Journal of Memory and Language 66 (2012) 416–437

Contents lists available at SciVerse ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Aging and IQ effects on associative recognition and priming in item recognition

Gail McKoon *, Roger Ratcliff

The Ohio State University, United States

ARTICLE INFO

Article history:
Received 22 April 2011
revision received 1 December 2011
Available online 4 February 2012

Keywords:
Priming
Item recognition
Associative recognition
Aging

ABSTRACT

Two ways to examine memory for associative relationships between pairs of words were tested: an explicit method, associative recognition, and an implicit method, priming in item recognition. In an experiment with both kinds of tests, participants were asked to learn pairs of words. For the explicit test, participants were asked to decide whether two words of a test pair had been studied in the same or different pairs. For the implicit test, participants were asked to decide whether single words had or had not been among the studied pairs. Some test words were immediately preceded in the test list by the other word of the same pair and some by a word from a different pair. Diffusion model (Ratcliff, 1978; Ratcliff & McKoon, 2008) analyses were carried out for both tasks for college-age participants, 60–74 year olds, and 75–90 year olds, and for higher- and lower-IQ participants, in order to compare the two measures of associative strength. Results showed parallel behavior of drift rates for associative recognition and priming across ages and across IQ, indicating that they are based, at least to some degree, on the same information in memory.

Introduction

Memory researchers generally agree on a distinction between item information and associative information. Memory for single items is usually measured with single-item tests: participants are asked to recognize whether single stimuli, such as words or pictures, appeared earlier in the experiment. Memory for associative information is usually measured with pairs tests: participants are asked to recognize whether two items of a pair were presented earlier in the experiment in the same or different pairs. For the research in this article, we also measured memory for associative information in a second way, priming in single-word test lists: some test words were immediately preceded in the test list by the word that had appeared in their same pair at study and others were preceded by a word from a different pair.

To give specific examples, suppose that the pairs "tree carpet" and "blue jealous" were presented in a list of pairs of words to be studied. Then for an associative recognition test, the correct response to the test pair "tree carpet" would be "studied in same pair" (i.e., "intact") and the correct response to the pair "tree jealous" would be "studied in different pairs" (i.e., "rearranged"). For a single-word recognition test, the correct response to "carpet" would be "yes" (i.e., "old"). For priming, "carpet" would be primed if the immediately preceding test word (the prime) was "tree" and unprimed if the immediately preceding test word was "blue". Note that in the single-word tests, to the participants, each word would appear to be a separate test: there is nothing that signals the relationship between pairs of words.

One of the goals for the experiment we describe here was to compare associative memory as measured by associative recognition to associative memory as measured by

* Corresponding author. Address: Department of Psychology, The Ohio State University, Columbus, OH 43210, United States. Fax: +1 614 688 3984.

E-mail address: mckoon.1@osu.edu (G. McKoon).

0749-596X/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2011.12.001

priming in item recognition, and to do so in terms of their interactions with two participants variables, IQ and age (college-age, 60–74 year olds, and 75–90 year olds), and one items variable: pairs were either related (e.g., "mortician cadaver") or unrelated ("tree carpet"). If priming and associative recognition have a common basis, then parallel effects of the three variables on them would be expected. This should occur even though age and IQ affect item recognition—the task with which priming effects are measured—and associative recognition differently. In other words, the effects of age and IQ on priming and associative recognition should be similar even as their effects on item recognition might be different. Age and IQ are, therefore, good candidates for teasing apart associative information as measured by pairs tests from associative information as measured by priming.

Performance on associative recognition has usually been measured in terms of accuracy, and priming effects in item recognition have usually been measured in terms of response times (RTs). This is because the effects of priming on accuracy tend to be small and the effects on RTs tend to be large. For associative recognition, the reverse is true: large effects in accuracy and small effects in RTs. Accuracy and RTs for the two tasks cannot be directly compared because they are measured on different scales, accuracy on a probability-correct scale and RTs on a time scale. Neither RT nor accuracy can be used alone as the basis of a model of performance. A model built solely on accuracy data would almost surely be invalidated by RT data, and a model built solely on RTs would almost surely be invalidated by accuracy data. Instead, we compared associative recognition and priming in terms of accuracy and RTs jointly using a diffusion model for two-choice decisions that allows accuracy and RTs to be mapped onto the same metric.

The model also allowed us to address another issue: in item recognition, the correct response to both primed and unprimed test words is "yes," but in associative recognition, the correct response to intact pairs is different than the correct response to rearranged pairs. Just as with accuracy and RTs, the model allows the different responses to be measured on the same metric.

The model we used was the well-established diffusion model for two-choice decisions (Ratcliff, 1978; Ratcliff & McKoon, 2008; Wagenmakers, 2009), which has been successful in explaining item and associative recognition. It has not previously been applied to priming in item recognition, so an important question was whether it could explain priming data at the same time that it explained data for item and associative recognition.

The model separates out the components of processing that underlie two-choice decisions. One component, called "drift rate," is the quality of the evidence from memory that drives the decision process. The conditions of an experiment that are different in terms of difficulty each have different drift rates. The other components are the speed/accuracy criteria that are set on the decision process, and the time taken up by encoding and response output processes. The logic in using the model was that it would provide a coherent account of processes that underlie item recognition, priming, and associative recognition, explaining and fitting accuracy, RTs for correct responses, RTs for errors, and the shapes of RT distributions. The model brings together RT and accuracy measures onto a single scale, drift rate, and so the quality of the information that drives item recognition, priming, and associative information can be directly compared.

In mapping RT and accuracy onto the same scale, the model can find effects of independent variables that are significant in terms of drift rates even when they are not significant for RTs or accuracy. In this way, use of the model gives more powerful data analyses than RTs or accuracy. For example, in the data below, age affected priming for neither RTs or accuracy, but drift rates did show a significant effect. The model can also reconcile dissociations among measures: while there was no significant effect of age on accuracy or RTs for priming, there was a significant effect of age on both for associative recognition.

Among item recognition, associative recognition, priming, IQ, and age, there are several interactions that are of particular interest. In the next paragraphs, we provide background for the issues raised by these interactions.

Associative recognition and priming in item recognition

The question of whether priming in item recognition depends on the same information in memory as associative recognition has not, to our knowledge, been considered previously. To further our understanding of memory processes, both measures need to be considered.

At first thought, it might be supposed that associative recognition and priming are mediated by the same associative memory, that is, the association encoded when a pair of items is studied. However, in terms of current models, this is not necessarily the case. In global memory models (e.g., Dennis & Humphreys, 2001; Gillund & Shiffrin, 1984; Humphreys, Bain, & Pike, 1989; McClelland & Chappell, 1998; Murdock, 1982; Shiffrin & Steyvers, 1997), the two items of a pair are stored in one representation. At retrieval, a single test item is matched against single-item information in the representation and a pair test item is matched against pair information. Without further assumptions (e.g., Ratcliff & McKoon, 1988), the models might suggest that single-item tests are not influenced by associative information.

Dual-process models handle the distinction between item and associative information differently. They propose two processes, familiarity and recollection, that rely on different information in memory. Recognition of a single test item is based largely on familiarity information, whereas recognition of a pair test item is based primarily on recollection information. Some dual-process models conceive of familiarity and recollection as two distinct processes (e.g., Mandler, 1980; Reder et al., 2000; Rotello, Macmillan, & Reeder, 2004; Yonelinas, 1994, 1997). Others see a single source of information that either represents a continuum from familiarity to recollection (Wixted, 2007; Wixted & Stretch, 2004) or a single undifferentiated dimension (Cohen, Rotello, & Macmillan, 2008; DeCarlo, 2002; Dunn, 2004, 2008; Humphreys & Bain, 1983; Starns & Ratcliff, 2008). For all of the dual-process models, to the extent that item recognition depends primarily on familiarity information and pair recognition primarily on recollection information, it is difficult to see how priming effects in item recognition could be obtained.

Age, associative recognition, and priming in item recognition

For college students, we know that, for pairs of words that have been studied, responses to primed words are facilitated relative to unprimed words (e.g., Balota & Duchek, 1989; McKoon & Ratcliff, 1979, 1980; Neely & Durgunoglu, 1985; Ratcliff & McKoon, 1978, 1981). This is true for unrelated pairs (e.g., "chair justice") and for weakly related pairs (e.g., "plainly see"). Older adults show priming effects when the words of studied pairs are strongly semantically related (e.g., "dog cat"), and the effects are as large or larger than for young adults (e.g., Balota, Dolan, & Duchek, 2000; Laver, 2009; Laver & Burke, 1993). But for unrelated pairs, comparisons of older to young adults show equivalent priming only when the pairs are presented for study more than once (e.g., Faust, Balota, & Spieler, 2001; Howard, Heisey, & Shaw, 1986; Laver, 2009), not when there is no repetition (Spaniol, Madden, & Voss, 2006).

Age, item recognition, and associative recognition

Associative recognition, as measured in pairs tests, declines more with age than item recognition (e.g., Balota & Duchek, 1989; Bastin & van der Linden, 2006; Naveh-Benjamin, 2000; Naveh-Benjamin, Guez, Kilb, & Reedy, 2004; Naveh-Benjamin, Guez, & Shulman, 2004; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003; Ratcliff, Thapar, & McKoon, 2011). The problems presented by associative information for older adults are commonly attributed to a deficit in the ability to bind one piece of information to another (e.g., Chalfonte & Johnson, 1996; Light, 1991; MacKay & Burke, 1990; Mitchell, Johnson, Raye, Mather, & D'Esposito, 2000; Naveh-Benjamin and colleagues, e.g., Old & Naveh-Benjamin, 2008). The claim is that when two items are presented together and therefore are in short-term memory at the same time, older adults are less able than young adults to establish connections between them. This disadvantage could come about in several ways: Processing might be generally slower for older adults (e.g., Salthouse, 1996), older adults might be more limited in their attentional resources (e.g., Craik, 1983), older adults might have less capacity in short-term memory (e.g., Hasher & Zacks, 1988), or older adults might be less able to inhibit irrelevant, interfering information (Hasher & Zacks, 1979).

Age is a particularly insightful variable with which to compare priming and associative recognition because it dissociates item recognition, upon which priming depends, from associative recognition. As just discussed, the effects of age on performance are small for item recognition but large for associative recognition. If more information is retrieved about single items when they are tested in item recognition than when they are tested in pairs in associative recognition, and this extra information allows the retrieval of more associative relations than are available to pairs, then performance for priming might decline with age less than performance for associative recognition.

Age, item recognition, associative recognition, and IQ

IQ is another variable that dissociates item recognition from associative recognition, although there are fewer studies than for the effects of age. For item recognition, Ratcliff et al. (2011) found that drift rates for higher-IQ participants were larger than drift rates for lower-IQ participants for all three age groups in their experiment (college age, 60–74 year olds, and 75–90 year olds). For associative recognition, drift rates depended on an interaction between age and IQ: drift rates were low, nearing chance, for the higher-IQ and lower-IQ older participants and for the lower-IQ young participants; only the higher-IQ young participants showed reasonably large drift rates. The question of interest for the experiment described in this article was whether age interacts with IQ in the same way for priming as for associative recognition.

Related and unrelated pairs of words

In the experiment reported here, there were two types of pairs, one for which the words were related (e.g., "plainly see") and one for which the words were unrelated (e.g., "tree carpet"). One reason to include this variable was to provide more constraints on application of the diffusion model than would be available from the interactions of age, IQ, item recognition, associative recognition, and priming alone.

A second reason was to investigate whether related words support priming or associative recognition or both more than unrelated words (e.g., for associative recognition, see Naveh-Benjamin et al., 2003). Associations to be learned in real-world settings are often associations between related pieces of information. If this is correct, then experiments that have used pairs of unrelated words may have given a more pessimistic view of associative learning than is warranted.

Ratcliff's diffusion model

As explained in detail in the next section, the model explains performance in terms of the components of processing that give rise to the responses participants make and the RTs with which they make them. As pointed out above, an integrated explanation of the differences among performance in item recognition, associative recognition, and priming and how the differences are affected by age and IQ requires a model that accommodates both accuracy and RTs. One reason is simply that, for any task, a complete explanation must cover all of the task's dependent variables. A second reason is that accuracy and RTs often behave differently as a function of independent variables. For item recognition, for example, older adults are nearly as accurate as young adults, but they are slower; for associative recognition, they are less accurate as well as slower. The diffusion model allows a unified interpretation of accuracy and RT effects because, in the model, the components of processing responsible for them are the same.

Scaling accuracy and RTs onto the same metric is one problem solved by the diffusion model. A second, also a problem of scaling, is that there are baseline differences in RTs between younger and older adults. Older adults are generally slower than younger adults and show larger differences among conditions. For example, for associative recognition, RTs for older adults might be 1200 ms for intact pairs and 1400 ms for rearranged pairs, a difference of 200 ms, whereas for younger adults, the RTs might be 1000 ms and 1100 ms, a difference of only 100 ms. The smaller effect for younger adults might simply be that younger adults' discrimination is as good as older adults,' and their smaller effect is due only to their shorter RTs. With the model, this possibility can be examined.

A third key feature of the model is that not only can it be used to explore differences in components of processing among groups of participants, it can also be used to determine differences among individual participants. This allows the calculation of correlations between components of processing and, for example, IQ and age, and it allows the calculation of correlations between components of processing for different tasks.

Tests of the diffusion model

While the diffusion model has provided explanations for item and associative recognition data (see Ratcliff and McKoon, 2008, for a review), it has not yet been applied to priming in item recognition. The experiment described here provides a strong test of the model: The model should be able to explain, simultaneously, priming effects on RTs and accuracy, RTs and accuracy for item recognition, and RTs and accuracy for associative recognition, and it should be able to do so for older and younger participants, higher and lower IQ participants, and related and unrelated pairs of words.

A second challenge for the model was provided by what we call the ''small-n'' problem. In a priming study, it is often the case that the number of observations per participant must be small so that participants do not discover the priming manipulation and adopt strategies based on it. In the experiment described here, we had only 30 primed words and 30 unprimed words per participant (further divided into related and unrelated pairs). Because the number is small, RT and accuracy values are noisy. Further noise is added by differences among participants in the speed/accuracy criteria they adopt and differences among them in the time taken by processes outside the decision process (e.g., encoding the stimulus, response execution). Differences in speed/accuracy criteria contribute to noise in RTs and accuracy, and differences in nondecision processes contribute to noise in RTs. In the data analyses below, we describe how the model can be used to address this problem.

The diffusion model

The diffusion model is designed to explain the cognitive processes involved in making simple two-choice decisions like those required by the associative recognition and item recognition tasks used in the experiment described here (Ratcliff & McKoon, 2008). Decisions are made by a noisy process that accumulates evidence from a starting point z toward one of two response criteria, or boundaries, a and 0. In a recognition memory experiment, the evidence would come from a test item's representation in memory. When a boundary is reached, a response is initiated. For stimuli for which the correct response is ''yes'' for item recognition or ''intact'' for associative information, drift rates have a positive value, and for stimuli for which the correct response is ''no'' or ''rearranged'', drift rates have a negative value. Within-trial variability (noise) in the accumulation of information results in processes with the same mean drift rate terminating at different times (producing RT distributions) and sometimes at the wrong boundary (producing errors). The top panel of Fig. 1 illustrates the accumulation of evidence for stimuli with a positive drift rate (e.g., ''old'' test items). For each of the conditions in an experiment that vary in difficulty, there is a different value of drift rate. Components of processing outside the decision process, such as stimulus encoding and response execution, are combined in the model in a single parameter, the ''nondecision'' parameter, which is labeled by its mean duration, T_{er}. Once a decision process begins, the drift rate and the criteria are fixed; they cannot change as the accumulation of evidence proceeds.

RTs and accuracy are naturally integrated by the model: RTs are determined by the time it takes for accumulated evidence to reach one of the criteria (plus nondecision time), and which criterion is reached determines which response is given. The model abstracts the components of processing from RT and accuracy data and so separates drift rates, nondecision processes, and criteria settings from each other.

The values of drift rate, the criteria, and the nondecision parameter vary from trial to trial. This assumption is required if participants cannot accurately set the same values from trial to trial (e.g., Laming, 1968; Ratcliff, 1978). Across-trial variability in drift rate is assumed to be normally distributed with *SD* η, across-trial variability in the starting point is assumed to be uniformly distributed with range s_z, and across-trial variability in the nondecision component is assumed to be uniformly distributed with range s_t. In addition to across-trial variability, performance also includes ''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 in the model, on some proportion of trials (p_o), a uniformly distributed random delay, with range between the minimum and maximum RT for each condition, is added to the decision RT. The assumption of a uniform distribution is not critical; recovery of diffusion model parameters is robust to the form of the distribution (Ratcliff, 2008).

In fitting the model to the data from an experiment, the values of all of the components of processing identified by the model are estimated simultaneously from the data (the starting point (z), the distance between the criteria (a), the nondecision component (T_{er}), the variability parameters (η, s_z, and s_t), the contaminant probability (p_o), and a drift rate for each condition in the experiment. The model must fully explain the data: mean correct and error RTs for each experimental condition, accuracy for each condition, and the shapes of the RT distributions for correct and error

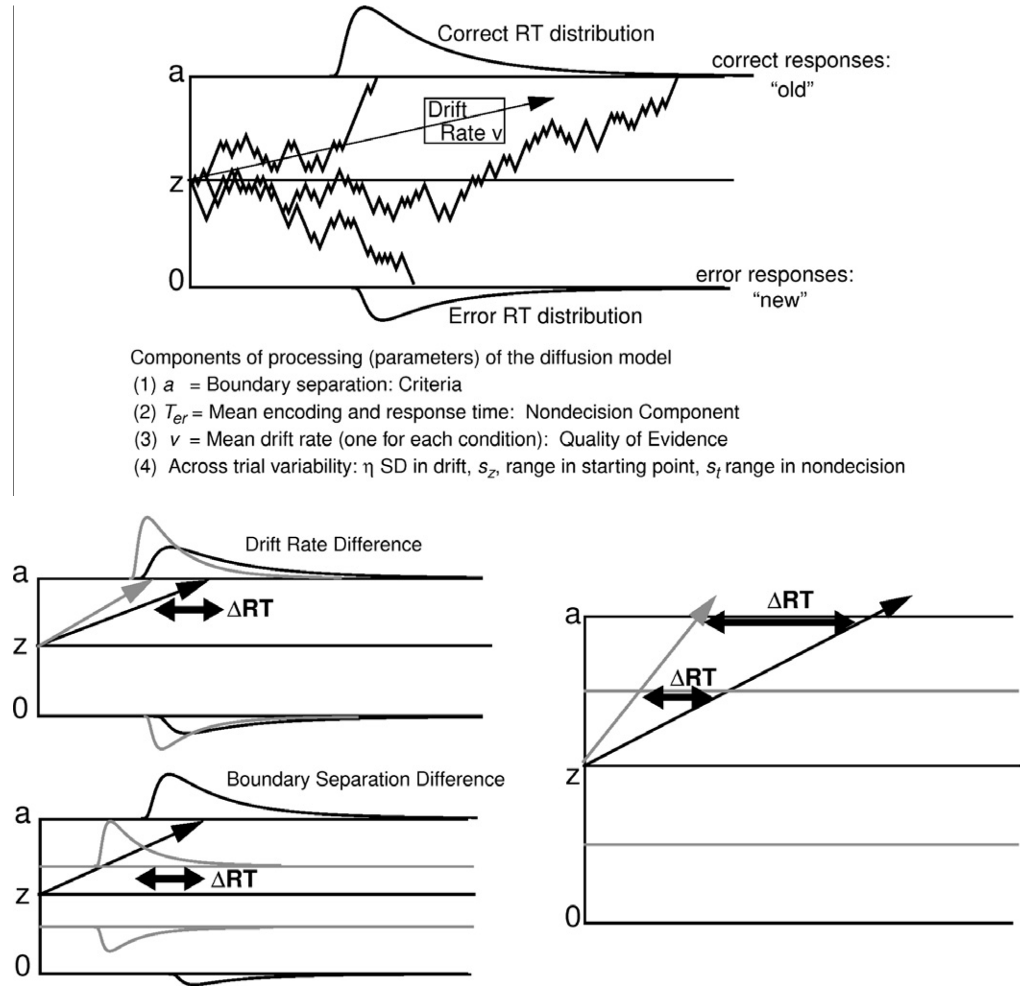

Fig. 1. An illustration of the diffusion model. The top panel shows three simulated paths. The two bottom left panels show the behavior of RT distributions when drift rate differs between two conditions and when boundary separation differs between two conditions. These two cases can give rise to equivalent RT differences even when accuracy and RT distributions differ (e.g., the double-ended thick arrows). The right-hand lower panel shows the effects on a difference in RTs when boundary separation increases: The mean RT difference increases (e.g., the double-ended thick arrows) and, for typical parameter values, accuracy increases only a few percent.

responses for each condition. In previous work, we have shown that, with a single 45-min session of data collection, the model can successfully fit data for individual participants with standard deviations in the parameter estimates for the distance between the criteria, nondecision time, and drift rate typically 3–5 times smaller than standard deviations across participants (which allows meaningful correlations to be computed).

When the model decomposes accuracy and RT data into components of processing, it is tightly constrained. The most powerful constraint comes from the requirement that the model fit the shapes of RT distributions (Ratcliff, 1978, 2002; Ratcliff & McKoon, 2008; Ratcliff, Van Zandt, & McKoon, 1999). A second constraint is that only drift rate can change across experimental conditions that vary in difficulty (and are randomly intermixed at test). There can be no changes in criteria settings or the nondecision parameter because making such changes would require knowing in advance what the correct response was for a stimulus. The change in drift rate alone must explain the changes from one condition to another in all the aspects of performance. The middle left-hand panel of Fig. 1 shows two conditions with different drift rates that produce different RT distributions (and hence different mean RTs).

Experimental conditions (or participants) that vary in their emphasis on speed versus accuracy are also explained with only one component of the model, the criteria settings (represented in the model by the distance between the two boundaries, a). For example, consider two conditions in an experiment with the same level of difficulty. For one, participants are instructed to respond as quickly as possible and for the other, as accurately as possible (Fig. 1 bottom left panel). Drift rate cannot be changed from one of the conditions to the other because the difficulty of the stimuli does not change (unless speed is highly stressed leading to a decrement in test item encoding). Differences in criterion settings must explain all the changes in the data. In the example, it is possible to get the same difference in mean RT from criterion changes as from drift rate changes but, as illustrated, the behavior of RT distributions is different for the two cases. Decreases in drift rate produce spreading of the RT distribution (and a reduction in accuracy) while increases in decision criteria produce a shift in the RT distribution in addition to spreading (and an increase in accuracy).

The bottom right panel of Fig. 1 illustrates a scaling effect. Drift rates are shown for two conditions (e.g., primed and unprimed words). The difference in RTs depends on the boundary settings, a larger difference with boundaries set further apart.

Age and Item recognition

The diffusion model has been used to explain a wide range of data for younger and older adults (Ratcliff, Thapar, & McKoon, 2001; Ratcliff, Thapar, & McKoon, 2003; Ratcliff, Thapar, & McKoon, 2004; Ratcliff, Thapar, Gomez, & McKoon, 2004; Ratcliff, Thapar, & McKoon, 2006a; Ratcliff, Thapar, & McKoon, 2006b; Ratcliff, Thapar, & McKoon, 2007; Ratcliff, Thapar, & McKoon, 2010; Thapar, Ratcliff, & McKoon, 2003, henceforth referenced as RTM). Item recognition data show large increases in RTs with age, coupled with smaller changes in accuracy or no change in accuracy at all (Ratcliff et al., 2004, 2006a, 2007, 2010). The large increases in RTs suggest large decrements in memory but the small differences in accuracy suggest only small decrements. As noted above, the diffusion model reconciles such seemingly inconsistent results by mapping the two dependent variables onto the same underlying decision processes. In the RTM studies, the large increases in RTs with age were due mainly to increases in criteria settings and the duration of the nondecision processes. Small or nonexistent deficits in accuracy were due to small or nonexistent decreases in drift rates. These findings led to the conclusion that, for memory for single items, drift rates (i.e., the strength of evidence available from memory) change little with age.

Age and associative recognition

Data behave differently for associative recognition than item recognition as a function of age (Ratcliff et al., 2011). With age, associative recognition performance suffers in both RTs and accuracy. The model explains this with drift rates decreasing with age, criteria settings moving farther apart, and the nondecision components slowing.

IQ, item recognition, and associative recognition

The effects of IQ on item and associative recognition were examined by Ratcliff et al. (2010). In both cases (for college age and 60–74 year olds), accuracy increased with IQ but IQ had only small effects on RTs. The model explained this with drift rates: drift rates increase with IQ, as would be expected, but the criteria and the nondecision parameter change relatively little with IQ.

Priming in item recognition

It is possible that older adults benefit more from priming or less from priming, in terms of drift rates, than young adults. But it is also possible that the only difference between older and young adults is that the older participants set their criteria further apart, leading to a scaling effect (bottom right panel of Fig. 1). This is illustrated with the simulation example shown in Table 1. RTs and accuracy were generated for two conditions, primed and unprimed. The drift rates were set to be the same for young and older adults: the drift rate for primed words was set at 0.3 and the drift rate for unprimed words was set at 0.2. The table shows three possible distances between the criteria: 0.08, 0.15, and 0.25. Despite the fact that drift rates were the same in the three cases, the priming effects were different: 9 ms, 38 ms, and 77 ms in RTs and .085, .095, and .103 in accuracy.

The issue of scaling groups with different baseline RTs and accuracy against each other has relevance to studies that have investigated semantic priming (e.g., priming between ''dog'' and ''cat''). For priming between semantically-related words in lexical decision, a number of researchers have shown that priming effects are larger for older than young adults (Laver & Burke, 1993; Myerson, Ferraro, Hale, & Lima, 1992; Myerson, Hale, Chen, & Lawrence, 1997). However, this conclusion does not take into account the possibility of scaling effects like those illustrated in Table 1 (see, e.g., Ratcliff et al., 2004).

Experiment

There were 12 blocks of test items. Each block began with a list of 16 pairs of words to be learned (3 s per pair). After the 16 pairs, participants were told whether the immediately following test items would be single words or pairs of words. For the pairs tests, participants were to respond ''intact'' or ''rearranged.'' For the single-word tests, they were to respond ''old'' or ''new.''

There were three groups of participants, college-age, 60–74 year olds, and 75–90 year olds. Each of these groups was divided into a lower-IQ group and a higher-IQ group. The experiment was designed to look at all the interactions among item recognition, associative recognition, priming, and age, and IQ. In addition, there were three types of materials: pairs of related words (e.g., ''plainly see''), unrelated pairs for which both words occur with high-frequency in English (Kucera & Francis, 1967), and unrelated pairs for which both words occur with low-frequency.

Method

Materials

There were 172 pairs of related words such as ''wine decanter,'' ''mortician cadaver,'' ''painter scaffold,'' ''prison knife,'' ''shower sky,'' ''monster insect,'' and ''passion argument.'' Forty-four of the related pairs were chosen from sentences used by Duffy, Henderson, and Morris (1989). The prime word of each pair was the subject of one of the sentences and the target word was the object of the

Table 1
Predictions from the diffusion model for primed and unprimed words as a function of boundary separation.

Boundary separation (a)	Primed mean RT (ms)	Unprimed mean RT (ms)	Priming effect (ms)	Primed accuracy	Unprimed accuracy
0.08	596	605	9	.795	.710
0.15	751	789	38	.858	.763
0.25	948	1025	77	.878	.775

Note. The parameter values for the predicted values were, starting point, $z = a/2$, nondecision component of response time, T_{er} = 500 ms, standard deviation in drift across trials, η = 0.25, range of the distribution of starting point, s_z = 0.05, proportion of contaminants, p_o = 0.002, range of the distribution of nondecision times, s_t = 200 ms, unprimed drift rate, v_1 = 0.2, and primed drift rate v_2 = 0.3.

sentence. The other 128 pairs were taken from McKoon and Ratcliff (1979, Experiment 3). The complete sets of words are presented in those two articles. The words in these pairs were chosen to make learning the pairs easier than pairs for which the words were completely unrelated. (Also, none of the words in the related pairs were listed as associates in the Nelson, McEvoy, & Schreiber (1998), free association norms.) The mean frequency (Kucera & Francis, 1967) of the first words of the pairs was 59.4 and the mean of the second words was 43.8.

There were also two pools of single words used to make up unrelated pairs, one pool of 378 high-frequency words (mean Kucera & Francis frequency, 265.8, with standard deviation 251.3) and one pool of 693 low-frequency words (mean 4.4 and standard deviation 0.5).

Procedure

The stimuli were displayed on a PC screen and responses were collected from its keyboard.

The to-be-learned lists of items, 16 pairs for each of 12 blocks, were constructed as follows: For six of the pairs, the two words were related; for four of the pairs, the words were unrelated and both were high frequency words; and for four of the pairs, the words were unrelated and both low-frequency words. The first and last pairs in the study list, randomly either two high-frequency words or two low-frequency words, were used as buffers. Pairs of words were presented one at a time, each pair on a single line of the PC screen with the words separated by two spaces, 3 s per pair, each pair followed by a clear screen for 100 ms and then the next pair to study.

Six of the to-be-learned lists of pairs were followed by a pairs test list and six were followed by a single-words test list (the pairs and single-words tests were presented in random order). At the beginning of a test list, participants were cued whether the test would be pairs (“Pairs- Does this pair exactly match what you studied?”) or single words (“Words- “Was this word in one of the pairs you studied?”). Participants were asked to press the ?/ key on the PC keyboard for positive responses and the zZ key for negative responses. They were instructed to respond as quickly and accurately as possible.

For a pairs test list, there were 16 test pairs. For the related pairs, for half the test lists there were two intact pairs and four rearranged pairs and for the other half there were four intact pairs and two rearranged (the two kinds of test lists were randomly ordered in the experiment). For the high-frequency pairs, two were intact and two were rearranged (both high-frequency words), and for the low-frequency pairs, two were intact and two were rearranged (both low-frequency words). The first two pairs of the test list were made up of words from the buffer study pairs, randomly either intact or rearranged. The first word of a test pair (always the left word of a studied pair) was displayed for 300 ms and then the second word appeared below it (this was done to reduce variability in RTs). If the response was longer than 280 ms, there was a 500 ms pause and then the next test pair. If the response was shorter than 280 ms, the message “TOO FAST” appeared on the screen for 1500 ms, then the 500 ms pause and then the next test item. “Too fast” responses sometimes occurred for the young participants.

A single-word test list consisted of 64 test items. Primed test words were always the second word of the pair in which they had been studied, immediately proceeded in the test list by the first word of the pair. Unprimed test words were also always the second word, preceded by the first word of a different pair.

For the related pairs, for half the test lists, there were four primed and two unprimed tests and for the other half, there were two primed and four unprimed tests (this makes 12 test items; the prime words for the unprimed test items came from other related pairs). For the high- and low-frequency (i.e., unrelated) pairs, for half the lists, there were two primed pairs and for the other half, two unprimed pairs (this makes four test items; the prime words came from other pairs with the same frequency). There were also eight high- and eight low-frequency words from studied pairs (this makes 16 test items) for which the immediately preceding test word was randomly either a word from another studied pair or a new, unstudied item. Finally, there were 16 high- and 16 low-frequency words (from the same pools of words as were used for the study lists) that were not studied in any list. We kept the proportion of primed pairs relatively low (0.08) to avoid any possibility that participants would adopt special strategies to try to recover the word with which each test word was studied (Ratcliff & McKoon, 1981).

Each word of a test list was displayed until a response was made. (The words were displayed in the same position on the PC screen as the left words of the pairs tests.) Then, just as with the pairs tests, if a response was correct and slower than 280 ms, there was a 500 ms pause and then the next test word. If the response was faster than 280 ms, the message “TOO FAST” appeared on the screen for 1500 ms, then the 500 ms pause, and then the next test item.

Participants

There were three groups of participants, college-age, 60–74 year olds, and 75–90 year olds. Each group was divided into a lower and a higher IQ group by splitting the participants available to us into two groups such that the groups had approximately the same number of participants. We refer to these two groups as the "lower" and the "higher" groups. For the college-age participants, there were 19 participants in the lower IQ group (IQ mean = 98.1, range = 80–114) and 20 participants in the higher IQ group (IQ mean = 122.6, range = 114–134. For the 60–74 year olds, there were 19 participants in the lower IQ group (IQ mean = 97.4, range = 80–114) and 20 participants in the higher IQ group (IQ mean = 122.3, range = 114–140). For the 75–90 year olds, there were 17 participants in the lower IQ group (IQ mean = 102, range = 83–114) and 17 participants in the higher IQ group (IQ mean = 124, range = 117–140).

The 39 college-age participants (18–25 years old) were recruited from The Ohio State University and Columbus, OH, area community centers. Twenty-nine responded to posted fliers and were paid $12 per session for their participation; 10 received credit for an introductory psychology course. The 39 60–74 year olds and the 34 75–90 year olds were recruited from local senior centers and were paid $15 per session (they were paid more than the young participants because they had to travel to the center). Each participant participated in one 50-min session for the experiment and one 1-h session to collect the background information described next.

All participants had to meet the following criteria to participate in the study: a score of 25 or above on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975), and no evidence of disturbances in consciousness, medical or neurological diseases causing cognitive impairment, head injury with loss of consciousness, or current psychiatric disorder. All participants completed the Vocabulary and Matrix Reasoning subtests of the Wechsler Adult Intelligence Scale-III (Wechsler, 1997), and estimates of full-scale IQ were derived from the scores of the two subtests (Kaufman, 1990). Participants also completed the Center for Epidemiological Studies-Depression Subscale (CES-D, Radloff, 1977), on which there were no significant differences among the three age groups. The means and standard deviations for these background characteristics are presented in Table 2.

Results

Responses longer than 4000 ms and shorter than 350 ms were excluded from analyses. This corresponded to 0.7%, 1.9%, and 3.8% of the data for item recognition for the three groups of participants, college age, 60–74 year olds, and 75–90 year olds, respectively; and 1.8%, 2.6%, and 4.8% of the data for associative recognition for the three groups. There were larger amounts of data excluded for older participants because we used the same fixed cutoffs for all groups, but these amounts are in the range of those typically excluded in experiments like the one reported here.

The data are reported in Tables 3–5. We conducted three analyses of variance, one for the item recognition data, one for the associative recognition data, and one for priming effects. The full results of the analyses of variance are given in Table 6. Some of the main effects are very large compared to the interactions in which they participate, which means that the main effects can be interpreted as though there were no qualifying interactions. To aid this, the large significant effects ($p < .001$) are denoted by # marks and the smaller significant effects ($p < .05$) by $*$.

In the paragraphs below, we begin with the main effects of age, IQ, and pair type on accuracy and median RTs and then discuss interactions among them.

For both item recognition and associative recognition, accuracy decreased with age ($F(1106) = 7.8$; $F(2106) = 5.0$) and it was lower for lower than higher IQ participants ($F(1106) = 49.5$; $F(1106) = 31.2$). For associative recognition, accuracy was better for related than unrelated pairs ($F(1106) = 239.9$). In contrast to the strong effects of age, IQ, and pair type on accuracy for item and associative recognition, these variables had no significant effects on accuracy for priming.

RTs increased with age for item and associative recognition ($F(2106) = 61.4$; $F(2106) = 50.2$), and decreased with IQ ($F(1106) = 11.8$; $F(1106) = 4.3$). For priming, neither age nor IQ significantly affected RTs. Although the main effects were not significant, age, IQ, and pair type did interact, as we discuss below. For both priming and associative recognition, RTs were shorter for related pairs than unrelated pairs, ($F(1106) = 22.3$; $F(1106) = 71.4$).

Table 2
Participant characteristics.

Measure	College age 18–25		60–74 year olds		75–90 year olds	
	M	*SD*	*M*	*SD*	*M*	*SD*
Mean age	20.6	2.13	68.4	3.93	80.8	4.11
Years education	13.4	1.95	17.5	3.39	13.65	2.53
MMSE	29.0	1.05	28.9	1.14	27.7	1.49
WAIS-III vocabulary (raw/unscaled score)	42.6	10.82	48.1	10.22	43.9	10.38
WAIS-III vocabulary (scaled score)	11.6	2.82	11.9	2.58	11.6	2.51
WAIS-III matrix reason ing (raw/unscaled score)	19.5	4.42	14.3	6.12	12.8	5.09
WAIS-III matrix reason ing (scaled score)	12.1	3.02	11.7	3.35	13.0	3.25
WAIS-III IQ	110.7	14.95	110.2	14.98	112.9	14.00
CES-D	10.1	6.33	9.5	7.09	8.3	5.40

Note. MMSE = Mini-Mental State Examination; WAIS-III = Wechsler Adult Intelligence Scale-3rd edition; CES-D = Center for Epidemiological Studies-Depression Scale.

Table 3
Mean accuracy and mean of median RTs for priming in item recognition as a function of test type, age, and IQ.

Age group	Mean of IQs	Accuracy					
		Primed related	Unprimed related	Primed unrelated	Unprimed unrelated	Priming difference related	Priming difference unrelated
College age	98.1	.775	.724	.747	.666	.051	.081
	121.6	.842	.781	.765	.724	.061	.041
60–74 year	97.1	.780	.753	.650	.656	.027	−.006
olds	121.4	.872	.850	.780	.762	.022	.018
75–90 year	102.0	.713	.675	.581	.588	.038	−.007
olds	123.2	.827	.791	.761	.712	.036	.059
		Means of median RTs (ms)					
College age	98.1	633	699	659	727	66	68
	121.6	558	676	655	707	118	52
60–74 year	97.1	872	925	976	972	53	−4
olds	121.4	756	874	869	890	118	21
75–90 year	102.0	998	1164	1265	1199	166	−66
olds	123.2	815	931	891	931	116	40

Item recognition has usually been compared to associative recognition in terms of accuracy, usually with unrelated pairs of words for associative recognition. The data reported here show the typical result, that accuracy decreases more with age for associative recognition than item recognition. For item recognition, relative to chance (50%), the drop from college-age to 75–90 year olds was 19%. For associative recognition, the drop was 68%. For related pairs, the drop in accuracy for associative recognition was much less, from college-age to oldest only 5%, indicating that relatedness supported associative memory for the older participants. The interaction between pair type and age was significant, $F(2106) = 12.2$. (This interaction was not significant for RTs.)

For accuracy for associative recognition, there was an interaction between age, IQ, and pair type. For unrelated pairs, higher-IQ participants' accuracy was much better than lower-IQ participants' for college-age and 60–74 year olds, but it fell to near chance for 75–90 year olds. For the college-age and 60–74 year olds, accuracy for the higher-IQ participants averaged 22% above chance but for the 75–90 year old participants, only 8% above chance. For the lower-IQ participants, accuracy was poor for all three age groups. It averaged 10% above chance for the college-age and 60–74 year olds and 4% above chance for the 75–90 year olds.

For the related pairs, the pattern was different. Accuracy for the lower-IQ participants was above chance for all three age groups (20%, 24%, and 20%, respectively), as was accuracy for the higher-IQ participants (34%, 36%, and 32%, respectively). In terms of proportion correct, the triple interaction between age, IQ, and pair type was not significant. However, in terms of percent values relative to chance, the pattern of effects is clear.

For RTs for associative recognition, there was an interaction between age, IQ, and pair type ($F(2106) = 3.2$) for which we have no principled explanation. For related pairs, higher-IQ participants were faster than lower-IQ participants for all age groups. For unrelated pairs, for the older participants, the higher-IQ participants were slower than the lower-IQ participants.

For RTs for priming, the triple interaction mentioned above between age, IQ, and pair type ($F(2106) = 3.6$) shows a pattern similar to that for accuracy in associative recognition. For unrelated pairs, there were (small) priming effects for the college-age participants and for the higher-IQ older participants. For the lower-IQ 60–74 year olds and 75–90 year olds, there was no significant priming effect. For related pairs, priming effects were reasonably large for all participants and there was no significant difference across age groups (consistent with earlier results, e.g., Balota et al., 2000).

In item recognition, the data for high- and low-frequency words show the effects that would be expected. Responses for high-frequency words were less accurate and slower than responses for low-frequency words, $F(1106) = 259.9$ and $F(1106) = 35.1$ (showing a mirror effect). For both accuracy and RTs, the difference between high- and low-frequency words was larger for the older participants (the interaction of age and frequency was significant for accuracy, $F(2106) = 3.1$, and for RTs, $F(2106) = 5.2$).

Supplementary study

In the experiment we have described, the conditions were such that if the two words in a test pair were related, then the correct response was always "intact." In the supplementary study, this was not the case: for some test pairs for which the words were related, the correct response was "rearranged." The results of this study, described in the Appendix, showed that, for related test pairs, participants were able to discriminate whether they were intact or rearranged.

Summary

There are five key points about the data:

1. Performance for unrelated pairs of words, as measured by accuracy, decreased with age much more for associative recognition than item recognition. For the oldest participants, accuracy for associative recognition was near chance for both higher- and lower-IQ.
2. Priming for unrelated pairs of words, as measured by RTs, decreased with age such that, for the older lower-IQ participants, there was no priming effect at all (though there is considerable variability in the 75–90 year old results).

Table 4
Mean accuracy and mean of median RTs for high and low frequency fillers in item recognition as a function of age and IQ.

Age group	Mean IQs	Accuracy					
		Old high freq.	Old low freq.	New high freq.	New low freq.	High freq. mean	Low freq. mean
College age	98.1	.657	.713	.712	.790	.685	.752
	121.6	.695	.805	.813	.895	.754	.850
60–74 year olds	97.1	.541	.711	.750	.843	.646	.777
	121.4	.698	.823	.830	.918	.766	.874
75–90 year olds	102.0	.534	.642	.704	.810	.619	.726
	123.2	.652	.726	.733	.875	.693	.801
		Means of median RTs (ms)					
College age	98.1	706	714	752	806	729	760
	121.6	694	670	814	692	754	681
60–74 year olds	97.1	1022	952	1085	1016	1055	984
	121.4	942	860	1031	944	987	902
75–90 year olds	102.0	1205	1182	1359	1286	1282	1229
	123.2	977	942	1168	1042	1073	992

Table 5
Mean accuracy and mean of median RTs for associative recognition as a function of test type and IQ.

Age group	Mean IQs	Accuracy					
		Related intact	Unrelated intact	Related rearranged	Unrelated rearranged	Related mean	Unrelated mean
College age	98.1	.677	.587	.734	.660	.701	.624
	121.6	.833	.748	.836	.745	.835	.747
60–74 year olds	97.1	.777	.565	.708	.586	.743	.576
	121.4	.901	.710	.809	.663	.855	.687
75–90 year olds	102.0	.748	.571	.643	.502	.696	.537
	123.2	.840	.623	.790	.543	.815	.583
		Means of median RTs (ms)					
College age	98.1	727	913	890	938	809	926
	121.6	673	806	789	931	731	869
60–74 year olds	97.1	947	1156	1499	1420	1223	1288
	121.4	921	1300	1280	1546	1101	1423
75–90 year olds	102.0	1246	1662	1892	1745	1569	1704
	123.2	997	1317	1447	1585	1222	1451

3. For related pairs, priming effects on RTs were reasonably large and accuracy for associative recognition was significantly above chance for all groups of participants.
4. Age did not affect accuracy in associative recognition in the same way that age affects RTs in priming. There were significant main effects of age on associative recognition (accuracy and RTs) but not on priming. This is a problem that application of the diffusion model will address to a large degree: Older adults adopted more conservative decision criteria. This led to longer RTs and magnified RT differences in priming. When drift rates were examined, as discussed below, the effects of age on priming and associative recognition were parallel.
5. There were large differences in baseline RTs: For both item recognition and associative recognition, the older participants were slower than the college-age participants, and the higher-IQ participants were faster than the lower-IQ participants.

Diffusion model analyses

In this section, we show that the diffusion model fits the experimental data well. In the next sections, we describe the effects of age, IQ, and pair type on the components of processing identified by the model.

The experiment provided a wide range of constraints on the model. Studied pairs were made up of related words, unrelated words, and high- and low-frequency words. Memory was tested with single words, priming between single words, and intact/rearranged pairs. Participants varied in age and IQ. The data encompass accuracy, RTs for correct responses, RTs for error responses, the relative speeds of correct and error responses, and the shapes of the RT distributions.

To fit the model to the data, the RT distributions were represented by 5 quantiles, the .1, .3, .5 (the median), .7, and .9 quantiles. The quantiles and the response proportions were entered into the minimization routine and the diffusion model was used to generate the predicted cumulative probability of a response occurring by that quantile RT. Subtracting the cumulative probabilities for each successive quantile from the next higher quantile gives the proportion of responses between adjacent quantiles. For a chi-square computation, these are the expected proportions, to be compared to 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). The proportions are multiplied by the number

Table 6
ANOVA results for accuracy, median RT, and drift rate.

	Factor	*df*	Priming		Associative recognition		Item recognition	
			F	MSE	*F*	MSE	*F*	MSE
Accuracy	Age	2	.67	.0187	4.98∗	.1042	7.79#	.0690
	IQ	1	.15	.0044	31.21#	.6531	49.48#	.4379
	Type	1	.13	.0037	239.86#	1.2130	259.93#	.5877
	AgexIQ	2	.10	.0029	.42	.6555	0.55	.0049
	AgexType	2	.02	.0006	12.16#	.0615	3.08∗	.0069
	IQxType	1	.02	.0007	1.67	.0084	.01	.00002
	AgexIQxType	2	.39	.0112	1.39	.0070	1.58	.0035
RT	Age	2	.97	15,764	50.21#	81,47,997	61.44#	33,83,917
	IQ	1	3.17	51,286	4.33∗	703,154	11.89#	654,995
	Type	1	22.25#	407,026	71.40#	15,82,224	35.13#	146,371
	AgexIQ	2	.23	3824	2.81	457,006	3.17∗	174,869
	AgexType	2	3.64∗	66,749	1.15	25,700	5.18∗	21,601
	IQxType	1	0.07	1329	9.91∗	219,698	2.28	9506
	AgexIQxType	2	3.64∗	66,602	3.21∗	71,219	0.09	405
Drift Rate	Age	2	3.19∗	.086	4.28∗	.044	9.20#	.367
	IQ	1	1.13	.030	14.16#	.148	40.49#	1.625
	Type	1	8.46∗	.183	41.58#	.420	207.52#	1.813
	AgexIQ	2	.94	.025	.78	.008	0.22	.008
	AgexType	2	1.05	.022	16.71#	.037	4.25∗	.014
	IQxType	1	.76	.016	3.19	.085	9.77∗	.064
	AgexIQxType	2	.28	.006	3.89∗	.003	0.42	.014

Note. For priming and associative recognition, type = relatedness, for item recognition, type = low versus high frequency words. The denominator degrees of freedom (df) is 106. # designates effects significant at the .001 level and ∗ designates effects significant at the .05 level.

of observations in the condition to give observed (O) and expected (E) frequencies and summing over $(O-E)^2/E$ for all conditions gives a single chi-square value to be minimized. The number of degrees of freedom in the data are the 12 proportions between and outside the quantiles minus 1 (because the sum must equal 1) multiplied by the number of conditions in the data. The model was fit to the data for each participant individually, using a chi-square minimization method as described by Ratcliff and Tuerlinckx (2002).

In fitting the model to the data, there were seven parameters that were fixed across experimental conditions for item recognition and seven that were fixed for associative recognition, for each participant in each of the six groups. The seven were: the distance between the criteria, the time taken up by nondecision processes, the starting point of the accumulation of evidence process, the range of the starting point, the range of the nondecision component, the across-trial *SD* in drift rates, and the probability of contaminant responses. For item recognition, in addition to the seven parameters just listed, there were nine drift rate parameters, one for each of the nine categories of data: primed words and their unprimed counterparts from related study pairs (correct responses "old"), primed words and their unprimed counterparts from unrelated study pairs (correct responses "old"), the words that were used as primes (correct response "old"), unprimed high-frequency words from studied pairs (correct response "old"), unprimed low-frequency words from studied pairs (correct response "old"), high-frequency words that were not studied (correct response "new"), and low-frequency words that were not studied (correct response "new"). For associative recognition, in addition to the seven parameters listed above, there were four parameters, one each for intact pairs for which the words were related, intact pairs for which the words were unrelated, rearranged pairs for which the words in their study pairs were related, and rearranged pairs for which the words in their study pairs were not related. In sum, there were 99–16 degrees of freedom for item recognition and 44–11 degrees of freedom for associative recognition.

Table 7 shows the best-fitting values, averaged across participants, of the seven parameters that were fixed across conditions and the standard deviations (*SD*s) in them. The values of the seven parameters did not vary significantly with IQ, so the table shows the means and *SD*s for the participants grouped only by age. The model fit the data well, as shown by the average chi-square values in Table 7. The chi-square values are close to the critical values for both tasks. Because the chi-square test is a very conservative test, these values show good fits of the model to the data Although there might be significant misses, the power of the chi-square statistic grows with the number of observations so that even a small difference between the proportions of observed and expected frequencies can become significant (see Ratcliff et al. (2004), for discussion).

For item recognition, Fig. 2 shows examples of how well the model accounts for accuracy and RTs for correct responses. For this figure, we chose two conditions: one for which the number of observations was large, specifically, non-studied low-frequency words (the plots for non-studied high-frequency words and both high- and low-frequency studied words look the same); and one for which the number of observations was small, specifically, words that were primed by their related pair-mates (the other primed and the unprimed conditions look the same). For non-studied low-frequency words, there were 96 observations per participant but for the primed words, there were only 18 (consequently, the accuracy values tend to line up vertically). The top panel of the figure plots accuracy as predicted from the best-fitting parameters of the model against accuracy from the data, for each participant in the experiment. The model's predictions are close to the data, as shown by their proximity to the straight line that represents a perfect match between the two. The lower panels show the 0.1, 0.5, and 0.9 quantiles of the RT distributions. Again, the values predicted from the best-fitting parameters are close to the line that represents a perfect match. Note that the .9 quantile RTs have much more variability than lower quantiles and so the quality of the fit is worse.

Fig. 2 also shows how well the model fares with the associative recognition data. Two of the four conditions were chosen for this figure: pairs for which the words were related at study, tested either intact or rearranged. As with the item recognition data, the model's predictions are close to the straight lines for both accuracy and quantile RTs. Because the fits are based on only 18 observations per condition, there is a relatively large amount of variability in the fits to the data, especially for the .9 quantile RTs.

In reporting the best-fitting parameter values for the model in the next sections, we report the mean of each one (averaged over participants) and the *SD*s in the parameter values across participants. The *SD*s in parameter values across individuals must be larger than the *SD* in the estimate of a parameter value for each individual in order to examine individual differences. Variability in parameter values for single participants comes about because of the inherent variability in the components of processing in the model, which arise from variability in data. Even if we generated simulated data from the model with exactly the same parameter values for each of several simulated experiments, the estimated parameter values would vary across experiments. For the experiment reported in this article, for item recognition, the *SD*s in parameter values for single participants were 3–4 times less than the *SD*s of the average across participants, and for associative recognition, they were 2–4 times less (values for the *SD*s in model parameters can be computed from tables in Ratcliff & Tuerlinckx (2002)).

Diffusion model parameters

Table 8 shows the best-fitting values of drift rates and their *SD*s across participants and Table 6 shows the results of analyses of variance. Figs. 3 and 4 summarize the patterns in the drift rates that are discussed in the next sections. For item recognition, the figures show drift rates that are the sums of the absolute values of the old and new drift rates divided by two, collapsed over high- and low-frequency words. For associative recognition, the drift rates are the sums of the absolute values of the intact and rearranged drift rates divided by two. For priming, the figures show the differences in drift rates between primed and unprimed words.

Associative recognition

As has been noted, the numbers of observations per condition for associative recognition were small. However, the probability of an "intact" response was above .4 for intact pairs and below .5 for rearranged pairs so the difference between the two probabilities was large enough that reasonably good estimates of model parameters could be obtained. Overall, the effects of age, IQ, and pair type on drift rates were significant, $F(2106) = 4.3$, $F(1106) = 14.2$, and $F(1106) = 41.6$, respectively (see Table 6).

As Fig. 3 shows, the drop in drift rates from the young to the older adults was much steeper for the unrelated pairs than the related pairs. The pattern is especially dramatic when the effects are viewed in terms of proportional decline to floor. For the unrelated pairs, the difference dropped to 58% of its value from college-age to 60–74 year olds and to 21% of its value from college-age to 75–90 year olds. In contrast, the equivalent numbers for related pairs were 88% and 56%. The interaction of age and pair type was significant, $F(2106) = 16.7$.

Fig. 4 shows drift rates for lower- and higher-IQ participants. For the college-age and 60–74 year olds, drift rates increased with IQ for both related and unrelated pairs. For the 75–90 year olds, the drift rates for both lower- and higher-IQ participants were near zero for unrelated pairs and near zero for the lower-IQ participants for related

Table 7
Diffusion model parameters.

Task	Subjects	a	T_{er} (ms)	η	s_z	p_o	s_t (ms)	z	χ^2
Item mean	College	0.159	0.498	0.245	0.054	0.002	0.231	0.063	112.4
	60–74	0.218	0.657	0.265	0.036	0.002	0.221	0.083	97.1
	75–90	0.231	0.710	0.216	0.054	0.005	0.249	0.089	102.5
Associative mean	College	0.180	0.507	0.167	0.077	0.002	0.229	0.078	38.0
	60–74	0.245	0.636	0.130	0.051	0.003	0.260	0.095	41.7
	75–90	0.248	0.715	0.106	0.041	0.002	0.325	0.094	43.7
Item *SD*	College	0.052	0.063	0.069	0.044	0.004	0.087	0.023	51.4
	60–74	0.075	0.085	0.050	0.039	0.008	0.081	0.030	25.0
	75–90	0.067	0.124	0.081	0.062	0.009	0.108	0.026	37.4
Associative *SD*	College	0.052	0.094	0.104	0.062	0.004	0.128	0.025	20.8
	60–74	0.073	0.122	0.089	0.071	0.006	0.142	0.031	16.4
	75–90	0.063	0.136	0.081	0.046	0.003	0.106	0.025	19.1

Note. a = boundary separation, z = starting point, T_{er} = nondecision component of response time, η = 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, and χ^2 is the chi-square goodness of fit measure (for item recognition, with 83° of freedom, the critical value is 110.1 and for associative recognition, with 33° of freedom, the critical value is 50.7).

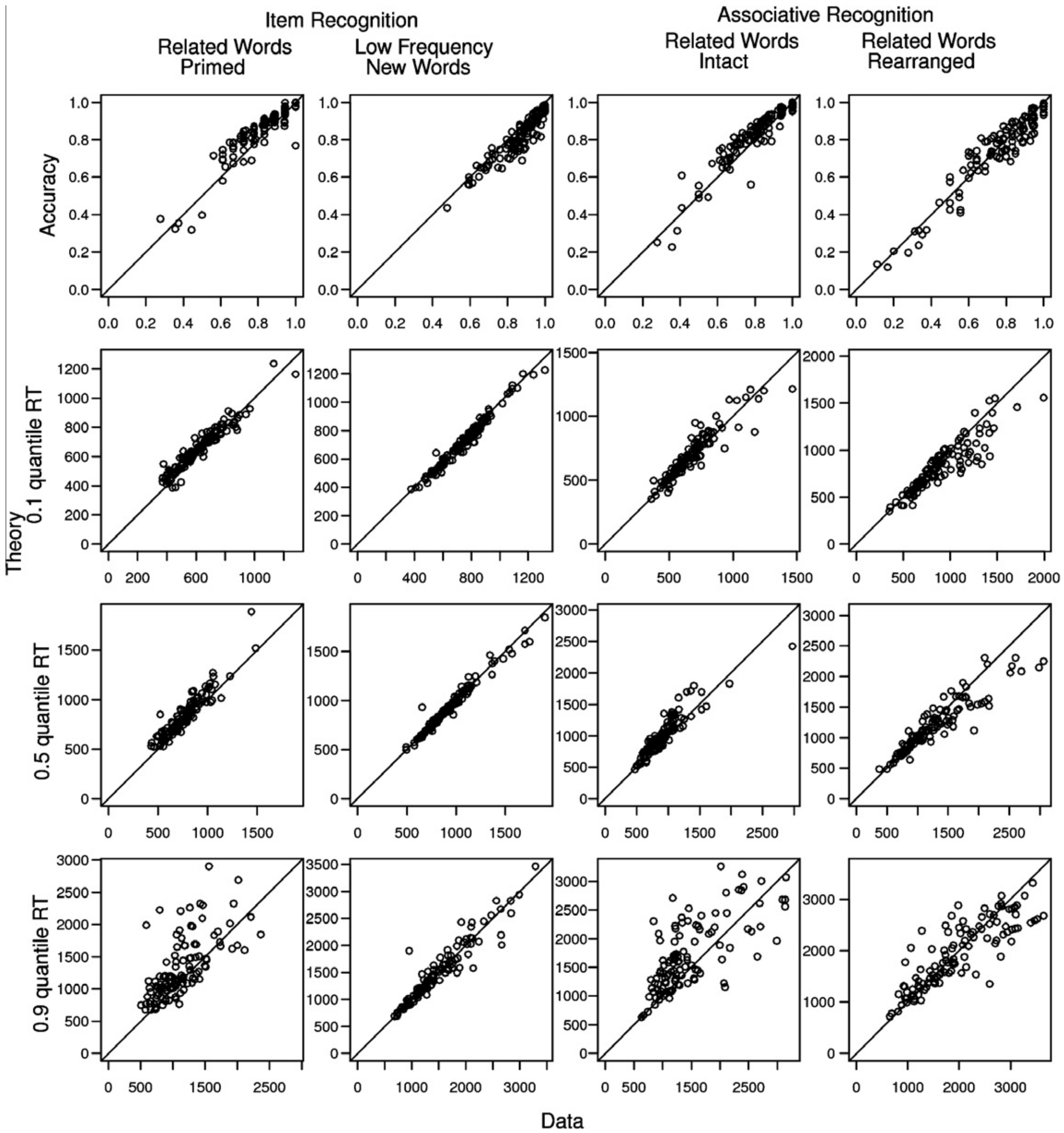

Fig. 2. Plots of accuracy and the .1, .3, .5 (median), .7, and .9 RT quantiles for data (*y*-axis) and predicted values from fits of the diffusion model (*x*-axis) for correct and error responses for two conditions for item recognition and two conditions for associative recognition for all participants in the three age groups (the data and fit for each participant is represented by an open circle). For item recognition, the data are for primed words (small numbers of observations) and low-frequency new words (larger numbers of observations), and for associative recognition, the data are for related intact pairs and for rearranged words from related study pairs.

pairs. The triple interaction between age, IQ and relatedness was significant, $F(1106) = 3.9$.

Priming

Just as for associative recognition, the numbers of observations were small. However, unlike associative recognition, the differences in accuracy between the primed and unprimed conditions were too small to yield significant effects (although the trends were in the expected directions, primed higher than unprimed).

The problem with the priming conditions is the small-n problem mentioned in the introduction. The number of observations per participant was small so that participants did not come to suspect the priming manipulation. With such a small number, the primed and unprimed accuracy and RT data are highly variable and represent a mixture of participants with different criteria settings and different nondecision times. This means that parameters of the model cannot be estimated from the data from the primed and unprimed conditions alone.

The solution to the problem comes from the other conditions of the experiment. When we fit accuracy and RTs for all the conditions simultaneously, the conditions with large n's (high- and low-frequency and old and new test words) largely determine the nondecision and criterion parameters and the across-trial variabilities. Since these

Table 8
Drift rate parameters.

Task	Participants	v_{pr}	v_{ur}	v_{pn}	v_{un}	v_p	v_{ho}	v_{lo}	v_{hn}	v_{ln}
Item mean	College	0.300	0.173	0.227	0.135	0.223	0.083	0.151	−0.225	−0.314
	60–74	0.335	0.224	0.156	0.140	0.228	0.044	0.168	−0.239	−0.344
	75–90	0.217	0.155	0.115	0.092	0.172	0.037	0.087	−0.156	−0.255
Item *SD*	College	0.190	0.115	0.178	0.133	0.110	0.105	0.120	0.131	0.162
	60–74	0.193	0.101	0.130	0.156	0.115	0.118	0.108	0.128	0.154
	75–90	0.156	0.123	0.121	0.148	0.109	0.097	0.090	0.112	0.146
		v_{ir}	v_{rr}	v_{iu}	v_{ru}					
Associative mean	College	0.211	−0.214	0.081	−0.121					
	60–74	0.224	−0.148	0.044	−0.073					
	75–90	0.125	−0.115	0.005	−0.037					
Associative *SD*	College	0.226	0.156	0.133	0.109					
	60–74	0.205	0.125	0.089	0.086					
	75–90	0.080	0.124	0.076	0.087					

Subscripts. *pr* = primed related, *ur* = unprimed using items from related pairs, *pu* = primed from non-related pairs, *uu* = unprimed from non-related pairs, *p* = primes, *ho* = high frequency old, *lo* = low frequency old, *hn* = high frequency new, *ln* = low frequency new, *ir* = intact from related pairs, *rr* = rearranged from related pairs, *iu* = intact non-related pairs, *ru* = rearranged from non-related pairs.

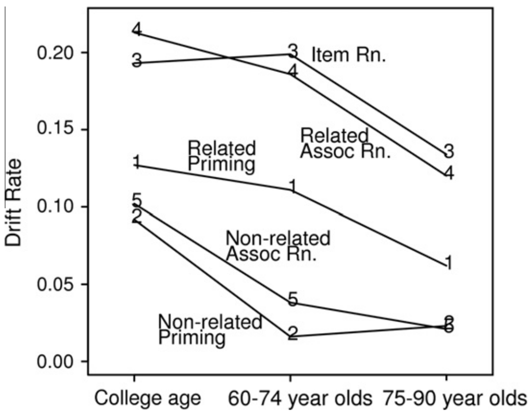

Fig. 3. Drift rates as a function of age and item type. The related and unrelated priming drift rates are the differences in drift rates between the primed and unprimed conditions, the associative recognition drift rates are half the differences between intact and rearranged pairs, and the item recognition drift rates are half the differences between the positive and negative fillers.

parameters cannot vary across experimental conditions, drift rates for the primed words and their unprimed counterparts are determined by their RT and accuracy data alone. The estimates of drift rates obtained in this way have less variability than would be obtained if the other parameters were not fixed across conditions, and less variability than in the raw accuracy and RT values.

White, Ratcliff, Vasey, and McKoon (2009, 2010a, 2010b) demonstrated this with studies of anxiety. The question was whether individuals with high trait anxiety could be differentiated from those with low trait anxiety in terms of their responses to what are called threat words (e.g., anger, fear). Testing such words in lexical decision, previous studies had failed to find consistent effects on either RTs or accuracy. However, when the diffusion model was applied, there were significant effects in drift rates: anxious individuals had larger drift rates for threat words compared to neutral words, whereas non-anxious individuals did not.

Application of the small-n method to the data reported here had the same consequence as for the White et al. studies: power was increased relative to the raw data. In particular, the effect of age that was not significant in accuracy or RTs became significant in drift rates.

Fig. 3 shows the effects of age and pair type on the differences in drift rates between primed and unprimed words. The main effects of both were significant, $F(2106) = 3.2$ and $F(1106) = 8.5$.

In the model, drift rates are determined more by accuracy than RTs. The difference in accuracy between the primed and unprimed conditions was small and not significant. This suggests that the drift rates for priming will be small relative to the drift rates for associative recognition, and this is what occurred (Figs. 3 and 4).

Despite the drift rates being smaller for priming than associative recognition, the pattern of age and pair type for priming looks much like the pattern for associative recognition. Like associative recognition, age and pair type interacted. For the unrelated pairs, the drift rate difference dropped to near floor for 60–74 year olds and 75–90 year olds (to 17% and 25%, respectively, of the difference for college-age participants). For the related pairs, the drift rate difference for the older participants was much larger. The drift rate difference for the 60–74 year olds was 87% of that for college-age participants, and for the 75–90 year olds, it was 48%, values that are well above floor. However, although the pattern of priming effects matched that of associative recognition, the interaction between age and pair type did not reach significance. The variability in the drift rates was too large even though there were large numbers of items in the other conditions of the experiment (high- and low-frequency words, old and new words).

Like the effects of age, the effects of IQ on priming were similar to the effects of IQ on associative recognition (Fig. 4). For related pairs, the drift rate difference between primed and unprimed test words increased with IQ for all three age groups. For unrelated pairs, the difference declined with IQ for the college-age participants, a finding for which we have no explanation except noise in the data. The difference increased with IQ for the 60–74 year olds.

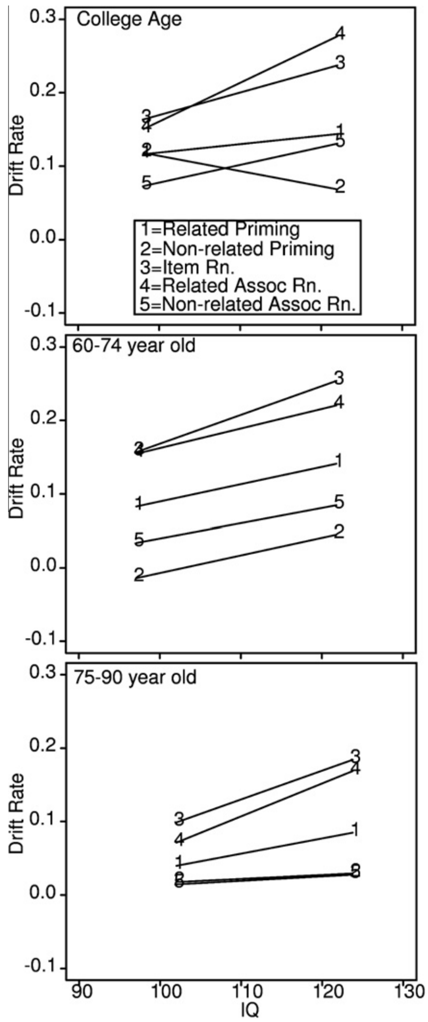

Fig. 4. Plots of drift rates for item type as a function of IQ and age. The related and unrelated priming drift rates are the differences in drift rates between the primed and unprimed conditions, the associative recognition drift rates are half the difference between intact and rearranged pairs, and the item recognition drift rates are half the differences between the positive and negative fillers.

For the 75–90 year olds, the drift rate difference was near zero for both the lower- and higher-IQ participants. This interaction, between IQ, age, and pair type, was not significant, a disappointing result that is partly due to the turnaround in IQ drift rates for the college-age participants. Also, compared to the three age groups, there were half as many participants in each IQ group.

Item recognition

For item recognition, there were significant decreases in drift rate with age, IQ, and word frequency ($F(2106) = 9.2$ for age, $F(1106) = 40.5$ for IQ, and $F(1106) = 207.5$ for frequency). Figs. 3 and 4 show that the effects were above floor and almost parallel for IQ and age. There were significant interactions between age and frequency ($F(2106) = 4.3$) and between IQ and frequency ($F(1106) = 9.8$).

Other diffusion model parameters

Results of analyses of variance are shown in Table 9. For associative recognition and item recognition, there were significant effects of age and IQ on RTs. For associative recognition, the effect of age was due to increases in boundary separation and nondecision time ($F(2106) = 14.1$ and $F(2106) = 30.7$), as well decreases in drift rates. The effect of IQ was due to decreases in nondecision time ($F(1106) = 4.4$), as well as increases in drift rates. The results for item recognition were the same: the effect of age was due to increases in boundary separation and nondecision time ($F(2106) = 12.4$ and $F(2106) = 60.7$), plus decreases in drift rates, and the effect of IQ was due to decreases in nondecision time ($F(1106) = 13.8$), plus increases in drift rates. The IQ effect on nondecision time was not a large effect and did not replicate Ratcliff et al. (2010). What it might show is that in this experiment, higher-IQ participants were a little faster at encoding, memory access, and/or response output.

For priming, the older participants adopted more conservative decision criteria for item recognition than the college-age participants to the extent that there was no significant effect of age on RTs. The older participants' differences between primed and unprimed drift rates were smaller than those of the college-age participants' but the older participants' wider boundaries translated into about the same size RT effects as the college-age participants.

For the variability parameters of the model (across-trial variability in drift rates, starting point, and nondecision time), there was larger variability in their estimates relative to drift rate, boundary separation, and nondecision time (for discussion see Ratcliff & Tuerlinckx, 2002) and none were significantly different across age or IQ groups.

Summary

There are four main findings (Figs. 3 and 4). First, for unrelated pairs, the decrease in drift rates with age for associative recognition was much larger than that for item recognition, with associative recognition drift rates near zero for the older participants, replicating Ratcliff et al. (2011).

Second, for unrelated pairs, the differences between primed and unprimed drift rates showed a pattern similar to that for associative recognition: the decrease with age was large, with drift rate differences near zero for the older participants.

Third, relatedness affected memory in the same way for associative recognition and priming, with drift rates for associative recognition and drift rate differences for priming both much larger for related pairs than unrelated pairs, and the drop off with age was much less.

Fourth, drift rates increased with IQ for item recognition and associative recognition with related and unrelated pairs, and drift rate differences increased with IQ for priming with related and unrelated pairs. There were two exceptions to this pattern. One was that priming for unrelated pairs for the college-age participants decreased with IQ, which we attribute to unexplained variability in the data. The other was that, for the 75–90 year olds for unrelated pairs, the drift rate difference for priming and the associative recognition drift rate were near zero for both levels of IQ.

Table 9
ANOVA results for boundary separation and nondecision time.

	Factor	*df*	Item recognition		Associative recognition	
			F	MSE	*F*	MSE
Boundary	Age	2	12.42#	.0542	14.13#	.0574
	IQ	1	.60	.0026	.56	.0023
	AgexIQ	2	.33	.0014	.95	.0038
Nondecision time	Age	2	60.68#	.4548	30.66#	.4079
	IQ	1	13.76#	.1031	4.40*	.0586
	AgexIQ	2	2.17	.0162	1.62	.0216

Note. The denominator degrees of freedom (*df*) is 106.

Correlations among model parameters and IQ

Drift rates

If the diffusion model's account of the data is appropriate, then there should be positive correlations among the drift rates for the various tasks and conditions (Table 10): a participant who does well in one should also do well in the others (e.g., Ratcliff et al., 2011). Also, there should be positive correlations among the drift rates and IQ. Although there are exceptions, the data overall follow these patterns, supporting the diffusion model's account of the data. In the next paragraphs, we discuss the observed correlations in more detail. The results show what would be expected if memory is correlated across individuals and if priming and associative recognition are tapping the same information in memory.

Table 10 shows correlations among drift rates. For item recognition, the drift rates used to calculate the correlations were the sums of the absolute values of the drift rates for old and new filler items, divided by 2 and collapsed over word frequency. For associative recognition, the drift rates used were the sum of the absolute values of drift rates for the intact and rearranged pairs, divided by 2. For priming, the drift rate differences between the primed and unprimed conditions were used.

The correlations that involve drift rates for item recognition were the largest because they were based on the largest numbers of observations. Indeed, except for priming for unrelated pairs for the older two groups, the correlations between item drift rates and IQ and between item drift rates, associative drift rates, and priming drift rates for related and unrelated pairs were significantly positive.

The correlations based on associative information were the next largest because, even though there were low numbers of observations, there were large differences in accuracy between intact and rearranged conditions. As should be the case, the correlations between drift rates for related and unrelated pairs were significant.

If associative recognition and priming depend on the same information in memory, then there should be positive correlations between their drift rates. However, such correlations might not be strong because drift rates for priming and for associative recognition depend on relatively few numbers of observations. Also, for priming, the primed and unprimed drift rates are close to each other which reduces correlations even more relative to those for the other drift rates. Nevertheless, for related pairs, the correlations between drift rates for associative recognition and drift rate differences for priming were significantly positive for the college-age and 60–74 year old participants. For the 75–90 year olds, the correlation was not significant because their performance was near floor. For unrelated pairs, there were no significant correlations between drift rates for associative recognition and differences in drift rates for priming because all these measures were near floor.

Drift rates and IQ

IQ correlated with drift rates for item recognition and (to a smaller degree) with drift rates for associative recognition, but not with drift rate differences for priming (with one exception). This is what would be expected given the relative values of the drift rate estimates for item recognition, associative recognition, and priming.

Nondecision and criteria parameters

For these parameters, other studies have found strong correlations across experiments (RTM). In this study, the nondecision parameters for item recognition and associative recognition correlated .68, .49, and .36 for college-age, 60–74 year old, and 75–90 year old participants, and the boundary separation parameters correlated .50, .79, and .73, respectively. Thus these components of processing appear to be consistent for participants across the two tasks.

Reliability of correlations

To determine the upper limit for correlation coefficients given the relative sizes of individual differences and variability in parameter estimates, we conducted a simple Monte Carlo simulation. First, we generated two sets of 40 random numbers from a normal distribution with *SD* = 2, with correlations between the two sets 1, .8, or .6. Then we added variability to each of these numbers by adding to the number, a value chosen from a normal distribution with the number as its mean and *SD* = 1. Calculating correlations between the original and the new numbers, the original correlations were well-recovered—.80, .64, and .48—but less than the original values because of the added variability. The conclusion from this demonstration is that if the *SD* in individual differences is twice the size of measurement error, the correlation coefficient is reduced by .8 (e.g., the ratios are .8/1, .64/.8, and.48/.6) of the correlation due to individual differences alone (i.e., without measurement error). If the *SD* in individual differences is more than twice the size of measurement error, the

Table 10
Correlations among drift rates differences and IQ.

Participant group	Condition	Related priming difference (Rp)	Non-related priming difference (Np)	Item difference old-new (I)	Associative difference related pairs (Ra)	Associative difference non-related pairs (Na)
College age	IQ	0.06	−0.05	**0.49**	**0.40**	**0.48**
	Rp		0.08	**0.37**	**0.33**	0.18
	Np			**0.34**	0.17	0.16
	I				**0.71**	**0.66**
	Ra					**0.73**
60–74 year olds	IQ	**0.31**	0.05	**0.58**	0.24	**0.48**
	Rp		0.19	**0.51**	**0.32**	**0.34**
	Np			0.20	0.06	0.15
	I				**0.68**	**0.66**
	Ra					**0.67**
75–90 year olds	IQ	0.27	−0.17	**0.68**	**0.68**	0.24
	Rp		0.05	**0.35**	0.18	−0.11
	Np			−0.20	−0.23	−0.12
	I				**0.54**	**0.33**
	Ra					**0.35**

Critical value of the correlation coefficient for 39 df = .30. The bold correlations are significant.

correlation coefficient will be much closer to the correlation due to individual differences alone.

Illustrating modeling priming and associative recognition with a memory model

As Fig. 3 shows, the effect of relatedness was larger for associative recognition than priming. This larger effect can be explained by current models of memory (e.g., Ratcliff & McKoon, 1988). As an example, consider the SAM model (Gillund & Shiffrin, 1984), illustrated in Table 11.

In this model, memory is made up of associations of varying strengths between representations of information in memory (“images”) and potential cues to memory, where the strengths between images and cues vary between 0.0 and 1.0. Table 11 shows an example for a memory store with five items. Cue A is associated with memory image B with strength 1.0, with memory image C with strength 0.2, with memory image D with strength 0.2, and so on. When a cue is presented to memory, the strength with which it evokes information in memory is the sum of the strengths over all the images in memory. If cue A is presented, the strengths between A and B, A and C, A and D, A and E, and A and F are summed to give a value of familiarity upon which a decision about A is made.

Ratcliff and McKoon (1988) used SAM as one way of implementing a “compound cue” model to explain item-to-item priming. The assumption was that the cue to memory is not just a single item but rather a compound made up of all the items in short-term memory. For present purposes, we assume that there are only two items in short-term memory, i and j in Table 11. Each item in short-term memory is weighted as appropriate for the task to which the model is applied (with the weights summing to 1.0). The total familiarity of the $i-j$ compound is $F=\sum_k S_{ik}^{w1}S_{jk}^{w2}$, i.e., it is the sum over memory images k of the product of the strengths from cue i to the image and from cue j to the image, with the strengths weighted by w_1 and w_2.

Table 11
Retrieval structure for SAM model for recognition.

Cues	Memory images					
	A	B	C	D	E	F
A	1	1	0.2	0.2	0.2	0.2
B	1	1	1	0.2	0.2	0.2
C	0.2	1	1	1	0.2	0.2
D	0.2	0.2	1	1	1	0.2
E	0.2	0.2	0.2	1	1	1
F	0.2	0.2	0.2	0.2	1	1

The larger effect of relatedness on associative recognition than on priming can be interpreted with this model. For item-to-item priming, the test words are presented sequentially, and so the most recent test item (the one about which a decision is to be made) is weighted more heavily than the preceding test item. In the example in Table 11, if the weightings are 0.9 on a test word and 0.1 on the test word that immediately precedes it, then the resulting values of familiarity are 4.28 if the preceding word is related to A and 4.05 if it is not, a priming effect of 0.23. In contrast, for associative recognition, the judgment is made about the two words simultaneously and so they are weighted equally (Clark & Shiffrin, 1992). The resulting values of familiarity are 4.09 for intact pairs and 3.47 for rearranged pairs, a difference of 1.42. Thus with reasonable assumptions for the compound cue model as implemented in the SAM memory model, the prediction is that relatedness has a larger effect on associative recognition than on priming. This prediction holds even though the strengths of the test words as cues to memory are the same for the two tasks.

We should note that this example is an illustration of how a compound cue model embedded within a global memory model might account for priming effects. It does not address any of the issues that have led to the development of the second wave of global memory models (Dennis & Humphreys, 2001; McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997).

The compound cue and SAM models can also be used to interpret the results of a study by Cohn and Moscovitch (2007). They used an associative recognition task ("were the two words of a test pair studied in the same pair at study") and what they called an associative reinstatement task (respond "old" if both members of tested pair were studied, "new" if not). For their associative recognition task, the two words of a test pair would be weighted equally, and memory would be probed by the compound cue. For their associative reinstatement task, memory would be probed with individual test items, perhaps with some small weight on a compound of the two cues with the two cues weighted equally (see Clark & Shiffrin, 1987). Thus, in the associative reinstatement task, the assumption would be that to a large degree, participants performed two separate item recognitions. (To be consistent with our assumption about priming, we could assume single item tests would be modeled in the same way as item recognition priming, by assuming that each item recognition probe was actually a compound with a low weighting of the other item of the pair.)

Cohn and Moscovitch conducted some additional manipulations in their studies. First, they had participants generate sentences in a deep encoding condition and this provided retrieval cues for recall strategies. There is absolutely no doubt that in associative recognition tasks, if the study material is encoded appropriately, then recall from one of the cues can be used to generate the other which then is evidence that the pair was related. Such recall processes are slow and in Cohn and Moscovitch's study, the associative recognition difference (intact-rearranged) was reduced by a factor of 5 with a 1000 ms deadline. This deadline is a little longer than the mean RT (938 ms) for our associative recognition results for college-age participants. In another experiment, they manipulated speed and accuracy instructions; in this case, even in the speed condition, mean RTs were over 1.2 s and the reduction in the intact/rearranged difference was much smaller. Thus, their conclusion about the use of recollection (recall) in this paradigm is likely to be correct. However, it is likely in paradigms (like ours) with speeded tests and no special encoding opportunities, associative recognition is based on associative information as assumed by the global memory models.

Discussion

There are three main findings from this experiment: the diffusion model met the challenges offered by the data; the data for priming behaved in much the same ways as the data for associative recognition; and the drift rates provided by the model showed a complex pattern of correlations and dissociations among item recognition, associative recognition, and priming. The model's ability to account for the data is prerequisite to other conclusions, so we begin with that.

Diffusion model analyses

Explaining the data

One goal was to determine whether the diffusion model could explain all of the data, including priming effects. While the diffusion model has frequently been applied to item recognition data and sometimes to associative recognition data, it has not previously been applied to priming in item recognition. On the surface, it might seem that, if the model can explain performance on item recognition and associative recognition, then, because priming is, in some sense, a combination of the two, depending on both item information and associative information, it should be able to explain priming effects. But, as discussed in the introduction, this is not necessarily the case. Some models of memory separate memory for item information from memory for associative information (e.g., Dennis & Humphreys, 2001; Gillund & Shiffrin, 1984; McClelland & Chappell, 1998; Murdock, 1982; Reder et al., 2000; Rotello et al., 2004; Shiffrin & Steyvers, 1997; Wixted, 2007; Wixted & Stretch, 2004; Yonelinas, 1994, 1997). If item and associative information were indeed separate in memory, then the fact that the diffusion model gives good accounts of item recognition and associative recognition does not necessarily imply that it would give a good account of priming.

However, as shown by the chi-square values in Table 7, the model did well in all three cases. The model gave a full account of all the data: accuracy and mean RTs for correct and error responses, the relative speeds of correct and error responses, and the shapes of the RT distributions. The model explained these data for older and younger participants, higher- and lower-IQ participants, related and unrelated pairs of words, and high- and low-frequency words.

We stress that priming effects represent a significant challenge for the model because they must be built on top of item recognition effects. That is, explanations of priming effects can only be considered if the effects of other variables on item recognition are accounted for (word frequency, age, and IQ).

It should be noted that the diffusion model is highly falsifiable. The strongest constraint on the model is that it must fit right-skewed RT distributions and the ways they change over conditions. Specifically, as difficulty increases, increases in mean RT come mainly from increases in the spread of the distribution, not shifts in the location of the distribution. Ratcliff (2002) generated several sets of data for which the distributions changed with difficulty in ways that are plausible but never observed in real data. The model could not fit any of them.

It should also be noted that it is likely that other sequential sampling models, such as the one presented by Usher and McClelland (2001), would produce interpretations similar to the diffusion model for the data that are reported here (Ratcliff, Thapar, Smith, & McKoon, 2005).

Scaling

The model addresses two criterial problems for data interpretation. First, accuracy and RTs are measured on different scales. For the experiment reported here, this was especially problematic because priming is usually measured in terms of RTs and associative recognition is usually measured in terms of accuracy. Also, priming effects were much larger on RTs than accuracy. Using the model to interpret the accuracy and RT data together puts the two measures on the same scale. The fact that the model can accommodate large RT effects for priming (with small accuracy effects) and large accuracy (and RT) effects for associative recognition, and offer insightful interpretations of them, provides strong new support for the model.

The second problem is that participants often differ in their baseline RTs. This means that the sizes of the effects of experimental variables on the performance of different individuals cannot be directly compared. This problem was evident in the priming data reported here. For instance, priming effects were considerably larger for the 75–90 year old participants (141 ms) than the 60–74 year old participants (87 ms).

The differences in priming effects between the age groups could have come about because the difference in drift rates between primed and unprimed test words for the older participants was larger than for the college-age participants (perhaps because the older participants exerted more effort in learning the pairs). Or the differences between the groups might have come about because the groups set different speed-accuracy criteria, with the older participants requiring that more information be accumulated before executing a response. It turned out that both were true: the older participants' criteria were set farther apart and the difference between their primed and unprimed drift rates was (somewhat) smaller than for the college-age participants.

Power

The model adds significantly to the power of an experiment to observe the effects of experimental variables. Because all of the parameters of the model are determined by accuracy and RTs together, effects not significant in terms of accuracy or RTs can become significant in terms of the parameters of the model. This is especially effective for drift rates: Accuracy and RTs that are quite noisy can be converted into much less noisy drift rates.

The fact that there was a significant effect of age on drift rates for priming illustrates the model's ability to handle the "small-n" problem. There were only 18 and 12 observations per condition for related and unrelated pairs, respectively. However, the data for the test words in the large-n conditions (high- and low-frequency words, old and new words) largely determined the boundary and nondecision parameters of the model, and so reliable estimates of drift rates could be extracted for the priming conditions. In addition, the model's treatment of drift rates for the small-n stimuli produced variabilities in parameter estimates that were small enough to allow comparisons of drift rates among individuals.

Dealing with the small-*n* problem is a major development for the model. The solution to the problem opens up large new areas of research for application of the model and all its benefits. For one instance, the White et al. (2010a, 2010b) studies described above show application of the model to issues in clinical psychology. For another instance, the model can now be applied to many issues in psycholinguistics, as we are currently doing to investigate differences among individuals in their abilities to draw inferences from texts they have read (McKoon & Ratcliff, in preparation).

Numbers of observations

For item recognition, for the four conditions made up of low- and high-frequency words and old and new words, the large number of test words meant that the parameters of the model were very well-estimated. The standard deviations in the estimates for individuals were small (e.g., 0.03–0.06 for drift rates, Table 4, scaled by the square root of the number of observations, Ratcliff & Tuerlinckx, 2002) relative to the standard deviations across participants (e.g., 0.1–0.2 for drift rates for college-age participants, Table 8).

The differences in drift rates between the primed and unprimed test words were less well estimated, in part because the numbers of observations were small and in part because the difference between primed and unprimed drift rates was small (e.g., for college-age participants, the mean difference was 0.127 with a *SD* of 0.136). This suggests that the numbers of observations for priming (12 for unrelated pairs, 18 for related pairs) were smaller than is desirable. Nevertheless, the significantly positive correlations between priming drift rates and item recognition drift rates support the conclusion that the model extracted meaningful individual differences for priming.

For associative recognition, drift rates were estimated well enough to provide individual differences analyses, with the standard deviations in the estimates for individual participants about two times smaller than the standard deviations across participants. In a previous study of associative recognition (Ratcliff et al., 2011), there were about 300 observations per participant. The reduction in number of observations for the experiment reported here produced larger standard deviations in parameter values for each individual participant by about 1.7 (square root of 3), but the standard deviations were still just small enough to allow comparisons among individuals. Also, the fact that the correlations of the associative recognition parameters with the item recognition parameters were significant supports the conclusion that the model extracted meaningful estimates for associative recognition. Note that a reduction in number of observations can introduce biases into parameter values (e.g., Ratcliff, 2008) but the biases are not large and they are consistent across participants, so they do not affect interpretations of data.

It should be noted that a small number of observations has less impact on associative recognition than on priming. For associative recognition, there are two different responses (intact vs. rearranged) with opposite signs of drift rates (positive vs. negative) and the large difference in drift rate between them allows better estimates in the differences than for priming.

Priming and associative recognition

The drift rates for associative recognition and priming suggest that they are based on the same information in memory. First, for the related pairs, there were significant correlations between associative recognition drift rates and priming drift rates (except for the 75–90 year olds, whose drift rates were near zero).

Second, drift rates for priming and associative recognition tracked each other as a function of age and whether the two items in a study pair were related. For unrelated pairs, for both measures, the decrease in drift rates from college-age participants to the older groups reached near floor (even though item recognition for items in these pairs was well above floor). For related pairs, for both measures, the decrease in drift rates was much less, with drift rates for 75–90 year olds no worse than 50% lower than those for college-age participants.

Third, drift rates for priming and associative recognition tracked each other as a function of IQ. For related pairs, drift rates increased with IQ in both cases. For unrelated pairs, drift rates for the 60–74 year olds increased with IQ (drift rates for the college-age participants decreased, a finding for which we have no explanation). Drift rates for the 75–90 year olds were near zero for both priming and associative recognition for both lower- and higher-IQ participants.

For these three reasons, the data and the model's application to them indicate that priming and associative recognition depend on the same associative information in memory. This finding is an achievement for the model because it allows comparison between associative recognition, which is usually measured by accuracy, and priming, which is usually measured by RT differences.

Using the compound cue model (Ratcliff & McKoon, 1988), we showed one way in which the larger effect of relatedness on associative recognition could come about. In the model, test items are cues to memory: a test item is matched against memory and the resulting value of familiarity determines its drift rate, which in turn determines (in part) the accuracy and RT of the response to it. In the compound cue model, what is matched against memory is not a single item but instead the combination of all the test items in short-term memory at the same time. The test items can be equally weighted in the calculation of familiarity or differentially weighted. The two words of a pair for a test of associative recognition are given the same weight because they are displayed at the same time. For the two words of a priming pair, the target is displayed after the prime, and so the prime is weighted less than the target. In this way, as we showed above, relatedness has a larger effect on associative recognition than on priming.

Correlations and dissociations

Comparing item recognition and associative recognition, the results of the drift rate analyses confirm results reported in Ratcliff et al. (2011). One result was that drift rates for item recognition show much smaller decrements with age than drift rates for associative recognition. A second was that the correlations between individuals' drift rates on item and associative recognition were strong. In the experiment reported here, they were significant for all three age groups, for related and unrelated pairs (ranging from .33 to .71). A third result was significant correlations between drift rates for item recognition and priming, which means that participants who are better at item recognition also have larger priming effects.

The combination of these data provides a challenging pattern for models of memory to explain. A model must provide correlated levels of drift rates between item recognition, associative recognition, and priming, but at the same time, it must dissociate drift rates for item recognition from drift rates for priming and associative recognition as a function of age. Models proposing that item and associative recognition are based on the same information in memory have problems because they cannot accommodate the differential changes with age between item information, on the one hand, and priming and associative memory on the other (e.g., Old & Naveh-Benjamin, 2008). At the same time, models that base item recognition on completely different information in memory than associative recognition and priming have problems because, at least on the face of it, they cannot accommodate the high correlations between drift rates for item recognition, associative recognition, and priming.

The priming results have added a layer of complexity to attempts to reconcile the patterns of correlations with the patterns of dissociations. Priming depends on item recognition, which means that associative information must be accessible to item recognition decisions. This, then, provides another reason that models cannot base item recognition on different information in memory than associative recognition.

Conclusion

From our data and analyses, we argue that priming and associative recognition have a common basis in memory. In one sense this might be an uncontroversial claim, but it is difficult to see how the claim could be supported except by the two main features of our analyses. First, the diffusion model allowed RTs and accuracy for item and associative recognition to be compared on a common metric (drift rate), and second, we used three different variables to compare priming and associative recognition (age, IQ, pair relatedness). Finding parallel effects of these variables on priming and associative recognition indicates that we should view priming and associative recognition as two complementary ways of examining new associations in memory.

Acknowledgment

Preparation of this article was supported by NIA Grant R01-AG17083.

A. Appendix. Supplementary experiment

In the main study, for the pairs tests, if two words were related at test (e.g., the test pair was "mortician cadaver)," then it was always the case that they had been studied in the same pair. This was changed for the supplementary study. The test lists included pairs of related words for which the two words of the pair had been studied in different pairs. The two words were then tested in the same pair, for which the correct response was "rearranged." For example, for the pair "mortician cadaver," "mortician" would be studied in one pair and "cadaver" in a different pair. Then they would be combined at test into the pair "mortician cadaver," for which the correct response would be "rearranged." The participants in the experiment were from the same pool as the older adults in the experiment above. As described below, they were well able to distinguish between test pairs for which the two related words had been studied together and test pairs for which the two related words had been studied in different pairs.

A.1. Method

The pools of words were the same as those in the main study, and the procedure was the same. All of the test lists were "pairs" test lists.

The to-be-learned lists of items, 16 pairs for each of 12 blocks, were constructed as follows: For four of the pairs, the two words were related; for four of the pairs, the words were unrelated and both were high-frequency words; for four of the pairs, the two words were unrelated and both low-frequency words; for two of the pairs, the left-hand member came from a related pair and the right-hand member was a high-frequency word; and for two of the pairs, the left-hand member was a high-frequency word and the right-hand member came from a related pair.

There were 14 pairs in each test list. For the pairs for which the two words were related and studied in the same pair, there were four intact pairs in half the test lists and four rearranged pairs in the other half. For the pairs made up of two high-frequency words, there were two intact and two rearranged, and for the pairs made up of two low-frequency words, there were two intact and two rearranged. The remaining two pairs were made up of related words that had been studied in different pairs (e.g., the test pair was "mortician cadaver").

The 20 participants were 60–74 year olds, all except four with IQ's above 106 (those below 106 were 72, 83, 91, and 97).

A.2. Results

The same cutoffs were used as in the main experiment, eliminating about 2% of the data. The results were similar to those of the main experiment. For pairs that were related at study, the probability correct for intact pairs was .79 (mean RT 1085 ms) and the probability correct for rearranged pairs was 0.89 (mean RT 1252 ms). For pairs that were not related at study or test, the probabilities correct were .57 (mean RT 1399 ms) and .72 (mean RT 1500 ms).

For related words that were studied in different pairs but tested together, the probability correct for responding "rearranged" was .63 (mean RT 1519 ms). If participants had been responding only to relatedness at test, they would have responded "intact" with probability .79 and "rearranged" with probability .21.

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). Oxford University Press.

Balota, D. A., & Duchek, J. M. (1989). Spreading activation in episodic memory: Further evidence for age independence. *Quarterly Journal of Experimental Psychology, 41A*, 849–876.

Bastin, C., & Van der Linden, M. (2006). The effects of aging on the recognition of different types of associations. *Experimental Aging Research, 32*, 61–77.

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. (1987). Recognition of multiple-item probes. *Memory & Cognition, 15*, 367–378.

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.

Cohn, M., & Moscovitch, M. (2007). Dissociating measures of associative memory: Evidence and theoretical implications. *Journal of Memory and Language, 57*, 437–454.

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.

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*, 452–477.

Duffy, S. A., Henderson, J. M., & Morris, R. K. (1989). Semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 791–801.

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.

Faust, M. E., Balota, D. A., & Spieler, D. H. (2001). Building episodic connections: Changes in episodic priming with age and dementia. *Neuropsychology, 15*, 626–637.

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.

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

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356–388.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.). *The psychology of learning and motivation* (Vol. 22, pp. 193–225). New York, NY: Academic Press.

Howard, D. V., Heisey, J. G., & Shaw, R. J. (1986). Aging and the priming of newly learned associations. *Developmental Psychology, 22*, 78–85.

Humphreys, M. S., & Bain, J. D. (1983). Recognition memory: A cue and information analysis. *Memory & Cognition, 11*, 583–600.

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.

Kaufman, A. (1990). *Assessing adolescent and adult intelligence*. Boston, MA: Allyn and Bacon.

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: Wiley.

Laver, G. D. (2009). Adult aging effects on semantic and episodic priming in word recognition. *Psychology and Aging, 24*, 28–39.

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging, 8*, 34–43.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology, 42*, 333–376.

MacKay, D. G., & Burke, D. M. (1990). Cognition and aging: New learning and the use of old connections. In T. M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 213–263). Amsterdam: North Holland.

Mandler, George (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*, 252–271.

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.

McKoon, G., & Ratcliff, R. (in preparation). Aging and predictive inferences using a diffusion model analysis.

McKoon, G., & Ratcliff, R. (1979). Priming in episodic and semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18*, 463–480.
McKoon, G., & Ratcliff, R. (1980). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning and Verbal Behavior, 19*, 369–386.
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. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89*, 609–626.
Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. *Psychology and Aging, 7*, 257–270.
Myerson, J., Hale, S., Chen, J., & Lawrence, B. (1997). General lexical slowing and the semantic priming effect: The roles of age and ability. *Acta Psychologica, 96*, 83–101.
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.
Naveh-Benjamin, M., Guez, J., Kilb, A., & Reedy, S. (2004). The associative deficit of older adults: Further support using face-name associations. *Psychology and Aging, 19*, 541–546.
Naveh-Benjamin, M., Guez, J., & Shulman, S. (2004). Older adults' associative deficit in episodic memory: Assessing the role of decline in attentional resources. *Psychonomic Bulletin & Review, 11*, 1067–1073.
Naveh-Benjamin, M., Hussain, Z., Guez, J., & Bar-On, M. (2003). Adult age differences in episodic memory: Further support for an associative-deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 826–837.
Neely, J. H., & Durgunoglu, A. (1985). Dissociative episodic and semantic priming effects in episodic recognition and lexical decision tasks. *Journal of Memory and Language, 24*, 466–489.
Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms*. <http://www.usf.edu/FreeAssociation/>.
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.
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. (2002). A diffusion model account of reaction time and accuracy in a two choice brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review, 9*, 278–291.
Ratcliff, R. (2008). The EZ diffusion method: Too EZ? *Psychonomic Bulletin & Review, 15*, 1218–1228.
Ratcliff, R., & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of sentences. *Journal of Verbal Learning and Verbal Behavior, 17*, 403–417.
Ratcliff, R., & McKoon, G. (1981). Does activation really spread? *Psychological Review, 88*, 454–462.
Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95*, 385–408.
Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*, 873–922.
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., Thapar, A., & McKoon, G. (2011). Individual differences, aging, and IQ in four memory tasks. *Journal of Experimental Psychology: General, 140*, 464–487.
Ratcliff, R., Thapar, A., Smith, P. L., & McKoon, G. (2005). Aging and response times: A comparison of sequential sampling models. In J. Duncan, P. McLeod, & L. Phillips (Eds.), *Speed, control, and age*. Oxford, England: Oxford University Press.
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 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.
Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403–428.
Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM: Retrieving effectively from memory. *Psychonomic Bulletin & Review, 4*, 145–166.
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.
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.
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.
Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*, 550–592.
Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology, 21*, 641–671.
Wechsler, D. (1997). *The Wechsler adult intelligence scale III*. Harcourt Brace, San Antonio, CA: Psychological Corporation.
White, C., Ratcliff, R., Vasey, M., & McKoon, G. (2009). Dysphoria and memory for emotional material: A diffusion model analysis. *Cognition and Emotion, 23*, 181–205.
White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010a). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology, 54*, 39–52.
White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010b). Anxiety enhances threat processing without competition among multiple inputs: A diffusion model analysis. *Emotion, 10*, 662–677.
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.