smaller.

Table 6 shows correlations of boundary separation and the nondecision component of processing between item and associative recognition and with cued and free recall accuracy. Boundary separation and the nondecision component correlated between

item and associative recognition. There were weaker correlations between boundary separation in one recognition task and the nondecision component in the other task. Few of the other correlations were significant.

Overall, the pattern of correlations suggests first that IQ, accuracy in the two recall tasks, and drift rates in the two recognition memory tasks may all represent a common memory ability that drives individual differences. Second, the finding that nondecision duration and boundary separation are not strongly related to the memory measures or to IQ (see Table 5) suggests that they are governed by different factors from the memory measures. Third, the high correlations between the two recognition tasks for boundary separation and the nondecision component suggest that these reflect some common decision mechanism, a mechanism separate from the drift rates that drive the decision process.

## Discussion

Our goal was to examine item recognition and associative recognition and to constrain explanations of performance on these

Table 5  
Correlations Within Tasks

Task, group, and measure	<i>a</i>	$T_{er}$	$v$	Accuracy	Median RT
Item recognition					
College age					
IQ	.135	.370	.522	.656	.245
<i>a</i>		-.028	.277	.215	.348
$T_{er}$			.543	.404	.498
$v$				.696	.255
Accuracy					.340
60–74 years old					
IQ	-.167	-.288	.638	.540	-.092
<i>a</i>		.407	.058	-.215	.640
$T_{er}$			-.180	-.197	.745
$v$				.738	.059
Accuracy					.023
75–90 years old					
IQ	-.135	-.291	.352	.333	-.281
<i>a</i>		.332	-.075	-.034	.580
$T_{er}$			-.138	-.110	.547
$v$				.458	-.080
Accuracy					.313
Associative recognition					
College age					
IQ	-.007	-.209	.169	.533	.256
<i>a</i>		.332	-.075	-.075	-.256
$T_{er}$			-.138	-.013	-.124
$v$				-.285	.143
Accuracy					.399
60–74 years old					
IQ	.131	-.079	.517	.657	-.471
<i>a</i>		-.015	-.105	.173	.470
$T_{er}$			.168	.231	.303
$v$				.707	-.490
Accuracy					-.345
75–90 years old					
IQ	-.320	-.171	.240	.331	-.387
<i>a</i>		.291	.143	.102	.579
$T_{er}$			.275	-.051	.446
$v$				.543	-.244
Accuracy					-.305

Note. *a* = boundary separation;  $T_{er}$  = nondecision component of RT;  $v$  = drift rate; RT = response time.

two tasks by jointly varying age and IQ. At the same time, we wanted to better understand the interactions of age and IQ.

First, before addressing these issues, we needed to show that the diffusion model explained the data well for the two recognition tasks—which it did for both item and associative recognition. The quality of the fit of the model to the data is shown by chi-square tests and by data such as those in Figure 4, where predicted and experimental values of accuracy and quantile RTs are plotted against each other.

Associative recognition provided an important test of the diffusion model in that the model had not previously been applied to associative recognition tasks. The fact that the model was successful expands greatly the theoretical and empirical tasks and issues that the model can address.

It is important that the model successfully predicts both accuracy and RT data. In many previous studies, item recognition has been compared with associative recognition only in terms of

accuracy, ignoring RTs. However, understanding the data from the experiment presented here, and likely from most other experiments, would not be possible without taking both measures into account. In the experiment presented here, we found that accuracy and RTs did not correlate significantly.

If one were to consider accuracy measures alone, then the conclusions would be as follows: no deficit in item recognition with age, a deficit in associative recognition with age for higher IQ participants, and a general deficit for lower IQ participants. In contrast, if one were to consider RT measures alone, then the conclusion would be a general speed-of-processing deficit for both item and associative recognition. The diffusion model allows these two patterns to be reconciled and explained within a single framework.

For the discussion in the next few paragraphs, we describe the data in terms of the components of processing the model abstracts from accuracy and RT data: the quality of the evidence on which a decision is based (drift rate), the amount of evidence that is required before a decision is made (boundary settings), and the nondecision component (which includes such processes as encoding the stimulus and response execution).

One striking result concerns the behavior of drift rates as a function of age and IQ. Drift rates for associative recognition for high-IQ participants decreased dramatically with age—to the point that their drift rates were nearly equivalent to those of low-IQ participants, for whom drift rates were only a little above zero at all age levels. This result is even more striking because it differed from the pattern for item recognition; for item recognition, the advantage of high-IQ participants over low remained approximately constant with age.

In more detail, this result is demonstrated in the interactions of age and IQ (see Figure 4). For item recognition, drift rates increased with IQ for all three age groups: doubling for the college-age participants and 60- to 74-year-olds (from about .15 to about .30) and increasing somewhat less for the 75- to 90-year-olds, from about .15 to about .20.

In contrast, for associative recognition, drift rates increased with IQ differentially for the three age groups. For college-age partic-

Table 6  
Correlations of *a*,  $T_{er}$ , and Cued and Free Recall

Group and variable	$T_{er}$ item	<i>a</i> associative	$T_{er}$ associative	Cued recall	Free recall
College age					
<i>a</i> item	-.028	.541	.088	.101	.292
$T_{er}$ item		.182	.554	.238	.186
<i>a</i> associative			.160	.168	.408
$T_{er}$ associative				.202	.306
60–74 years old					
<i>a</i> item	.407	.789	.008	-.117	-.178
$T_{er}$ item		.381	.539	-.240	-.294
<i>a</i> associative			-.015	.137	-.040
$T_{er}$ associative				.048	.128
75–90 years old					
<i>a</i> item	.332	.590	.176	-.142	-.162
$T_{er}$ item		.441	.456	-.062	-.260
<i>a</i> associative			.291	-.213	-.265
$T_{er}$ associative				-.032	-.025

Note.  $T_{er}$  = nondecision component of response time; *a* = boundary separation.

ipants, the increase with IQ was about fourfold (about .05 to about .20). For the 60- to 74-year-olds, the increase was about .05 to about .10, only a twofold increase, and for the 75- to 90-year-olds, IQ had no significant effect at all, with drift rates about .03 for all levels of IQ. We discuss this contrast between item and associative recognition next.

Figure 7 collapses over IQ to summarize the impact of age on drift rates for item and associative recognition and accuracy for cued and free recall. Each line represents performance of the 60- to 74-year-olds or the 75- to 90-year-olds relative to performance of the college-age participants (normalized to 1). To a good approximation, as expected, age affected associative memory but not item memory. The figure shows a drop in drift rates and accuracy from the college-age participants to the 75- to 90-year-olds in all four tasks. The decline for item recognition is much smaller than the decline for associative recognition and cued recall, and there is little decline from college-age participants to 60- to 74-year-olds in item drift. (As noted earlier, the small decline for free recall was likely the result of the task being easy, with only eight words to recall, each studied twice.)

One question that might arise is: Could younger adults be more accurate than older adults if they adopted more conservative decision criteria (for speed–accuracy manipulations, see Ratcliff et al., 2001, 2003; Ratcliff, Thapar, & McKoon, 2004). Figure 1 showed that accuracy is quite similar for the college-age participants and the 60- to 74-year-olds, but the latter adopt wider decision criteria. If the college-age participants adopted more conservative decision criteria, their RTs would slow and their accuracy would increase. It is instructive to see how large the effects might be. We took the parameter values from the college-age participants (see first line in Table 2) and generated predictions

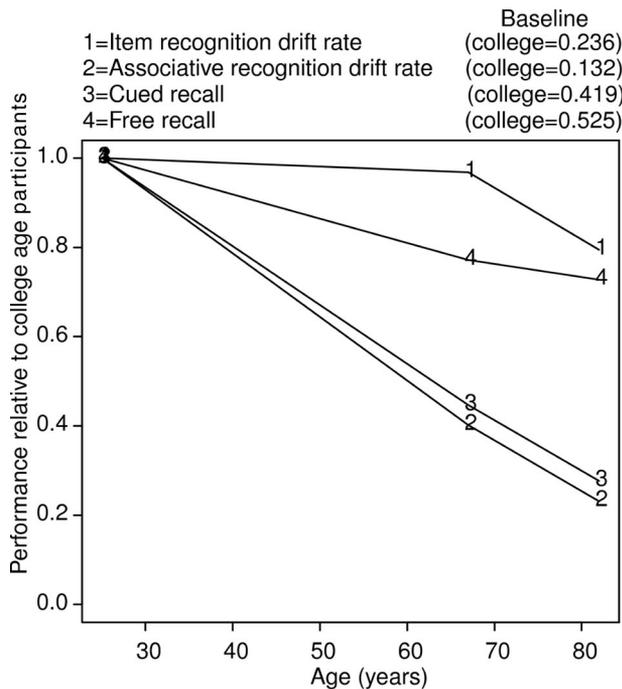


Figure 7. Relative values of drift rates for item and associative recognition, and accuracy values for cued and free recall, as a function of age.

for accuracy values and mean RTs. For the four conditions, accuracy values were 0.73, 0.86, 0.79, and 0.84 and mean RTs were 602, 586, 606, and 597 ms. Then we repeated generating the predictions but substituting the boundary separation and starting point values for the 60- to 74-year-olds (see second line in Table 2). With the boundaries increased, accuracy values were 0.74, 0.89, 0.88, and .92 with mean RTs 720, 673, 684, and 664 ms. (There is a small amount of bias in the college-age participants' value of  $z$ , which makes the accuracy differences between the first two and second two conditions not quite symmetric). The conclusions drawn from this exercise are that moving boundaries out increases accuracy by only about 4% but changes mean RT by about 89 ms. Thus changes in speed–accuracy settings have a relatively large effect on mean RT but a relatively small effect on accuracy. In Figure 1, the increase in accuracy is hardly noticeable except for perhaps the condition for new high-frequency words (the third condition).

We note in passing that our results cannot be described by the phrase “use it or lose it,” which is commonly used as a description of an approach to preventing cognitive decline (Hultsch, Hertzog, Small, & Dixon, 1999). “Use it or lose it” would predict that higher IQ participants' drift rates would decline less with age than would lower IQ participants' for both item and associative memory. This is because high-IQ participants likely engage in more intellectual activity (e.g., reading books, solving crossword puzzles) on a daily basis than do low-IQ participants. However, this was not the case: Drift rates suffered more for the higher than the lower IQ participants in associative recognition, although less so for item recognition. The lack of significantly differential decline for item information is consistent with work by Lindenberger and Baltes (1997); Rabbitt, Chetwynd, and McInnes (2003), and Singer, Verhaeghen, Ghisletta, Lindenberger, and Baltes (2003), who found little differential decline as a function of ability (but see Deary, MacLeanan, & Starr, 1998, who did find evidence for differential age-related declines as a function of ability).

Turning to the other components of processing—boundary separation and the nondecision component—the effects of age were similar to those in other published results (Ratcliff et al., 2001, 2003; Ratcliff, Thapar, & McKoon, 2004; Ratcliff et al., 2006a, 2006b, 2007, 2010). For both item and associative recognition, the distance between the boundaries was larger for the older participants than for the college-age participants, and the duration of their nondecision component was longer. However, neither boundary separation nor the nondecision component was affected by IQ (see also Ratcliff et al., 2010).

Starns and Ratcliff (2010) provided an explanation for the difference in the decision criteria adopted by younger and older adults. They found that younger adults are more likely to focus on getting the maximum number of responses correct per unit time, and to do this they are willing to sacrifice some degree of accuracy. In contrast, older adults are unwilling to make errors that they could avoid, that is, responses for which they might realize a few moments later that they were incorrect.

We calculated, in addition to the results already described, correlations among all the performance measures and all the components of the model. The first result was that the model separated components of processing and measures of performance into two groups: Drift rates correlated with accuracy and IQ but not RTs, and boundary separation and the nondecision component corre-

lated with RTs but not accuracy or IQ. This provides a clear dissociation between components of processing in the model and age and individual differences.

The second result was that item and associative recognition appear to share those mechanisms of decision making that are outside of drift rate: For both boundary separation and the nondecision component, there were strong correlations between the two tasks.

The third, and most interesting, result was the high correlations between item and associative drift rates, around .5 to .6. This suggests that the two tasks rely on shared information in memory. But this result conflicts with the finding described earlier, namely that drift rates in item recognition change little with age, whereas drift rates for associative recognition decline substantially. We discuss this conflict in the next section.

The failure to find significant correlations between IQ and RTs might be surprising. It might be expected that RTs would slow as IQ decreased (see e.g., Detterman, 1987), but this was not the case in this study. The diffusion model explains this finding: Different participants adopt different speed–accuracy criterion settings (i.e., different distances between the two boundaries; Ratcliff et al., 2010), and these settings are not a function of IQ. If participants could be encouraged to adopt similar speed–accuracy settings, then the correlation between IQ and RT would be greater.

## Memory Models

Concerning hypotheses about memory systems, the data from the item and associative recognition tasks provide a challenging pattern for models to accommodate. A model must provide correlated levels of drift rates between the two tasks, but it also must dissociate drift rates in item recognition from drift rates in associative recognition (and cued recall) as a function of age.

Models that base item and associative recognition on exactly the same information in memory have problems because they cannot accommodate the differential changes with age. At the same time, models that base item and associative recognition on completely different information in memory have problems because they cannot accommodate the high correlations between drift rates for the two tasks. The main conclusion is that memory models must allow for differential memory strength for item and associative information as well as large correlations between item and associative information.

We consider, as examples of memory models, Murdock's (1982) model and Shiffrin and Steyvers's (1997) model. In Murdock's model, items are represented as vectors of features. When a memory is formed, item vectors are added into a memory vector. The probability that a particular feature will be added into memory is a parameter of the model (e.g., the probability is larger with longer study times per item). An association is represented as a convolution between two-item vectors, which itself is a vector. The probability with which features from the convolution are encoded is different from the probability for item information. Thus, in the model, associative information is weighted separately from item information. This would allow the model to accommodate the sparing of item information along with the large drop in associative information with age by assuming that associative information is not stored as well, and the difference becomes greater with age. However, there is no a priori explanation for why

this would happen. The model would also have to add a relationship across individuals between the probability of storing a feature of an item and storing a feature from an association in order to account for the high correlations between item and associative information in each age group.

In Shiffrin and Steyvers's (1997) retrieving effectively from memory model, items are encoded as vectors of features. To form pairs, two items are concatenated. Associative recognition is modeled by matching a test pair against memory, and item recognition is modeled by matching a single item against memory. In this model, it is difficult to see how item and associative information could decline differentially as a function of age. However, hypotheses about associative information have not yet been fully developed and investigated.

The main problem with using memory models to address issues of age and individual differences is that, to our knowledge, there have been no attempts to do so. In fact, in the cognitive modeling literature in general, there have been few attempts to examine and understand individual differences, especially across different participant populations and how they might be similar or different across tasks.

## Dual-Process Models

The research reported here is relevant to a recent view of memory that is quite different from the global memory models, a view that has been labeled *dual process*. Unlike the global memory models for which item and associative recognition depend on the same representation of information in memory, in dual-process models item recognition is based on two different representations that are stored separately. In setting the context of our research in terms of the global models, we argue against the dual-process approach. In the dual-process approach, a contrast is drawn between “familiarity” and “recollection.” Subjectively, recollection is a process by which details of a study episode are available to support recognition performance. Familiarity is simply a feeling of “knowing” that an item was presented in an earlier study episode.

In support of the dual-process view, Jacoby (1991; Jacoby & Kelley, 1992) developed a method for separating familiarity and recollection, the “include/exclude” method. The paradigm is an item recognition experiment in which participants study two lists of words, followed by a test list. In the “include” test condition they respond “old” if a word appeared on either list, and in the “exclude” test condition they respond “old” only if the word appeared in one of the lists but not the other. Correct “old” responses in the include condition are taken to rely on both familiarity and recollection, and incorrect “old” responses in the exclude condition are taken to be failures of recollection. The probability that responses depend on recollection information is then estimated as  $P(R) = P(I) - P(E)$ , where  $P(I)$  is the probability of an “old” response in the inclusion condition and  $P(E)$  is the probability of an incorrect “old” response in the exclusion condition. The probability that items are recognized on the basis of familiarity is  $P(F) = P(E)/[1 - P(R)]$ . Although we disagree that this paradigm allows the separation of two distinct processes, it is likely the case that exclude decisions do require the use of recollection. It is the conclusions drawn about item recognition with which we disagree.

One problem with this process dissociation method, pointed out by Ratcliff, Van Zandt, and McKoon (1995), is that it cannot fail to produce separate, independent estimates of familiarity and recollection. There are two parameters in the model and two data points. The method will always produce a split of data between familiarity and recollection whether that is appropriate or not, and so the method is not falsifiable.

Ratcliff et al. (1995) used simulated data to demonstrate this problem. They showed that the method cannot be used to determine whether two sorts of information or only one underlie performance, nor, if there are two, can it guarantee accurate recovery of their relative contributions. Ratcliff et al. performed simulations for which performance was determined by only one process (using the search of associative memory model; Gillund & Shiffrin, 1984). Applying the process dissociation method still yielded separate estimates of familiarity and recollection. One conclusion that has been taken from this demonstration is that a single-process model can always explain results from experiments used to support dual-process models. However, that is not exactly the claim that was made by Ratcliff et al. Rather, what was claimed was that the process dissociation method cannot be used to determine whether both familiarity and recollection or only one of them is responsible for performance.

A major point of contention has been the use of receiver operating characteristics (ROC) functions and their  $z$ -transformations to argue for and against dual-process and single-process models of item recognition. Yonelinas (1994) extended the dual-process model to confidence judgments and ROCs. In his model, the familiarity process is represented by an equal-variance signal-detection model (if equal variances are not assumed, the model is not identifiable). Recollection is represented by an all-or-none (high threshold) process. If recollection remains constant at all levels of the criterion setting and if familiarity and recollection are independent, then  $P(\text{"yes"}|\text{"old"}) = P(R) + P(F > \text{crit}) - P(R)P(F > \text{crit})$ , where  $R$  signifies recollection and  $F$  signifies familiarity. By moving the criterion, an ROC function can be swept out. The result is  $z$ -transformed ROC functions that are often approximately linear with slopes less than one. However, Ratcliff et al. (1995) showed that the key assumption that recollection is constant across levels of criterion settings is not supported by data.

Since this initial work, a large amount of research has been conducted, and the conclusions vary: Some support the dual-process view, others provide a different version of it, still others assume that only a single dimension (not separate familiarity and recollection) underlies performance. Some models use a single source of familiarity based on an unequal-variance signal-detection model to explain  $z$ -ROC functions (see the debate in Cohen, Rotello, & McMillan, 2008; DeCarlo, 2002; Dunn, 2004, 2008; Rotello, MacMillan, & Reeder, 2004; Starns & Ratcliff, 2008; Wixted, 2007; Wixted & Stretch, 2004; Yonelinas, 1997). For some researchers, there are two kinds of information that are independent of each other; for others, there is a continuous dimension between the two kinds; and for others, there is only one kind of information.

There have been two recent studies that challenge the dual-process view. Ratcliff and Starns (2009) designed a model for recognition memory that can explain the RTs of confidence judgments for recognition memory, as well as the probabilities with which responses in the different confidence categories are made. In

their model, the decision process is composed of racing diffusion processes, one for each confidence category.  $z$ -ROC slopes less than 1 can be explained in terms of differences in decision criterion settings, even if the memory familiarity distributions for old and new items have equal  $SD$ s.

Starns, Ratcliff, and McKoon (2010) conducted an item recognition experiment that manipulated speed/accuracy instructions, the proportions of old and new test items (five different proportions), word frequency, and number of repetitions. Four participants were tested for 20 sessions each. When instructions stressed speed, most responses were made under 570 ms and the  $z$ -ROC slope was less than 1. To produce a less-than-1 slope, Yonelinas's (1994) dual-process model requires that familiarity and recollection both contribute to performance. However, responses were too fast for recollection to contribute. It follows that the unity or nonunity of  $z$ -ROC slopes does not provide a diagnostic for single-versus dual-process theories.

Both the Ratcliff and Starns (2009) and Starns et al. (2010) studies pushed this domain of research to jointly account for RT distributions and ROC functions. Attempting to explain both will produce a new set of modeling issues that will supersede both accuracy-based signal-detection models and dual-process models, and this will require complete reworking of those models or even completely new models.

There are several other empirical results that challenge familiarity/recollection dual-process hypotheses. Experiments by Gronlund and Ratcliff (1989; also Ratcliff & McKoon, 1982) mapped out the time course of availability of information using a response-signal procedure. Participants studied lists of pairs of words, with a test list of pairs following each study list. In one condition, the test pairs were made up of either two old words, one old word and one new word, or two new words, and participants responded "old" if the two words had appeared on the study list and *new* if they had not, irrespective of whether they had occurred in the same pair at study. For associative recognition, they responded "intact" if the two words had been studied in the same pair and "rearranged" otherwise. To examine the time course of processing, a response signal was given after each test pair was presented, with the signals varying from 50 ms to 2,500 ms after the pair. Participants were instructed to respond immediately at the signal.

Gronlund and Ratcliff's (1989) data raise two issues for dual-process hypotheses. First, if associative recollection information was used in the item recognition task, then accuracy for item recognition should have been better for intact than rearranged pairs. Although Gronlund and Ratcliff did find such a difference, it was small ( $d' = 0.34$ ), and it was much smaller than the difference in accuracy between intact and rearranged pairs when the task was associative recognition (for which the  $d'$  difference varied between 1.24 and 1.96 in different experiments). Clark and Shiffrin (1992) found a similar result: a larger difference in accuracy between intact and rearranged pairs when the task was associative recognition than when it was item recognition. Taking the data from Gronlund and Ratcliff's study at face value, there are between 1.24 and 1.96  $d'$  units of associative information available, yet only 0.34 units are used in item recognition. If participants were using an explicit recall strategy, then one recall attempt would be needed for intact pairs, whereas two successful attempts would be needed for rearranged pairs (to make sure each word in the pair had an associate). The probability of two successful recall

attempts would be lower than that of one attempt, and the difference would depend on the cued recall probability. The findings of both studies suggest, contra simple dual-process hypotheses, that participants had more associative information available to them than was used in item recognition.

The second issue raised by Gronlund and Ratcliff's (1989) study was that item and associative information became available at different points in time. Accuracy for decisions requiring only item information was above chance by 330 ms. Accuracy for decisions requiring associative information did not begin to rise above chance until after 570 ms (see also McElree, Dolan, & Jacoby, 1999; Ratcliff & McKoon, 1989).

Like Gronlund and Ratcliff (1989), Clark and Shiffrin (1992), and Starns et al. (2010), Gillund and Shiffrin (1984) found data inconsistent with associative recollection information being used in item recognition decisions. They investigated the effects of three variables that would be thought to affect recollection (or *search processes*, as they termed the recollective process): the number of times an item was repeated in a study list, the depth of encoding of the studied items, and the type of distractor that was used in the test list. They also manipulated participants' decision times. According to the dual-process view, larger effects of the three variables should have been observed when decision time was long, but instead there were no interactions of the three variables with short versus long decision times.

The results of all of these studies strongly suggest that when associative information might be used to help an item recognition decision, it is not. The failure to use associative information might occur because it requires more active effortful retrieval processes, processes that would be used only if a task required this sort of information (as in "intact"/"rearranged" decisions).

Buchler, Light, and Reder (2008) conducted a study that found that participants can access different sources of information in associative recognition; for example, they can access information about each word of a pair separately, and they can access information about the association. The results were interpreted in terms of a dual-process model (source of activation confusion [SAC]; Reder et al., 2000), but unlike the dual-process models based on process dissociation, SAC has structure and makes predictions across a range of experimental variables. However, like the global memory models, it is mute on the relationship between accuracy and RT.

The dual-process models based on process dissociation contrast with global memory models in terms of their goals. The global models attempt to explain the effects of many experimental variables on performance in many experimental tasks. In contrast, dual-process models focus on explaining performance in terms of two processes that rest on the two different sorts of information, familiarity and recollection. These dual-process models, as currently instantiated, do not specify processes or representations in detail, as the global memory models do, and so they typically do not produce predictions about how performance will change as a function of independent variables or how performance on one task relates to another.

Furthermore, the dual-process models provide no account of the interdependence between the two dependent variables, RT and accuracy, in either item recognition or associative recognition tasks. In contrast, the diffusion model (Ratcliff, 1978) explains the

two variables simultaneously. Decisions are based on a single, continuously available source of information.

There is also a variety of electroencephalograph (EEG) and functional magnetic resonance imaging (fMRI) data that address the issue of single-process versus dual-process views of item recognition. In EEG studies, event-related potential (ERP) components have been identified by subtracting the ERP signals of one brain region from the ERP signals of another as a function of experimental variables such as studied/not studied, correct/incorrect, and encoding strength (e.g., deep versus shallow encoding; Rugg, 1995; Rugg & Curran, 2007). Two ERP components have been reliably estimated in several studies: a frontal component, identified with familiarity, that occurs in the range of 300–500 ms after onset of a test item, and a parietal component, identified with recollection, that occurs in the range of 400–500 ms after onset (Eichenbaum, Yonelinas, & Ranganath, 2007; but see MacKenzie & Donaldson, 2007; Yovel & Paller, 2004). Many of these studies have shown that different brain areas become active during item recognition and that these areas seem to correspond to different kinds of information. However, activity in a particular region of the brain does not mean that information corresponding to that activity is used in making a decision.

In another line of research, a *mnemonic accumulator* hypothesis has been proposed, mainly on the basis of fMRI evidence. The hypothesis is that there is a link between activity in the posterior parietal cortex and memory retrieval (Cabeza, Ciaramelli, Olson, & Moscovitch, 2008; Konishi, Wheeler, Donaldson, & Buckner, 2000; Wagner, Shannon, Kahn, & Buckner, 2005). One possibility is that activity in the parietal regions represents the integration of memory-strength signals from separate processes and/or brain areas for familiarity and recollection. However, because of the temporal uncertainty in the fMRI signal, it is not clear when and what are the sources of information actually entering the decision process.

Most recently, in an fMRI study, Johnson, McDuff, Rugg, and Norman (2009) provided a multivoxel pattern-classification analysis of an item recognition task. In this task, participants were required to use a scale that was made up of a remember judgment and four confidence judgments to be used if a remember judgment was not made. They trained pattern classifiers on the fMRI signals for three encoding tasks for words that denote objects (describe how an artist would draw the object and rate the difficulty, report how many functions the object has, and report how difficult it would be to read the word backward). At test, they found reinstatement of the encoding patterns both when the participants indicated that they remembered the item and when they had only a feeling of familiarity. They argued that both sources of information are present and combined in making a recognition judgment and that even recollective information is better viewed as continuously available information as opposed to threshold information. However, they offered the caveats mentioned in the previous paragraph: that fMRI cannot be used to determine the time course of reinstatement and that fMRI signals may reflect processing that occurs after a response is made.

In sum, we see four major problems with the dual-process approach. One is that, as Ratcliff et al. (1995) showed, there is no way to falsify the claim that item recognition is composed of two processes. Even if data are generated from a single process, the process-dissociation method produces estimates of two processes

(an exception is the SAC model [Reeder et al., 2000], which predicts patterns of results across many more than two conditions; see also Norman & O'Reilly, 2003). The second problem is that experiments designed to find effects of associative information on item recognition have failed to show more than a modest contribution (Clark & Shiffrin, 1992; Gillund & Shiffrin, 1984; Gronlund & Ratcliff, 1989). The third problem is that, by focusing on separating familiarity and recollection in item recognition tasks (either behaviorally or in terms of brain activity), dual-process theory has taken a step backward. There has been little attempt to understand the actual processes underlying familiarity and recollection. Dual-process approaches contrast with the global memory models' attempts to understand how retrieval processes operate and how information is represented in memory. Fourth, the available EEG and fMRI data do not show conclusively that there are two components of processing associated with different brain regions that both contribute to item recognition decisions.

We argue that, in contrast to dual-process views that propose that recollection and familiarity both contribute to item recognition decisions, item recognition depends on only a single source of information in memory (as long as a participant is performing relatively passively, as is usually the case in the tasks relevant here). There is absolutely no doubt that, for many test items, participants could recover information about them (as in recollection) and base a decision on the recovered information. It is also not clear that a binary distinction between sources of information is the correct one. Item information and associative information are associated, as in the correlations across individuals presented earlier, but they dissociate with age. Source memory seems to be different from item information, and in the models that link item information to context (see e.g., Dennis & Humphreys, 2001), the representation of source (context) and associative information are quite different. Accessing source or associative information is a slow, effortful process, and we see little evidence for this being used routinely in simple item recognition tasks.

### Associative Deficits

Naveh-Benjamin (2000; see also Burke & Light, 1981; Chalfonte & Johnson, 1996; Spencer & Raz, 1995) has proposed that problems in binding are responsible for associative deficits with age; this has been called the associative deficit hypothesis (ADH). Binding is said to involve the creation and retrieval of links between units of information, such as the link between two items of a pair or the link between an item and its context. Examples of the sorts of links that have been studied empirically are words in a particular color, words in positions on display screens, words presented in a male or female voice, the temporal order of two items, and item pairings for verbal and nonverbal material. The ADH focuses on the distinction between memory for single units and memory for associations between units and provides an account for why item information remains relatively intact with age, whereas associative information is reduced. But the ADH does not propose specific mechanisms that might be responsible for deficits for some kinds of information but not others.

Li, Naveh-Benjamin, and Lindenberger (2005) proposed a specific, three-layer connectionist model that allows the difference between item and associative information as a function of age to be explained in a single architecture. Items of a pair are encoded into

separate vectors in the input layer, and three separate sets of hidden units map from the input vectors to two output vectors. Two sets of hidden units map the input item vectors to the output item vectors, and the third set of hidden units maps from both input vectors to both output vectors. Item recognition is represented by how well an input vector for an item is produced at output (the dot product between the input vector and the output vector), and associative recognition is represented by how well an input vector for a pair is produced at output (the dot product between the vector pair at input and output). The deficit in aging was modeled by a change in a single gain parameter that modulated the mapping from input to output (via a sigmoidal function). The model was able to account for the data from Naveh-Benjamin's (2000) Experiment 2. In that experiment, participants were instructed to study pairs as single items or as pairs, and item and associative recognition were tested.

This kind of model is a candidate to explain the differences that have been found between item and associative recognition, but it needs to be developed further and evaluated in ways that are standard in the memory models literature; for example, it needs to address catastrophic interference (McCloskey & Cohen, 1989; Ratcliff, 1990), mirror effects (Glanzer & Adams, 1990), the list strength effect (Ratcliff, Clark, & Shiffrin, 1990), and so on. For an example of a model that addresses these issues, see Shiffrin and Steyvers (1997).

### Cognitive Ability Studies

In research on IQ and cognitive abilities, intelligence has been divided into fluid intelligence and crystallized intelligence (see e.g., Carroll, 1993). To relate our results to this approach, and because we included WAIS measures of fluid intelligence (matrix reasoning) and crystallized intelligence (vocabulary), one might assign associative recognition to the fluid intelligence class of tasks (cf. Li et al., 2004) because associative recognition declines with age. Also, because item recognition declines relatively little with age, it might be assigned to the crystallized class of tasks. This classification would then predict higher correlations across individuals between WAIS matrix reasoning and associative recognition and between WAIS vocabulary and item recognition than would the other two comparisons (see Table 4). In fact, apart from the lower correlation between matrix reasoning and item recognition drift rate for college-age participants, the correlations between item and associative recognition drift rates and the two IQ measures appear to be similar across the three age groups.

Schmiedek et al. (2007) analyzed data from eight choice-RT tasks (including verbal, numerical, and spatial tasks) from Oberauer, Suß, Wilhelm, and Wittmann (2003) using an ex-Gaussian analysis (Hohle, 1965; Ratcliff & Murdock, 1976) and an EZ-diffusion model analysis (Wagenmakers et al., 2007). They found that drift rates in the diffusion model mapped onto working memory, speed of processing, and reasoning ability measures (each of these was measured by aggregated performance on several tasks). These results are similar to the relationship between IQ measures and drift rates found in this article and in Ratcliff et al. (2010). However, they also found significant correlations between drift rates, boundary separation, and nondecision time. Our study (like Ratcliff et al., 2001, 2003; Ratcliff, Thapar, & McKoon, 2004; Ratcliff et al., 2006a, 2006b, 2007, 2010) found only weak corre-

lations between boundary separation and nondecision time and no significant correlations between drift rates and either boundary separation or nondecision time. There are some important differences between the Oberauer et al. study and the study reported here. First, we used a much larger range of IQs (Oberauer et al., 2003, used university students); second, we used several age ranges; and third, we had smaller numbers of participants per group (the latter would not be responsible for the lower correlations) but many more observations for item and associative recognition. It is possible that the low number of observations and use of the EZ method in the Oberauer et al. study have produced spurious correlations in the model fits (see Ratcliff, 2008), but simulation studies would have to be conducted to evaluate this.

One of the main features of our study is that it allows decomposition of speed of processing and accuracy into model components within a unified single-model structure. This allows factors responsible for individual differences in RT and accuracy to be identified. In traditional abilities approaches, only one of these dependent variables is used in analyses. Because the different measures behave in different ways, analyses based on only one of them produce an incomplete picture of individual differences.

In individual differences applications to aging, a psychometrics approach is useful in a practical sense to identify the general categories of deficits that appear with age. However, from a theoretical perspective, it is limited because it averages over tasks and measures and so cannot provide an account of the processes involved in the individual tasks. Also, because performance is averaged over tasks, it is not a priority to understand processing differences within tasks within an "ability." This, in a sense, is the converse problem to memory models' current lack of attempt to deal with individual differences.

In other research, the effects of task and condition difficulty have been shown to interact with age and ability (see e.g., Baltes & Kliegl, 1992; Salthouse, 1992). There are undoubtedly some abilities that decline with age in a way that older adult performance cannot improve to the level of younger adults (see e.g., Baltes & Kliegl, 1992). However, other studies have demonstrated abilities that show little decline with age (Ratcliff et al., 2007). It may be that performance declines with the increased complexity of a task (see e.g., Salthouse, 1992) because resources needed to perform the task (e.g., working memory) decline with age. Our results complement this research by showing that one component of memory (item recognition) shows little decline with age, even at all performance levels, but large differences with IQ, whereas associative recognition shows large declines with IQ and age.

Our results show decreased correlations of IQ with advanced age. Results of this kind have been argued to show a depletion of resources for cognitive tasks so that as performance declines, differences based on ability (e.g., IQ) are reduced and differences based on genetic factors become a larger proportion of individual differences (see e.g., Lindenberger et al., 2008; Nagel et al., 2008). It might seem that the results presented here are consistent with this hypothesis because correlations are smaller for the 75- to 90-year-old group. But there are only three correlations that do not have measures in which performance is approaching floor (if performance is near floor, correlations are automatically reduced). These are correlation between IQ vocabulary, item recognition drift rate, and free recall proportion. Of these, the correlation of item recognition drift rate and free recall does not show a decline.

An important demonstration by our study is that behavioral data can be much richer than what is typically used in genetic studies. For example, Papassotiropoulos et al. (2006) showed that the polymorphism rs17070145 in the KIBRA gene was associated with episodic memory performance. A follow-up study by Schaper, Kolsch, Popp, Wagner, and Jessen (2008) replicated Papassotiropoulos et al.'s study, but one by Need et al. (2008) failed to replicate. In these studies, the behavioral data were limited; for example, Papassotiropoulos et al. used a single list of items for recall. Schaper et al. had participants learn a single list of 15 nouns, and these were recalled five times, then recalled 30 min later, and then recognition was tested. In the Need et al. study, one test involved list learning and recall, and another test involved story recall. In the Schaper et al. study, the effect size corresponded to a recall difference of about 13%, and in the Need et al. study, the effect size able to be detected was about 3% with power 90%, so the null result was not because of lack of power.

The assumption made in these genetic studies is that any memory task (even with extremely limited data per participant) taps into a single ability, namely episodic memory (see Thapar, Petrill, & Thompson, 1994). Our experiment shows that data in item recognition, associative recognition, and recall are multifaceted. It shows individual differences and values that change across age that are much larger than the differences reported as a function of genetic differences. Our study also shows that some aspects of performance decline at different rates as a function of age and that some are related to IQ and some are not. Interestingly, the correlations of memory measures with IQ support the notion that single tests can index general memory ability within a single age group of participants. Thapar et al. (1994) investigated the heritability of memory ability in the Western Reserve Twin Project (Thompson, Detterman, & Plomin, 1991). This study is unique in that its test battery includes measures of memory ability that tap memory span, verbal memory, and visual memory. Results showed low intercorrelations across the memory measures, and the estimates of genetic and environmental influences on memory ability varied across memory measures. It seems that genetic studies of memory are in their infancy, and the reliability of the effects needs to be more systematically investigated. Moreover, the amount and quality of the behavioral measures are impoverished relative to what are usually seen in cognitive research.

Our study also provides practical points that should be considered in conducting new research. First, researchers investigating individual differences in aging will need to carefully consider the ages of participants because the data clearly show different trends for younger older adults relative to very old adults (Lindenberger & Baltes, 1997). Mixing the two might mean averaging over different patterns of results. Second, the performance of young adults is also highly variable and impacted by ability. Although this is not necessarily surprising, it does mean that it is critical for researchers to classify young adults into ability groups when examining how ability is affected by aging. Third, the ADHD proposed by Naveh-Benjamin (2000) will need to address the performance of low-IQ young adults. To our knowledge, there has been no discussion of how low-ability young adults might struggle with associative recognition tasks.

## Conclusion

The results of this study begin to provide a theoretical framework to show how speed of processing is related to memory performance across age groups and individuals. The experiment also provides archival results for the relationship between ability measures, memory, and speed of processing in memory tasks.

This study reports the first account by a sequential sampling model of associative recognition accuracy and RT. The results show dramatic declines in memory as a function of age in associative recognition but small declines in item recognition (as is well known in the aging literature; see e.g., Naveh-Benjamin, 2000). The results also show dramatic differences in performance on both item and associative recognition as a function of IQ.

The results show strong regularities in individual differences in memory tasks. Drift rate, accuracy in recall tasks, and IQ measures are all highly intercorrelated. In contrast, boundary separation is correlated across the two tasks, as is the duration of the nondecision component, but neither is correlated with IQ or drift rates. Boundary separation and nondecision duration are affected by age but not IQ. For item recognition, the effect of age on drift rate is small, but the effect of IQ is large. The effects of age and IQ on drift rate for associative recognition are both large. In sum, the differences in item and associative recognition as a function of IQ and age present important targets for memory models.

## References

- Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy young and older adults. In E. Tulving & F. I. M. Craik (Eds.), *Handbook of memory* (pp. 395–410). New York, NY: Oxford University Press.
- Baltes, P. B., & Kliegl, R. (1992). Further testing of limits of cognitive plasticity: Negative age differences in a mnemonic skill are robust. *Developmental Psychology, 28*, 121–125.
- Bowles, N. L., & Poon, L. W. (1982). An analysis of the effect of aging on recognition memory. *Journal of Gerontology, 37*, 212–219.
- Buchler, N. G., Light, L. L., & Reder, L. M. (2008). Memory for items and associations: Distinct representations and processes in associative recognition. *Journal of Memory and Language, 59*, 183–199.
- Buchler, N. E. G., & Reder, L. M. (2007). Modeling age-related memory deficits: A two-parameter solution. *Psychology and Aging, 22*, 104–121.
- Burke, D. M., & Light, L. L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin, 90*, 513–514.
- Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: An attentional account. *Nature Reviews Neuroscience, 9*, 613–625.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytical studies*. New York, NY: Cambridge University Press.
- Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition, 24*, 403–416.
- Clark, S. E., & Shiffrin, R. M. (1992). Cuing effects and associative information in recognition memory. *Memory & Cognition, 20*, 580–598.
- Cohen, A. L., Rotello, C. M., & MacMillan, N. A. (2008). Evaluating models of remember-know judgments: Complexity, mimicry, and discriminability. *Psychonomic Bulletin & Review, 15*, 906–926.
- Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London B, 302*, 341–359.
- Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performances* (pp. 409–422). Amsterdam, the Netherlands: Elsevier.
- 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: Erlbaum.
- Craik, F. I. M., & McDowd, J. M. (1987). Recall, recognition, and aging. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 474–479.
- Deary, I. J., MacLeannan, W. J., & Starr, J. M. (1998). Is age kinder to the initially more able? Differential ageing of a verbal ability in the Healthy Old People in Edinburgh Study. *Intelligence, 26*, 357–375.
- DeCarlo, L. T. (2002). Signal detection theory with finite mixture distributions: Theoretical developments with applications to recognition memory. *Psychological Review, 109*, 710–721.
- Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review, 108*, 452–478.
- Detterman, D. K. (1987). What does reaction time tell us about intelligence? In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 177–200). Norwood, NJ: Ablex.
- Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review, 111*, 524–542.
- Dunn, J. C. (2008). The dimensionality of the remember-know task: A state-trace analysis. *Psychological Review, 115*, 426–446.
- Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30*, 123–152.
- Erber, J. T. (1974). Age differences in recognition memory. *Journal of Gerontology, 29*, 177–181.
- 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.
- Geddes, J., Ratcliff, R., Allerhand, M., Childers, R., Wright, R. J., Frier, B. M., & Deary, I. J. (2010). Modeling the effects of hypoglycemia on a two-choice task in adult humans. *Neuropsychology, 24*, 652–660.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1–67.
- Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 5–16.
- 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.
- Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 846–858.
- Healy, M. R., Light, L. L., & Chung, C. (2005). Dual-process models of associative recognition in young and older adults: Evidence from receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 768–788.
- Hockley, W. E. (1991). Recognition memory for item and associative information: A comparison of forgetting rates. In S. Lewandowsky & W. E. Hockley (Eds.), *Relating theory and data: Essays in human memory in honor of Bennet B. Murdock* (pp. 227–248). Hillsdale, NJ: Erlbaum.
- Hockley, W. E. (1994). Reflections of the mirror effect for item and associative recognition. *Memory & Cognition, 22*, 713–722.
- Hockley, W. E., & Cristi, C. (1996). Tests of encoding tradeoffs between item and associative information. *Memory & Cognition, 24*, 202–216.
- Hohle, R. H. (1965). Inferred components of reaction times as a function of foreperiod duration. *Journal of Experimental Psychology, 69*, 382–386.
- Hultsch, D. F., Hertzog, C., Small, B. J., & Dixon, R. A. (1999). Use it or lose it: Engaged lifestyle as a buffer of cognitive decline in aging? *Psychology and Aging, 14*, 245–263.

- Humphreys, M. S. (1976). Relational information and the context effect in recognition memory. *Memory & Cognition*, 4, 221–232.
- Humphreys, M. S. (1978). Item and relational information: A case for context independent retrieval. *Journal of Verbal Learning and Verbal Behavior*, 17, 175–187.
- Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review*, 96, 208–233.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513–541.
- Jacoby, L. L., & Kelley, C. M. (1992). A process-dissociation framework for investigating unconscious influences: Freudian slips, projective tests, subliminal perception, and signal detection theory. *Current Directions in Psychological Science*, 1, 174–179.
- Johnson, J. D., McDuff, S. G. R., Rugg, M. D., & Norman, K. A. (2009). Recollection, familiarity, and cortical reinstatement: A multivoxel pattern analysis. *Neuron*, 63, 697–708.
- Kaufman, S. B., DeYoung, C. G., Gray, J. R., Brown, J., & Mackintosh, N. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*, 37, 374–382.
- Kausler, D. H. (1994). *Learning and memory in normal aging*. San Diego, CA: Academic Press.
- Kelley, R., & Wixted, J. T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 701–722.
- Konishi, S., Wheeler, M. E., Donaldson, D. I., & Buckner, R. L. (2000). Neural correlates of episodic retrieval success. *NeuroImage*, 12, 276–286.
- Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Laming, D. R. J. (1968). *Information theory of choice reaction time*. New York, NY: Wiley.
- Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science*, 15, 155–163.
- Li, S.-C., Naveh-Benjamin, M., & Lindenberger, U. (2005). Aging neuro-modulation impairs associative binding. *Psychological Science*, 16, 445–450.
- Light, L. L., Patterson, M. M., Chung, C., & Healy, M. R. (2004). Effects of repetition and response deadline on associative recognition in young and older adults. *Memory & Cognition*, 32, 1182–1193.
- Lindenberger, U., & Baltes, P. B. (1997). Intellectual functioning in old and very old age: Cross-sectional results from the Berlin Aging Study. *Psychology and Aging*, 12, 410–432.
- Lindenberger, U., Nagel, I. E., Chicherio, C., Li, S.-C., Heekeren, H. R., & Bäckman, L. (2008). Age-related decline in brain resources modulates genetic effects on cognitive functioning. *Frontiers in Neuroscience*, 2, 234–244.
- MacKenzie, G., & Donaldson, D. I. (2007). Dissociating recollection from familiarity: Electrophysiological evidence that familiarity for faces is associated with a posterior old/new effect. *NeuroImage*, 36, 454–463.
- Malmberg, K. J., & Xu, J. (2007). On the flexibility and on the fallibility of associative memory. *Memory & Cognition*, 35, 545–556.
- 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.
- McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 24, pp. 109–165). New York, NY: Academic Press.
- McEree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information in item recognition: A time-course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 563–582.
- Merriam-Webster. (1990). *Merriam-Webster's ninth new collegiate dictionary* (9th ed.). Springfield, MA: Author.
- Mitchell, K. J., Johnson, M. K., Raye, C. L., Mather, M., & D'Esposito, M. (2000). Aging and reflective processes of working memory: Binding and test load deficits. *Psychology and Aging*, 15, 527–541.
- Murdock, B. B. (1974). *Human memory: Theory and data*. Potomac, MD: Erlbaum.
- Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, 89, 609–626.
- Murdock, B. B. (1992). Item and associative information in a distributed memory model. *Journal of Mathematical Psychology*, 36, 68–99.
- Nagel, I. E., Chicherio, C., Li, S.-C., von Oertzen, T., Sander, T., Villringer, A., . . . Lindenberger, U. (2008). Human aging magnifies genetic effects on executive functioning and working memory. *Frontiers in Human Neuroscience*, 2, Article 1.
- 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). *Human Memory: An introduction to research, data, and theory*. Pacific Grove, CA: Brooks/Cole.
- Need, A. C., Attix, D. K., McEvoy, J. M., Cirulli, E. T., Linney, K. N., Wagoner, A. P., . . . Goldstein, D. B. (2008). Failure to replicate effect of Kibra on human memory in two large cohorts of European origin. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, 147B, 667–668.
- Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7, 308–313.
- Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, 104, 611–646.
- Oberauer, K., SuB, H.-M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, 31, 167–193.
- Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of aging on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, 23, 104–118.
- Papassotiropoulos, A., Stephan, D. A., Huentelman, M. J., Hoerndli, F. J., Craig, D. W., Pearson, J. V., . . . de Quervain, D. J.-F. (2006, October 20). Common KIBRA alleles are associated with human memory performance. *Science*, 314, 475–478.
- Rabbitt, P., Chetwynd, A., & McInnes, L. (2003). Do clever brains age more slowly? Further exploration of a nun result. *British Journal of Psychology*, 94, 63–71.
- 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. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review*, 97, 285–308.
- 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. (2008). The EZ diffusion method: Too EZ? *Psychonomic Bulletin & Review*, 15, 1218–1228.

- Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). The list-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 163–178.
- Ratcliff, R., Love, J., Thompson, C. A., & Opfer, J. (in press). Children are not like older adults: A diffusion model analysis of developmental changes in speeded responses. *Child Development*.
- Ratcliff, R., & McKoon, G. (1982). Speed and accuracy in the processing of false statements about semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 16–36.
- Ratcliff, R., & McKoon, G. (1989). Similarity information versus relational information: Differences in the time course of retrieval. *Cognitive Psychology*, *21*, 139–155.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.
- Ratcliff, R., & Murdock, B. B., Jr. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*, 190–214.
- Ratcliff, R., Perea, M., Coleangelo, A., & Buchanan, L. (2004). A diffusion model account of normal and impaired readers. *Brain and Cognition*, *55*, 374–382.
- Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in letter identification. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 127–140.
- Ratcliff, R., Schmiedek, F., & McKoon, G. (2008). A diffusion model explanation of the worst performance rule for reaction time and IQ. *Intelligence*, *36*, 10–17.
- Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, *116*, 59–83.
- Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, *19*, 278–289.
- Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, *16*, 323–341.
- Ratcliff, R., Thapar, A., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & Psychophysics*, *65*, 523–535.
- Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*, *50*, 408–424.
- Ratcliff, R., Thapar, A., & McKoon, G. (2006a). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin & Review*, *13*, 626–635.
- Ratcliff, R., Thapar, A., & McKoon, G. (2006b). Aging, practice, and perceptual tasks: A diffusion model analysis. *Psychology and Aging*, *21*, 353–371.
- Ratcliff, R., Thapar, A., & McKoon, G. (2007). Application of the diffusion model to two-choice tasks for adults 75–90 years old. *Psychology and Aging*, *22*, 56–66.
- Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology*, *60*, 127–157.
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating the parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*, 438–481.
- Ratcliff, R., & Van Dongen, H. P. A. (2009). Sleep deprivation affects multiple distinct cognitive processes. *Psychonomic Bulletin & Review*, *16*, 742–751.
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1995). Process dissociation, single process theories, and recognition memory. *Journal of Experimental Psychology: General*, *124*, 352–374.
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, *106*, 261–300.
- Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 294–320.
- Rotello, C. M., MacMillan, N. A., & Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review*, *111*, 588–616.
- Rugg, M. D. (1995). Memory and consciousness: A selective review of issues and data. *Neuropsychologia*, *33*, 1131–1141.
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*, 251–257.
- Salthouse, T. A. (1992). Why do adult age differences increase with task complexity? *Developmental Psychology*, *28*, 905–918.
- Schaper, K., Kolsch, H., Popp, J., Wagner, M., & Jessen, F. (2008). KIBRA gene variants are associated with episodic memory in healthy elderly. *Neurobiology of Aging*, *29*, 1123–1125.
- Schmiedek, F., Oberauer, K., Wilhelm, O., Suß, H.-M., & Wittmann, W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*, 414–429.
- Schonfield, 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 & Review*, *4*, 145–166.
- Singer, T., Verhaeghen, P., Ghisletta, P., Lindenberger, U., & Baltes, P. B. (2003). The fate of cognition in very old age: Six-year longitudinal findings in the Berlin Aging Study (BASE). *Psychology and Aging*, *18*, 318–331.
- Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 101–117.
- Spencer, W. D., & Raz, N. (1995). Effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging*, *10*, 527–539.
- Starns, J. J., & Ratcliff, R. (2008). Two dimensions are not better than one: STREAK and the univariate signal detection model of remember/know performance. *Journal of Memory and Language*, *59*, 169–182.
- Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, *25*, 377–390.
- Starns, J. J., Ratcliff, R., & McKoon, G. (2010). *Adding response times to models of recognition memory receiver operating characteristics (ROCs): RTCON and the diffusion model*. Manuscript submitted for publication.
- Thapar, A., Petrill, S. A., & Thompson, L. A. (1994). The heritability of memory in the Western Reserve Twin Project. *Behavior Genetics*, *24*, 155–160.
- 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.
- Thompson, L. A., Detterman, D. K., & Plomin, R. (1991). Associations between cognitive abilities and scholastic achievement: Genetic overlap but environmental differences. *Psychological Science*, *2*, 158–165.
- Unsworth, N. (2010). On the division of working memory and long-term memory and their relation to intelligence: A latent variable approach. *Acta Psychologica*, *134*, 16–28.
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*, 1206–1220.
- Wagenmakers, E.-J., Van Der Maas, H. L. J., & Grassman, R. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*, 3–22.
- Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal

- lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, 9, 445–453.
- Wahlin, A., Backman, L., & Winblad, B. (1995). Free recall and recognition of slowly and rapidly presented words in very old age: A community-based study. *Experimental Aging Research*, 21, 251–271.
- Wechsler, D. (1997). *The Wechsler Adult Intelligence Scale—Third edition*. San Antonio, TX: Psychological Corporation.
- White, C., Ratcliff, R., Vasey, M., & McKoon, G. (2009). Dysphoria and memory for emotional material: A diffusion model analysis. *Cognition & Emotion*, 23, 181–205.
- White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010a). Anxiety enhances threat processing without competition among multiple inputs: A diffusion model analysis. *Emotion*, 10, 662–677.
- White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010b). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology*, 54, 39–52.
- Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114, 152–176.
- Wixted, J. T., & Stretch, V. (2004). In defense of the signal-detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review*, 11, 616–641.
- Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1341–1354.
- Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, 25, 747–763.
- Yovel, G., & Paller, K. A. (2004). The neural basis of the butcher on the bus phenomenon: When a face seems familiar but is not remembered. *NeuroImage*, 21, 789–800.

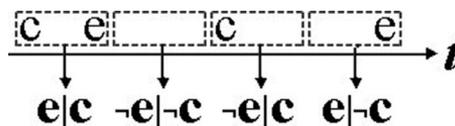
Received June 30, 2010

Revision received February 15, 2011

Accepted February 16, 2011 ■

### Correction to Greville and Buehner (2010)

In the article “Temporal Predictability Facilitates Causal Learning,” W. James Greville and Marc J. Buehner (*Journal of Experimental Psychology: General*, 2010, Vol. 139, No. 4, pp. 756–771), Figure 2 (p. 759) contained an error. The terms  $e|c$  and  $-e|c$  were mislabelled as  $-e|c$  and  $e|c$ . The corrected figure appears below.



*Figure 2.* The effect of attribution shift in parsing an event stream with a fixed temporal window:  $c \rightarrow e$  intervals that are longer than the temporal window simultaneously decrease impressions of  $P(e|c)$  and increase impressions of  $P(e|-c)$ .

DOI: 10.1037/a0024735