



Published in final edited form as:

Cogn Psychol. 2012 May ; 64(3): 127–160. doi:10.1016/j.cogpsych.2011.11.001.

Modeling Fan Effects on the Time Course of Associative Recognition

Darryl W. Schneider and John R. Anderson
Carnegie Mellon University

Abstract

We investigated the time course of associative recognition using the response signal procedure, whereby a stimulus is presented and followed after a variable lag by a signal indicating that an immediate response is required. More specifically, we examined the effects of associative fan (the number of associations that an item has with other items in memory) on speed–accuracy tradeoff functions obtained in a previous response signal experiment involving briefly studied materials and in a new experiment involving well-learned materials. High fan lowered asymptotic accuracy or the rate of rise in accuracy across lags, or both. We developed an Adaptive Control of Thought–Rational (ACT-R) model for the response signal procedure to explain these effects. The model assumes that high fan results in weak associative activation that slows memory retrieval, thereby decreasing the probability that retrieval finishes in time and producing a speed–accuracy tradeoff function. The ACT-R model provided an excellent account of the data, yielding quantitative fits that were as good as those of the best descriptive model for response signal data.

Keywords

response signal procedure; fan effect; associative recognition; memory retrieval; cognitive modeling

Associative recognition—the process of determining whether two items were previously experienced together—is not instantaneous. It takes time to probe memory for associative information and there are many variables that affect the time and accuracy of retrieval. In the present study we focus on the effects of associative fan, which refers to the number of associations that an item has with other items in memory. Past research has shown that the time taken to recognize an item becomes longer as its fan increases, a finding known as the fan effect (Anderson, 1974; for reviews, see Anderson, 2007; Anderson & Reder, 1999). As we discuss below, the fan effect is thought to be due to a decrease in associative activation that slows memory retrieval. However, little is known about the fine-grained temporal structure of the slowed retrieval process.

To address this issue, we investigated fan effects on the time course of associative recognition using the response signal procedure (Doshier, 1976, 1979; Reed, 1973, 1976;

© 2011 Elsevier Inc. All rights reserved.

Address correspondence to: Darryl W. Schneider Department of Psychology Carnegie Mellon University Pittsburgh, PA 15213 USA
Phone: 1-412-334-7758 dws@cmu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Darryl W. Schneider and John R. Anderson, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA.

Schouten & Bekker, 1967; Wickelgren, 1977). In this procedure, a stimulus is presented and followed after a variable lag by a signal indicating that an immediate response is required. Varying the response signal lag allows one to map out the time course of processing in the form of a speed–accuracy tradeoff function that shows how accuracy changes over time. Below, we provide an overview of the response signal procedure and we review two previous studies involving fan manipulations (Doshier, 1981; Wickelgren & Corbett, 1977). We then report the results of a new response signal experiment involving well-learned materials in a paradigm that is more typical of fan-effect studies.

At the heart of the present study is the development and evaluation of a formal model of fan effects on the time course of associative recognition. We describe a model based on the Adaptive Control of Thought–Rational (ACT-R) theory, which has a long history of success in cognitive psychology (see Anderson, 2007; Anderson & Lebiere, 1998). More specifically, we show how the extant ACT-R model of the fan effect, which applies to mean reaction time (RT), can be extended in a straightforward manner to account for data from the response signal procedure. We demonstrate that our ACT-R model not only accounts for fan effects on the time course of associative recognition, but it does so with quantitative fits that are as good as those of the best descriptive model for response signal data.

The Fan Effect and ACT-R

The fan effect is often demonstrated in the fact retrieval paradigm (Anderson, 1974), wherein subjects memorize a set of fictional facts (e.g., person–location pairs):

The hippie is in the park.

The hippie is in the factory.

The detective is in the library.

The tourist is in the factory.

Some items occur in only one fact (e.g., *detective* occurs in only the third fact) whereas other items occur in more than one fact (e.g., *hippie* occurs in the first and the second facts). The number of facts in which an item occurs is the fan of that item (e.g., *hippie* has a fan of 2). After memorizing the facts during a study phase, subjects perform a recognition task during a test phase in which they have to distinguish between targets (studied facts) and foils (non-studied facts, which are usually rearranged items from the studied facts; e.g., *The detective is in the park*).

The main finding from fan manipulations in the fact retrieval paradigm is the fan effect: recognition takes longer for items with higher fans (e.g., Anderson, 1974; King & Anderson, 1976; Pirolli & Anderson, 1985; for reviews, see Anderson, 2007; Anderson & Reder, 1999). Recognition accuracy also tends to be lower for items with higher fans, as reflected in higher false alarm rates but mixed effects on hit rates (Dyne, Humphreys, Bain, & Pike, 1990; Verde, 2004; see also Postman, 1976). Fan effects have been observed not only in the retrieval of fictional facts, but also in the retrieval of real-world knowledge (Lewis & Anderson, 1976), faces (Anderson & Paulson, 1978), and alphabet-arithmetic facts (White, Cerella, & Hoyer, 2007; Zbrodoff, 1995). In addition, fan effects have been observed not only in behavioral data (e.g., RT and accuracy), but also in neuroimaging data (e.g., differences in brain activation; Danker, Gunn, & Anderson, 2008; Sohn, Goode, Stenger, Carter, & Anderson, 2003; Sohn, Goode, Stenger, Jung, Carter, & Anderson, 2005). Thus, the fan effect is a robust phenomenon and consistent with the general principle of cue overload, which is the idea that a retrieval cue becomes less effective as it becomes associated with more items in memory (Surprenant & Neath, 2009; Watkins & Watkins, 1975).

The fan effect played an important role in the early development of the ACT-R theory of cognition (Anderson, 1976, 1983). ACT-R is a cognitive architecture in which a production system coordinates the activity of modules associated with perception, memory, and action (Anderson, 2007; Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004). Research on the fan effect helped shape the structure and functioning of the declarative memory module in ACT-R, which is a repository of knowledge ranging from fictional facts learned in an experiment (e.g., *The hippie is in the park*) to real-world information (e.g., *Ottawa is the capital of Canada*). Knowledge is represented in units called chunks, which can be retrieved from declarative memory and placed in the module's buffer for use by the rest of the ACT-R system.

The mechanism for retrieving chunks from declarative memory is formally specified in ACT-R (Anderson, 2007; Anderson & Lebiere, 1998). Each chunk has an activation level in declarative memory that is given by:

$$A_i = B_i + \sum_j W_j S_{ji}, \quad (1)$$

where A_i is the total activation of chunk i , B_i is its base-level activation, and the last term is the associative activation that the chunk receives from all sources j that are used as retrieval cues. Base-level activation reflects the frequency and recency with which the chunk has been used in the past, which provides an indication of how likely the chunk will be needed in the future (Anderson & Schooler, 1991). Associative activation reflects the strength of association between chunks in declarative memory:

$$S_{ji} = S + \ln [P(i|j)], \quad (2a)$$

where S_{ji} is the strength of association between chunks j and i , S is the maximum associative strength, and $P(i|j)$ reflects learning about the probability that chunk i will be needed when chunk j is used as a retrieval cue (based on the rational analysis of Anderson, 1990, 1991; see also Anderson & Reder, 1999). If all the chunks associated with chunk j occur with equal probability, which is a reasonable assumption in many contexts, then $P(i|j) = 1/fan_j$, where fan_j is the fan of chunk j . The strength of association between chunks can then be expressed in terms of fan:

$$S_{ji} = S - \ln(fan_j). \quad (2b)$$

Equation 2b indicates that as chunk j becomes associated with more chunks (i.e., its fan increases), its strength of association with each of those chunks decreases. The amount of associative activation for a chunk (the last term in Equation 1) is determined by weighting the strength of association by the amount of activation allocated to each source j (W_j) used as a retrieval cue for chunk i . Source activation is typically partitioned equally among all sources and sums to a constant (W), which implies that $W_j = W/J$, where J is the number of sources (Anderson, Reder, & Lebiere, 1996). Associative activation is summed across all sources and added to a chunk's base-level activation to give its total activation (Equation 1).

The total activation of a chunk determines the time taken to retrieve the chunk from declarative memory:

$$t_{\text{retrieve}} = F e^{-A}, \quad (3)$$

where t_{retrieve} is the retrieval time, A is the chunk's activation, and F is a parameter that scales retrieval time. Considering Equations 1–3 together, as the fan of a source increases, its strength of association with each chunk in declarative memory decreases (Equation 2b), resulting in less associative activation and, by extension, less total activation (Equation 1), yielding a longer retrieval time (Equation 3).

Equations 1–3 are the standard equations for declarative memory retrieval in ACT-R and represent the basic model for the fan effect. To understand how the model produces the fan effect, we return to the aforementioned fact retrieval paradigm where the recognition task is to decide whether a specific person–location probe was studied. The person and the location serve as sources of activation for retrieving facts from declarative memory. Each source provides an amount of associative activation based on the source activation allocated to it and its strength of association with facts in memory, with the latter being negatively related to the source's fan (Equation 2b). The associative activation summed across both sources contributes to the total activation of a specific fact (Equation 1) and determines the time taken to retrieve it (Equation 3). The model retrieves the fact that has the greatest total activation and compares it with the probe. If there is a match, as in the case of a target, then the model makes a “yes” response. If there is a mismatch, as in the case of a foil, then the model makes a “no” response. Thus, the model implements a recall-to-reject strategy for foils, consistent with extant theorizing about associative recognition (e.g., Malmberg, 2008; Rotello & Heit, 2000; Rotello, Macmillan, & Van Tassel, 2000).

The fan effect arises in the model from differences in associative activation from probes that have different fans. For example, from the list of person–location pairs presented earlier, the probe *The detective is in the library* involves person and location sources that each have a fan of 1 because each source occurs in only one fact, whereas the probe *The hippie is in the factory* involves person and location sources that each have a fan of 2 because each source occurs in two facts. We refer to these as Fan 1 and Fan 2 probes, respectively. From Equation 2b, the strength of association between probe sources and fact chunks in memory will be greater for the Fan 1 probe than for the Fan 2 probe. Consequently, the Fan 1 probe will produce more associative activation than will the Fan 2 probe (Equation 1), resulting in more total activation and a shorter retrieval time (Equation 3). The difference in retrieval times for probes that have different fans is the fan effect produced by the model.

The fan effect in the preceding example applies to targets, but the model produces a fan effect for foils in a similar way. For example, the probe *The tourist is in the library* is a foil that involves person and location sources that each have a fan of 1. However, unlike targets, both sources do not provide activation for the same fact in memory because that specific person–location pair was not studied. Instead, *tourist* is a source of activation for *The tourist is in the factory* and *library* is a source of activation for *The detective is in the library*. One of these two facts is retrieved and compared with the probe, yielding a mismatch that is used to reject the foil. In this example, each source has a fan of 1, but one could construct other foils that have higher fans. Given that retrieval works the same way for foils as for targets, the model produces a similar fan effect for foils.

The ACT-R model has produced good quantitative fits to empirical fan effects in several studies (e.g., Anderson, 1974; Anderson & Reder, 1999; Pirolli & Anderson, 1985; for a review, see Anderson, 2007). In addition, the basic principles governing associative activation in the model have been applied successfully to other cognitive phenomena,

including a variety of list memory effects (Anderson, Bothell, Lebiere, & Matessa, 1998; Anderson & Matessa, 1997) and set-size effects in multiple-choice behavior (Schneider & Anderson, 2011). However, an important limitation of the model's account of the fan effect in past studies is that its predictions applied only to mean RT. Attempts to account for RT data in greater detail and the relationship between RT and accuracy have been rare (for early exceptions concerning the latter, see Anderson, 1981; King & Anderson, 1976). The main objective of the present study was to take a step toward addressing this limitation by extending the model to account for fan effects on the time course of associative recognition, as reflected in speed–accuracy tradeoff functions obtained using the response signal procedure.

The Response Signal Procedure

It is well known that people can trade speed for accuracy in task performance, slowing down to make fewer errors and speeding up at the risk of making more errors (Pachella, 1974; Wickelgren, 1977). A popular method for investigating speed–accuracy tradeoff functions in recognition is the response signal procedure (Doshier, 1976, 1979; Reed, 1973, 1976; Schouten & Bekker, 1967; Wickelgren, 1977). In this procedure, a stimulus is presented for a yes–no recognition task and followed after a variable lag by a signal indicating that an immediate response is required (usually within 200–300 ms). The main dependent variable is accuracy as a function of the time available for task processing. Accuracy is often expressed as a d' measure to control for response bias and plotted against total processing time (lag + mean RT, where RT is defined as the time from response signal onset to the response). Varying the response signal lag allows one to map out the time course of processing in the form of a speed–accuracy tradeoff function.

An idealized example of a speed–accuracy tradeoff function from the response signal procedure is illustrated in Figure 1. At very short lags, accuracy is at chance because not enough time has elapsed for task processing to yield any useful response information. At very long lags, accuracy is at a high asymptote because enough time has elapsed for task processing to finish, often resulting in selection of the correct response. At intermediate lags, there is an intercept time at which accuracy begins to rise above chance, then accuracy continues to grow in a negatively accelerated manner until it reaches asymptote (see Figure 1). This accuracy data pattern is often described by a shifted exponential function (SEF):

$$d'(t) = \lambda \{1 - \exp[-\beta(t - \delta)]\}, \text{ if } t > \delta, \text{ else } 0, \quad (4)$$

where λ is the asymptote for accuracy, δ is the intercept time marking the transition from chance to above-chance accuracy, and β is the rate at which accuracy rises from chance to asymptote. The time variable t equals lag + mean RT to address the likely possibility that task processing does not stop precisely at the time the lag elapses and to account for any changes in RT as a function of lag. Indeed, RT typically becomes shorter as the lag becomes longer.

The SEF has been shown to provide a very good characterization of speed–accuracy tradeoff functions obtained using the response signal procedure in several studies (e.g., Doshier, 1976, 1981; Gronlund & Ratcliff, 1989; Hintzman, Caulton, & Levitin, 1998; Hintzman & Curran, 1994, 1997; McElree & Doshier, 1989; Wickelgren & Corbett, 1977; Wickelgren, Corbett, & Doshier, 1980). An alternative to Equation 4 that yields similar fits to response signal data is an expression for monotonic growth to a limit derived from the diffusion model (see Ratcliff, 1978, 1980; for applications, see Doshier, 1981, 1984b; Gronlund & Ratcliff, 1989; McElree & Doshier, 1989; Ratcliff & McKoon, 1982; Rotello & Heit, 2000).

Typically, the SEF is fit to the speed–accuracy tradeoff functions associated with different experimental conditions, allowing one or more of its parameters to vary across conditions. For example, if there are two conditions, then there are eight possible SEF variants based on whether each parameter (intercept, rate, or asymptote) is the same or different across conditions. Each SEF variant is fit to the data and model comparison techniques are used to determine which variant provides the best fit without excessive model complexity.

The parameters of the best SEF variant are often interpreted in terms of memory strength and retrieval dynamics. Differences in asymptote among conditions are thought to reflect differences in memory strength, such that a condition in which items are strongly represented in memory (e.g., as a consequence of extensive learning) will have a higher asymptote than a condition in which items are weakly represented. Differences in intercept and rate among conditions are thought to reflect retrieval dynamics—the nature of the process by which items are retrieved from memory. A condition with a shorter intercept (reflecting earlier onset of retrieval) or a higher rate (reflecting faster speed of retrieval) than another condition is considered to have faster retrieval dynamics.

The response signal procedure and fits of the SEF to speed–accuracy tradeoff functions have been used to investigate fan effects on the time course of associative recognition in two previous studies (Doshier, 1981; Wickelgren & Corbett, 1977). In Wickelgren and Corbett's experiment, subjects studied lists of words organized into pairs and triples. Pairs consisted of two words (denoted here as A and B), with A appearing on the left and B appearing on the right. Given that each word was associated with only one other word, pairs can be classified as Fan 1 items. Triples consisted of three words (denoted here as A, B, and C), with A appearing on the left and B and C appearing on the right. Given that each word was associated with two other words, triples can be classified as Fan 2 items.¹ Each pair or triple was presented for 3 s during a study phase, then associative recognition judgments were made in a test phase involving the response signal procedure.

The empirical speed–accuracy tradeoff functions from Wickelgren and Corbett (1977), averaged across subjects, are presented as points in Figure 2. It is clear that Fan 1 items have higher asymptotic accuracy than do Fan 2 items, but it is less obvious whether there are rate or intercept differences between fan conditions. The solid and dashed lines in Figure 2 represent the mean predictions from individual-subject fits of an ACT-R model and the best SEF variant, respectively, both of which we discuss in detail later. At this point, we simply note that the best SEF variant is one in which the asymptote and the intercept (but not the rate) differ between fan conditions, consistent with what Wickelgren and Corbett found with their fits. However, a very similar fit is obtained when the asymptote and the rate (but not the intercept) differ between fan conditions.

Wickelgren and Corbett's (1977) basic findings were replicated by Doshier (1981). In Doshier's experiment, subjects studied a list of three pairs of words (denoted here by letters) for a study–test sequence involving one of two conditions. In the independent condition, the pairs had no words in common (e.g., A–B, D–E, and F–C), which meant that each word had a fan of 1. In the interference condition, the first and third pairs shared a word (e.g., A–B, D–E, and A–C), which meant that the A word had a fan of 2. Each pair was presented for 3 s during a study phase, then associative recognition judgments were made in a test phase involving the response signal procedure. The critical results from Doshier concern the speed–accuracy tradeoff functions for Fan 1 and Fan 2 items in the independent and interference conditions, respectively. To avoid redundancy with our presentation of Wickelgren and

¹Wickelgren and Corbett (1977) instructed subjects that they would never be tested on the B–C association, but such an association was likely formed to some degree during the study phase.

Corbett's results, we simply note that Doshier found that the best SEF variant for her data was one in which the asymptote and either the rate or the intercept (but not both) differed between fan conditions, mirroring the findings of Wickelgren and Corbett.

The results from Wickelgren and Corbett (1977) and Doshier (1981) concerning fan effects on the time course of associative recognition are generally consistent with the standard ACT-R model of the fan effect. A difference in asymptote between fan conditions suggests a difference in memory strength, which is concordant with the ACT-R interpretation of fan affecting the strength of association between items in memory. A small difference in intercept or rate between fan conditions suggests a modest effect of fan on retrieval dynamics. Given that differences in strength of association affect the time for declarative memory retrieval in ACT-R, a fan effect on retrieval dynamics is consistent with the theory. However, it is an open issue as to whether ACT-R can account for the quantitative (not just qualitative) pattern of response signal data observed in previous studies. Before we address this issue, we report the results of a new experiment in which response signal data were collected in the context of the fact retrieval paradigm with well-learned materials.

A New Response Signal Experiment on the Fan Effect

One possible reason for the modest effects of fan on retrieval dynamics in the studies by Wickelgren and Corbett (1977) and Doshier (1981) is that their experiments involved brief study phases in which there was a limited opportunity to learn each study list (i.e., each item on a list was presented just once for only 3 s). A practical advantage of the brief study phase is that accuracy was kept below ceiling, making it easier to detect differences in asymptotic accuracy and facilitating the calculation of d' (which poses definitional problems when accuracy is perfect). However, a potential disadvantage is that the fan manipulation may not have been very strong because the representations of items in memory and the associations between them may have been somewhat weak due to limited learning. This raises the issue of whether the results of these previous studies can be replicated with well-learned materials that may elicit stronger fan effects.

To address this issue, we conducted a multi-session response signal experiment involving the fact retrieval paradigm. In a study phase at the start of the first session, subjects were presented with a list of person–location facts, half with a fan of 1 and the other half with a fan of 2. Subjects then completed a cued recall test in which they answered questions of the form *Where is the person?* and *Who is in the location?*. Each question had to be answered correctly three times, thereby ensuring that the facts were well-learned (Rawson & Dunlosky, 2011). In subsequent sessions there was an abbreviated study phase in which the cued recall test involved answering each question correctly once. Following the study phase, subjects completed a recognition test phase involving the response signal procedure. A person–location probe was presented on each trial and subjects had to distinguish between targets (studied facts) and foils (non-studied facts, which were rearranged persons and locations that maintained their fan status). The test probe was followed after one of eight lags by a response signal. Thus, our experiment was similar in many respects to the experiments of Wickelgren and Corbett (1977) and Doshier (1981), with the main difference being a more extensive study phase designed to promote better learning of the materials.

Method

Subjects

Ten individuals from the Carnegie Mellon University community each participated in five sessions for monetary compensation. There was one session per day for five consecutive days. The first session was 2 h in duration and subsequent sessions were each 1 h. Subjects

were paid at a rate of \$10/h, plus a bonus based on their performance (see below; the mean bonus payment was \$4 per session).

Apparatus

The experiment was conducted using Tscope (Stevens, Lammertyn, Verbruggen, & Vandierendonck, 2006) on computers that displayed stimuli on monitors and registered responses from QWERTY keyboards. Text was displayed onscreen in white 14-point Courier font on a black background. Auditory response signals were presented over headphones.

Materials

Each study fact was of the form *The person is in the location*. Study facts were created from lists of 24 persons and 24 locations (see Appendix A). Word length was 3–9 letters ($M = 6.25$, $SD = 1.59$) for persons and 4–10 letters ($M = 6.25$, $SD = 1.39$) for locations. Each subject received a random assignment of persons and locations to a study list of 32 person–location facts. Half of the facts had a fan of 1 (i.e., the person and the location each occurred in only one fact) and the other half had a fan of 2 (i.e., the person and the location each occurred in two different facts). The facts on the study list were the targets in the recognition test phase of the experiment. The foils in the test phase were drawn from a list of 32 non-studied facts created by rearranging persons and locations from studied facts such that their fan status was maintained (e.g., a Fan 2 foil was created using a person and a location from different Fan 2 studied facts).

Procedure

Subjects were seated at computers in private testing rooms after providing informed consent. Written instructions were presented to subjects and explained by the experimenter during the first session. The instructions were available for subjects to review in subsequent sessions if necessary. Each session was divided into a study phase and a test phase.

In the study phase for the first session, the study list of 32 facts was presented. Each fact appeared in the center of the screen for 5000 ms and was followed by a 500-ms blank screen. Subjects were instructed to read each fact and make an initial effort to memorize it. After all the facts were presented, subjects completed a cued recall test in which they answered questions of the form *Where is the person?* and *Who is in the location?*. There were 24 person questions and 24 location questions representing all studied items (see Appendix A). On each trial, a single question appeared in the center of the screen with an answer prompt below it. Subjects had to type the appropriate answer to the question based on the facts in the study list (e.g., when asked about a person, they had to recall the location(s) the person was in). Questions about Fan 1 and Fan 2 facts required one- and two-word answers, respectively. Multiple words were separated by commas and the Enter key was pressed to submit the answer. There was no time limit for giving the answer. If the answer was correct, there was no feedback and the next question appeared after a 500-ms blank screen. If the answer was incorrect, the word *INCORRECT* and the correct answer were displayed for 2500 ms, followed by the 500-ms blank screen.

An answer to a question about a Fan 1 fact was judged as incorrect if: (a) the answer was blank; (b) a single word was submitted, but it did not match the studied word; or (c) multiple words were submitted, even if one of them matched the studied word. An answer to a question about a Fan 2 fact was judged as incorrect if: (a) the answer was blank; (b) a single word was submitted, even if it matched one of the studied words; (c) two words were submitted, but one or both did not match the studied words; or (d) more than two words were submitted, even if one or two of them matched the studied words. The words for a Fan

2 answer could be entered in any order. Subjects were instructed to submit a blank answer if they could not recall anything. As noted above, they would then receive the correct answer as feedback, which served as a learning opportunity. They were also instructed to use any mnemonics or mental strategies that might help them memorize the facts (e.g., forming a semantic association between the person and the location in a fact).

The questions were presented in random order according to a dropout procedure: If a question was answered correctly, it was dropped from the list; if a question was answered incorrectly, it was presented again later, after all the other questions had been asked. A block of trials ended when each of the 48 questions had been answered correctly once. Subjects completed three blocks of trials in this manner; thus, each question was answered correctly three times by the end of the study phase of the first session. The remaining sessions involved an abbreviated study phase in which the initial presentation of the study list was omitted and the cued recall test involved only one block of trials, which served as an assessment of subjects' memory for the study list from one session to the next.

In the test phase, subjects performed a recognition task in blocks of trials involving the response signal procedure. Each trial began with a fixation cross presented in the center of the screen for 1000 ms, after which time the cross disappeared and a test probe of the form *Is the person in the location?* was presented in the center of the screen. The test probe was followed after one of eight lags (200, 500, 800, 1100, 1400, 1800, 2400, or 3000 ms) by a response signal consisting of an 800 Hz tone for 50 ms and offset of the test probe (resulting in a blank screen). Subjects were instructed to make a yes–no recognition response within 300 ms after the tone by pressing either the “/” key to respond “yes” or the “Z” key to respond “no” with their right or left index fingers, respectively. If they had determined their response before hearing the tone (which was likely at the longest lags), they were instructed to wait until the tone and then make the response immediately. If they had not determined their response by the time of the tone (which was likely at the shortest lags), they were instructed to immediately make whichever response they thought was most likely. It was emphasized that they should always try to respond within 300 ms after the tone.

After making their response, subjects received feedback about their trial performance for 1500 ms, followed by a 500-ms blank screen before the next trial commenced. There were three pieces of feedback. First, subjects were informed as to whether their response was correct or incorrect. Second, they were informed of their RT from response signal onset. Third, they received a message characterizing their performance. If they responded before the response signal, the message was *TOO EARLY*. If they responded longer than 300 ms after the response signal, the message was *TOO LATE*. If they responded within 300 ms after the response signal and correctly, the message was *BONUS*. If they responded within 300 ms after the response signal but incorrectly, there was no message. The *BONUS* message referred to a bonus system designed to motivate compliance with the response signal procedure and accurate performance. A correct response within 300 ms after the response signal earned one bonus point. Bonus points were accumulated over trials and converted to bonus pay at the end of the experiment (1 bonus point = 1 cent).

The test phase was divided into nine blocks, with 64 trials per block. At the start of each block, subjects were informed of the number of bonus points earned in the previous block and their cumulative number of bonus points earned in the session. The 64 trials in each block consisted of the 32 targets and 32 foils described earlier, presented in random order subject to the constraint that no person or location was repeated across consecutive trials. The eight response signal lags were randomly assigned to probes such that each lag occurred twice with each combination of probe (target or foil) and fan (1 or 2). Thus, each condition in the full 2 (probe) \times 2 (fan) \times 8 (lag) experimental design was represented twice per block,

giving a total of 18 observations per condition in each session and 90 observations per condition in the entire experiment. Excluding the first session and the first block of each subsequent session as practice, there were 64 observations per condition for each subject's experimental data.

Results and Discussion

Study phase

The mean frequency with which questions about Fan 1 and Fan 2 facts were asked during each block of the cued recall test in each session is provided in Table 1. Question frequency can be interpreted as a measure of the difficulty in learning the facts (with higher frequency indicating greater difficulty). The minimum possible frequency was equal to 1 because each question had to be answered correctly at least once per block. In the first session, subjects initially had more difficulty learning Fan 2 facts than Fan 1 facts, but the difference in frequency essentially disappeared by the second and third blocks, suggesting that all the facts had been memorized equally well by the end of the session. In subsequent sessions, question frequency was at or near the minimum of 1 and approximately equal for Fan 1 and Fan 2 facts, suggesting that subjects maintained good memory for the facts from one session to the next.

These observations are supported by the results of two repeated-measures analyses of variance (ANOVAs). First, a 3 (block) \times 2 (fan) ANOVA on question frequency for the first session revealed a significant main effect of block, $F(2,18) = 32.55$, $MSE = 0.42$, $p < .001$, $\eta_p^2 = .78$, a significant main effect of fan, $F(1,9) = 25.45$, $MSE = 0.21$, $p < .01$, $\eta_p^2 = .74$, and a significant interaction between block and fan, $F(2,18) = 15.29$, $MSE = 0.21$, $p < .001$, $\eta_p^2 = .63$, reflecting the decrease in the difference in question frequency between fan conditions across blocks. Second, a 5 (session) \times 2 (fan) ANOVA on question frequency for the third block of the first session and the blocks of all subsequent sessions revealed no significant effects, all $ps > .1$, reflecting the stable, near-minimum question frequency across sessions.

Test phase

The data from the first session and the first block of each subsequent session were excluded as practice. Trials with RTs shorter than 100 ms or longer than 350 ms were also excluded from all analyses (following Hintzman et al., 1998; Hintzman & Curran, 1994, 1997; Rotello & Heit, 2000) because the shorter RTs likely reflect anticipations and the longer RTs may reflect substantial post-lag task processing (indicating non-compliance with the instructions to respond immediately after the response signal). Only 5.7% of trials were excluded for not meeting the RT criteria, indicating that subjects were generally compliant with the demands of the response signal procedure. We present the group data in the figures below and we provide the individual-subject data in Appendix B.

Mean accuracy (proportion correct) is plotted as a function of lag + mean RT in Figure 3A for each combination of probe (target or foil) and fan (1 or 2). At the shortest lag, accuracy was near chance, but it increased rapidly with more processing time until it was at or near ceiling at the longest lag in all conditions. Accuracy was slightly higher for foils than for targets at the shortest lag, reflecting an initial bias to respond "no," but the difference disappeared at later lags. Accuracy was higher for Fan 1 items than for Fan 2 items, predominantly at the intermediate lags, although there were small numerical differences at the longest lags. These results represent evidence of a fan effect on the time course of associative recognition.

These observations are supported by the results of a 2 (probe) \times 2 (fan) \times 8 (lag) repeated-measures ANOVA on accuracy (proportion correct). There was a significant main effect of fan, $F(1,9) = 61.20$, $MSE = 0.004$, $p < .001$, $\eta_p^2 = .87$, a significant main effect of lag, $F(7,63) = 180.01$, $MSE = 0.007$, $p < .001$, $\eta_p^2 = .95$, and a significant interaction between fan and lag, $F(7,63) = 7.28$, $MSE = 0.003$, $p < .001$, $\eta_p^2 = .45$. There were no significant effects involving probe (target versus foil), all $ps > .2$.

To illustrate the fan effect in an alternative way, mean accuracy is plotted in terms of d' , as a function of lag + mean RT, as points in Figure 4 for each fan condition (d' was calculated separately for each individual subject, then averaged across subjects). To permit calculation of d' when accuracy was perfect, the raw accuracy values (proportion correct) were adjusted by adding 0.5 to the number of “yes” responses and dividing by the total number of responses + 1.0 (Hintzman & Curran, 1994, 1997; Rotello & Heit, 2000; Snodgrass & Corwin, 1988). Figure 4 shows there were typical speed–accuracy tradeoff functions in both fan conditions, with what appears to be a higher rate of rise to asymptote for Fan 1 items than for Fan 2 items. There was also a difference in d' at the longest lag for Fan 1 and 2 items, but this difference should be interpreted with caution because d' still appears to be rising at the longest lag in the Fan 2 condition, suggesting that the asymptote had not yet been reached. Moreover, the model fits we present later do not provide evidence in favor of different asymptotes.

Mean RT is plotted as a function of lag in Figure 3B for each combination of probe (target or foil) and fan (1 or 2). The most prominent effect was that RT became shorter as the lag became longer, which is typical of response signal data. However, there were also differences in RT between the fan conditions that are not particularly evident in Figure 3B. To make the fan differences clearer, Figure 5A shows the data collapsed over targets and foils, and Figure 5B shows the fan effect (the difference in RT between Fan 2 and Fan 1 items) as a function of lag. There was a small but statistically significant fan effect of 3 ms and it varied across lags, being almost non-existent at the shortest lag, emerging at the second-shortest lag, and then gradually decreasing back to zero by the longest lags (see Figure 5B). Thus, our experiment revealed a fan effect not only on accuracy, but also on the lag function for RT.

These observations are supported by the results of a 2 (probe) \times 2 (fan) \times 8 (lag) repeated-measures ANOVA on mean RT. There was a significant main effect of fan, $F(1,9) = 66.18$, $MSE = 9.9$, $p < .001$, $\eta_p^2 = .88$, reflecting the 3-ms fan effect. Given that RTs in the response signal procedure are very short and have low variance, it is not unusual for small differences between conditions to be statistically significant (e.g., see Hintzman et al., 1998). There was also a significant main effect of lag, $F(7,63) = 172.65$, $MSE = 86.2$, $p < .001$, $\eta_p^2 = .95$, and a significant interaction between fan and lag, $F(7,63) = 5.50$, $MSE = 19.2$, $p < .001$, $\eta_p^2 = .38$, reflecting the modulation of the fan effect across lags. The only significant effect involving probe (target versus foil) was an interaction between probe and lag, $F(7,63) = 4.36$, $MSE = 19.5$, $p < .01$, $\eta_p^2 = .33$.

Summary

The results of our experiment complement those of Wickelgren and Corbett (1977) and Doshier (1981) by revealing fan effects on the time course of associative recognition, but with materials that were well-learned instead of briefly studied. The data from the study phase support the conclusion that all the studied facts were represented strongly in memory by the end of the cued recall test in the first session and in subsequent sessions (see Table 1). The data from the test phase indicate that our materials and procedure yielded typical speed–accuracy tradeoff functions, with fan primarily affecting the rate of rise to asymptotic accuracy (see Figures 3A and 4). In addition, we found that RT became shorter as the lag

become longer (see Figure 3B) and there was a small fan effect on RT that varied across lags (see Figure 5). Collectively, the results constitute an important challenge for the ACT-R model of the fan effect, which has not addressed response signal data. In the next section, we present an extension of the ACT-R model that overcomes this limitation and we assess how well it accounts for speed–accuracy tradeoff functions compared with variants of the SEF.

An ACT-R Model for the Response Signal Procedure

Model Description

We extended the standard ACT-R model of the fan effect in a straightforward manner to account for data from the response signal procedure. Memory retrieval is still a central process in the model, but other processes involved in task performance need to be included to generate a time course for associative recognition. The four kinds of processes in the model—encoding, retrieval, guessing, and responding—are summarized in Table 2 and described below in the context of the fact retrieval paradigm. To aid in the understanding of when each process is active during a trial, Figures 6A and 6B illustrate the organization of the model's processing stages at long and short lags, respectively.

Encoding—There are two things that need to be encoded on every trial: the probe stimulus and the response signal. Stimulus encoding begins at stimulus onset (see Figure 6) and results in a representation of the probe. The time to encode the stimulus, t_{stim} , is a free parameter. Response signal encoding begins at response signal onset (see Figure 6) and results in a representation of the signal. Given that the response signal is typically a simple tone, we set the time to encode the response signal, t_{signal} , equal to 50 ms, which is the default time for simple tone detection in ACT-R (based on work by Meyer & Kieras, 1997).

Retrieval—Memory retrieval involves using the encoded stimulus to probe declarative memory and retrieve a fact. Targets and foils retrieve matching and nonmatching studied facts, respectively, with a match resulting in selection of a “yes” response and a nonmatch resulting in selection of a “no” response. For simplicity, we assume that retrieval is an error-free process (e.g., a target will never retrieve an alternative studied fact), although there are variants of the retrieval process in ACT-R that allow errors to arise from partial matching (Anderson & Lebiere, 1998). Consequently, if retrieval has time to finish, then the model will always produce a correct response.

Retrieval begins as soon as stimulus encoding has finished (see Figure 6) and retrieval time, $t_{retrieve}$, is determined by Equation 3, which we repeat here:

$$t_{retrieve} = F e^{-A}. \quad (3)$$

Recall that the activation of fact i (A_i) in memory is determined by its base-level activation (B_i) and associative activation (see Equation 1). Given that facts are typically tested equally often and practice effects are generally not a concern, we assume equal base-level activation for all facts and set $B_i = 0$. As described earlier, associative activation varies as a function of the total source activation (W), the maximum associative strength (S), and the fan from the person and location sources used as retrieval cues (see Equations 1 and 2b). We set $W = 1.0$, which is the default value in ACT-R (Anderson et al., 1996), and we set $S = 1.5$, which is an arbitrary value that we have used previously (Schneider & Anderson, 2011). We fixed S because it trades off with F , which is a free parameter that scales retrieval time in Equation 3.

A critical feature of the response signal procedure is that there is limited time available for retrieval (determined by the response signal lag). If there were a constant retrieval time in the model, then its resulting time-course function would be a step function: Retrieval would never have time to finish at very short lags, resulting in low accuracy, and it would always finish at very long lags, resulting in high accuracy. An abrupt shift from low to high accuracy would occur at some intermediate lag where retrieval time equaled the time available for retrieval. However, it is clear from response signal data that empirical time-course functions are not step functions (e.g., see Figures 2 and 4).²

For the model to produce a more appropriate time-course function, there has to be some variability in retrieval time. Variability can be introduced by adding noise to the activation of facts in memory. The default in ACT-R is for activation noise to be distributed logistically (Anderson, 2007) with a mean equal to 0 and variability controlled by a free parameter s . In this situation, retrieval times will follow a log-logistic distribution (Schneider & Anderson, 2011). Having a distribution of retrieval times means that, for a given response signal lag, there will be some probability that retrieval has finished. To determine this probability, one must first determine the time available for retrieval. It is not simply equal to the lag because some time is needed to encode the probe stimulus and the response signal. As mentioned earlier, we assume that retrieval begins when stimulus encoding has finished (see Figure 6), so stimulus encoding time must be subtracted from the lag. We also assume that retrieval can go on in parallel with response signal encoding (tone detection) because the two processes involve different modules in ACT-R (the declarative memory module and the auditory module, respectively). Consequently, response signal encoding time can be added to the lag. The total time available for retrieval, t_{avail} , can be expressed as (see Figure 6):

$$t_{avail} = lag + t_{signal} - t_{stim}. \quad (5)$$

Knowing the time available for retrieval and that retrieval time follows a log-logistic distribution, one can determine the probability that retrieval has finished in the time available, $p_{retrieve}$, from the cumulative distribution function (CDF) of the log-logistic distribution:

$$p_{retrieve} = \frac{1}{1 + (t_{retrieve}/t_{avail})^{1/s}}, \text{ if } t_{avail} > 0, \text{ else } 0. \quad (6)$$

Examples of the log-logistic CDF (with $t_{retrieve} = 500$ ms and $s = 0.2$ or 0.4) are shown in Figure 7. Comparing the two solid curves, one can see how the retrieval function becomes closer to a step function as s decreases (i.e., as the noise becomes less variable). The crossing point of the two curves—which is where the step would occur for a step function—corresponds to the median of the retrieval time distribution, which is equal to $t_{retrieve}$ (Equation 3). The two sets of broken lines show how the probability of retrieval changes as a function of the time available for retrieval in the context of a specific retrieval time distribution (indicated by the solid black curve). When $t_{avail} = 700$ ms (dashed line), $p_{retrieve} = .70$, meaning that retrieval would not finish in the time available on 30% of trials (in which case the model resorts to guessing, as described below). Assuming that stimulus encoding time and response signal encoding time do not vary with lag, then lengthening the

²The curved time-course functions for the group data in Figures 2 and 4 are not artifacts from averaging individual-subject data. No subject in Wickelgren and Corbett's (1977) experiment or in our experiment had a data pattern that resembled a step function (see Appendix B for individual-subject data).

lag by 300 ms results in $t_{avail} = 1000$ ms (dotted line) and $p_{retrieve} = .85$, showing that retrieval is more likely to finish at longer lags. As noted earlier, we assume that retrieval is an error-free process; therefore, a higher probability of retrieval results in a higher level of accuracy (an equation for the model's accuracy that involves $p_{retrieve}$ is presented below). Moreover, retrieval is a discrete, all-or-none process because it either finishes or does not finish in the time available. There is no partial information from an unfinished retrieval process to guide the model's selection of a response. We discuss the issue of retrieval being a discrete (as opposed to a continuous) process and the absence of partial information further when we compare our model with other models of response signal data in the General Discussion.

The preceding discussion is based on the assumption that the lag is used to generate an external deadline (t_{avail}) for the retrieval process (Equation 5). As a result, given a long enough lag, retrieval will always finish and accuracy will be perfect. This might be reasonable for a situation in which facts are well-learned (e.g., our experiment) and subjects know that a fact can ultimately be retrieved if there is enough time to do so. However, it is less reasonable for a situation in which facts are briefly studied (e.g., Wickelgren & Corbett, 1977) and, as a result, have low activation in memory and may take a long time to retrieve. If subjects have some perception of the degree to which facts are active in memory, then they may set an internal deadline to limit the time spent on futile attempts to retrieve facts with low activation that might ultimately never be retrieved. If an internal deadline is set for the retrieval process, then retrieval fails if it is taking too long. A standard aspect of the retrieval process in ACT-R is that retrieval fails if it takes longer than an internal deadline, $t_{failure}$, given by:

$$t_{failure} = F e^{-\tau}, \quad (7)$$

which is a variant of Equation 3 in which activation (A) is replaced by an activation threshold (τ). To determine the probability that retrieval finishes before the failure time, one simply replaces t_{avail} with $t_{failure}$ in Equation 6. The two sets of broken lines in Figure 7 can then be reinterpreted as illustrating the effect of varying the threshold, with a high threshold (dashed line) resulting in a shorter failure time and a lower probability of retrieval than a low threshold (dotted line).

If the internal and external deadlines represented by $t_{failure}$ and t_{avail} , respectively, are both present during a trial, then the probability of retrieval is determined by whichever deadline is shorter. For example, if the dashed line in Figure 7 is associated with $t_{failure}$ and the dotted line is associated with t_{avail} (i.e., $t_{failure} < t_{avail}$), then $p_{retrieve}$ is determined by $t_{failure}$ and equals .70. Similar logic applies when $t_{avail} < t_{failure}$. Critically, when $t_{failure}$ falls within the range of available times bounded by the shortest and longest lags (i.e., it is between the shortest and the longest values of t_{avail}), it determines the asymptotic level of accuracy achieved by the model. Continuing with the example of the dashed line in Figure 7 being associated with $t_{failure}$, the mean accuracy of the retrieval process would never exceed .70—even at the longest lags—because failure time is not affected by lag (compare Equation 7 with Equation 5). For modeling asymptotic levels of accuracy that are below ceiling, τ is a free parameter. For modeling asymptotic levels of accuracy that are at or near ceiling, we show in our model fits below that one can justify setting $\tau = -\infty$, which is similar to having no internal deadline or, more precisely, an infinitely long internal deadline that would never be reached, which is functionally equivalent to having no internal deadline. More generally, one could set τ equal to any value that makes the internal deadline exceed the longest external deadline. In this situation, accuracy will be perfect whenever t_{avail} is long enough for retrieval to finish.

At very short lags there may not be any time available for retrieval (i.e., $t_{avail} \leq 0$) because stimulus encoding has not finished. For example, the shortest lag in Wickelgren and Corbett's (1977) experiment and in our experiment was 200 ms, which is most likely insufficient for stimulus encoding. Indeed, in a fan experiment involving associative recognition of person–location facts in which eye movements were monitored, Anderson, Bothell, and Douglass (2004) found that mean first-gaze duration for the probe stimulus was 400 ms, suggesting that it may have taken that long to encode the stimulus. In our model, if stimulus encoding has not finished then retrieval is not initiated; instead, the model resorts to guessing. This is true even in the case where there may be sufficient time to encode one part of the stimulus (e.g., the person) but not the other (e.g., the location). Anderson et al. found no evidence that retrieval began with the first gaze, suggesting that both parts of the stimulus have to be encoded before retrieval is initiated. This seems sensible because even if retrieval were initiated by part of the stimulus (e.g., the person) and was able to finish in time, it would not be possible to respond accurately because the retrieved fact (e.g., a person–location pair) would match the encoded part of the stimulus regardless of whether the stimulus was a target or a rearranged foil (i.e., item recognition alone is insufficient). Accurate responding is possible only when the retrieved fact is matched against the entire stimulus, and for that to happen, both parts of the stimulus have to be encoded.

Guessing—If the stimulus has been encoded and retrieval has enough time to finish (i.e., $t_{retrieve}$ is shorter than the shortest deadline represented by $t_{failure}$ or t_{avail}), then the model waits until the response signal is encoded before it executes a response based on the match or mismatch between the retrieved fact and the probe (see Figure 6A). If retrieval is not initiated because the stimulus has not been encoded or it is initiated but does not finish because it either fails (i.e., it reaches the internal deadline represented by $t_{failure}$) or there is not enough time available (i.e., it reaches the external deadline represented by t_{avail}), then the response is determined by guessing. Guessing starts after the response signal has been encoded and occurs only if retrieval has not finished or been initiated (see Figure 6B). It does not occur in parallel with retrieval (cf. Meyer, Irwin, Osman, & Kounios, 1988) and the response is not determined by retrieval even if retrieval happens to finish while guessing is in progress. The contingency of guessing on unfinished retrieval is based on the assumption that subjects do not guess unless the need arises.

We assume a simple guessing process whereby the model guesses “yes” with a probability equal to $bias$ and “no” with a probability equal to $1 - bias$. For predicting d' , which controls for response bias, as a simplifying assumption we set $bias = 0.5$, although it would be a free parameter when predicting other accuracy measures such as proportion correct for targets and for foils. Note that the $bias$ parameter is fixed across lags and the guessing process does not change as a function of lag because an unfinished retrieval process provides no partial information. Given that retrieval is error-free, guessing is the sole source of errors in the model. For simplicity, we set the time to take a guess, t_{guess} , equal to 50 ms, which is the default time for firing a single production (condition–action rule) in ACT-R (Anderson, 2007). Thus, our model implements a guessing process (cf. Ollman, 1966; Yellott, 1967, 1971) that is triggered by the response signal and contingent on unfinished retrieval.

Responding—Once a response has been determined either by retrieval or by guessing, it is executed (see Figure 6). For simplicity, we assume that response execution is an error-free process. The time to execute a response, t_{resp} , is a free parameter.

Predictions—To understand how the model predicts speed–accuracy tradeoff functions, consider what happens at long and short lags. At a long lag (see Figure 6A), stimulus encoding and memory retrieval both have time to finish before the lag elapses. Consequently, the model has determined the correct response during the lag and simply

waits until the response signal is encoded, then the response is executed. Accuracy is high because retrieval is an error-free process. At a short lag (see Figure 6B), there may be insufficient time available to finish stimulus encoding and memory retrieval. For the example in Figure 6B, retrieval is only about halfway done by the time the response signal is encoded. Given that retrieval is unfinished, the model resorts to guessing and then executes the guessed response. Accuracy is low because guessing is an error-prone process.

The accuracy (proportion correct) predicted by the model is determined by the probability that retrieval finished (Equation 6) and the probability of a correct guess, p_{guess} , which is controlled by the *bias* parameter (i.e., $p_{guess} = bias$ for targets and $p_{guess} = 1 - bias$ for foils):

$$accuracy = p_{retrieve} (1) + (1 - p_{retrieve}) p_{guess}. \quad (8)$$

In other words, overall accuracy is a mixture of perfect accuracy based on retrieval (reflected in the multiplication of $p_{retrieve}$ by 1) and imperfect accuracy based on guessing ($p_{guess} < 1$). Given that the probability of retrieval increases as the lag becomes longer, accuracy increases across lags. The shape of the speed–accuracy tradeoff function produced by the model is based on the log-logistic CDF for the probability of retrieval (Equation 6).

Besides accuracy, the model also makes predictions for RT. When retrieval has time to finish, RT is the sum of response signal encoding time and response execution time (see Figure 6A). When retrieval does not finish, RT is the sum of response signal encoding time, response execution time, and guessing time (see Figure 6B). The mean RT predicted by the model is a mixture of times that do and do not include guessing time, as determined by the probability that retrieval finished (Equation 6):

$$RT = p_{retrieve} (t_{signal} + t_{respx}) + (1 - p_{retrieve}) (t_{signal} + t_{respx} + t_{guess}). \quad (9)$$

Given that the probability of retrieval increases as the lag becomes longer, RT gets shorter across lags due to the reduced contribution of guessing time. As noted earlier, RT typically becomes shorter across lags in the response signal procedure (e.g., see Figure 3B). It follows from Equation 9 that the difference in RT between the shortest and the longest lags will be approximately equal to the guessing time (assuming retrieval never finishes at the shortest lag and always finishes at the longest lag), a point we consider further in the General Discussion.

Summary—Our ACT-R model includes all the processes taking place in the time course of associative recognition in the response signal procedure: encoding, retrieval, guessing, and responding (see Figure 6 and Table 2). Stimulus encoding results in a representation of the probe that is used to access memory. Retrieval is an all-or-none, error-free process that involves retrieving a fact from memory for comparison with the probe. The probability of retrieval is determined by the distribution of retrieval times (which is affected by the amount of associative activation and the level of activation noise) and the time available for retrieval (which is affected by lag and the activation threshold). When retrieval does not finish, the model resorts to guessing, which is an error-prone process subject to response bias. Responding involves executing the response determined by retrieval or by guessing. The model is completely specified by Equations 1–3 and 5–9, allowing for precise quantitative predictions. Its predictions for accuracy and for RT are strongly determined by the probability of retrieval, which increases as the lag becomes longer, resulting in higher accuracy and shorter RT across lags.

Modeling Details

We evaluated our ACT-R model of the response signal procedure by fitting it to the individual-subject data from Wickelgren and Corbett (1977) and from our experiment (see Figures 2 and 4), then comparing its fits to those of different SEF variants.³ Before presenting the modeling results, we summarize the ACT-R model and SEF parameters, describe the fitting process, explain how models were compared, and discuss simulations of the ACT-R model.

Model parameters—For fitting the ACT-R model to d' accuracy data, there are only four free parameters: t_{stim} , F , s , and τ . The only constraints on free parameter values were that t_{stim} and F had to be greater than 0 because time cannot be negative. In the modeling results presented below, we also consider a variant of the model in which we set $\tau = -\infty$, which is equivalent to having no internal deadline for the retrieval process. We set the other parameters in the model equal to the fixed values mentioned earlier (see Table 2). Thus, there was a maximum of four free parameters for fitting the 16 d' accuracy data points in each experiment (2 fan conditions \times 8 lags; see Figures 2 and 4). Note that there are no free parameters that distinguish between fan conditions or vary as a function of lag. Thus, fan effects on accuracy and changes in accuracy across lags are predictions of the model that reflect its basic structure and functioning.

For fitting the ACT-R model to RT data, we set all encoding, retrieval, and guessing parameters equal to their fixed values or to their best-fitting values from the fit to accuracy data. This leaves only one free parameter, t_{resp} , which functions merely as an intercept shift for overall RT when fitting the 16 mean RT data points in our experiment (see Figure 5). Once again, note that there no free parameters that distinguish between fan conditions or vary as a function of lag. Thus, fan effects on RT and changes in RT across lags are predictions of the model that reflect its basic structure and functioning. The model was fit sequentially to d' accuracy and RT data for our experiment,⁴ but only to d' accuracy for Wickelgren and Corbett's (1977) experiment because they did not report individual-subject RT data.

For fitting the SEF to d' accuracy data, there are between three and six free parameters depending on the SEF variant. As noted earlier, if there are two conditions, as is the case with the fan manipulation of interest here (i.e., Fan 1 versus Fan 2), then there are eight possible SEF variants based on whether each parameter (intercept, rate, or asymptote) is the same or different across conditions. The model with the fewest free parameters assumes no difference between fan conditions for any parameter and the model with the most free parameters assumes a difference between fan conditions for every parameter, with the other models falling in between these extremes. There were no constraints on free parameter values.

Model fitting—For fitting the ACT-R model and the SEF to d' accuracy data, we followed an approach involving maximum likelihood estimation (Liu & Smith, 2009; for overviews, see Lewandowsky & Farrell, 2011, Ch. 4; Myung, 2003), using a likelihood function based on the probability density function for the normal distribution and minimizing the negative log-likelihood value.⁵ For fitting the ACT-R model to mean RT data, we performed least-

³We did not fit any data from Doshier (1981) because she did not report individual-subject data; however, her group data were similar to the data of Wickelgren and Corbett (1977).

⁴We followed a sequential model-fitting approach because we were primarily interested in the fit of the ACT-R model to accuracy data and how that compared with fits of SEF variants. Given that the SEF only addresses accuracy data, it seemed fair to use the fit of the ACT-R model only to accuracy data for comparison purposes, then to evaluate the fit of the ACT-R model to RT data (constrained by its earlier fit to accuracy data).

squares estimation, minimizing the root mean squared deviation (RMSD) between data and model predictions. The ACT-R model and the SEF were both implemented in MATLAB (The MathWorks, Inc., Natick, MA) and we used an implementation of the simplex algorithm (Nelder & Mead, 1965) for parameter optimization. We used multiple sets of starting parameter values covering a large region of the parameter space to avoid local minima. All the modeling results reported below are based on fits to individual-subject data, although we also include the results of fits to group data. We report individual-subject parameter values and fit indices in tables, but to simplify the visual presentation of the results, the means of the individual-subject predictions are shown in figures. Individual-subject modeling results are provided in Appendix B.

Model comparison—We conducted model comparisons to determine the best SEF variant and to assess the adequacy of the ACT-R model. Model comparisons for fits to response signal data often involve computing an adjusted R^2 statistic for each model that takes into account the number of free parameters:

$$R^2 = 1 - \frac{\sum_{i=1}^n (d_i - \widehat{d}_i)^2 / (n - k)}{\sum_{i=1}^n (d_i - \bar{d})^2 / (n - 1)}, \quad (10)$$

where d_i is data point i , \widehat{d}_i is the model's prediction for data point i , \bar{d} is the mean of the data points, n is the number of data points, and k is the number of free parameters in the model. For models with identical fits, Equation 10 will yield a higher R^2 value for the model with fewer free parameters. Although we report R^2 values, we focus on a more principled approach to model comparison involving information criterion measures that quantify goodness of fit while penalizing model complexity (Lewandowsky & Farrell, 2011; Liu & Smith, 2009; Wagenmakers & Farrell, 2004). The two measures we use are the Akaike information criterion (AIC), given by:

$$AIC = -2 \ln L + 2k + \frac{2k(k+1)}{(n-k-1)}, \quad (11)$$

and the Bayesian information criterion (BIC), given by:

$$BIC = -2 \ln L + k \ln n, \quad (12)$$

where L is the likelihood value, k is the number of the free parameters, and n is the number of data points. Equation 11 is a variant of the AIC that includes a correction for small sample size (the last term in the equation), which is recommended when the ratio of data points to free parameters (n/k) is less than 40 (Wagenmakers & Farrell, 2004). The AIC and the BIC are commonly used for model selection and comparison purposes, allowing one to choose the simplest model that provides an adequate fit to the data (the model with the lowest AIC or BIC value is to be preferred). In the present context, model selection involves determining which SEF variant is the best descriptive model of the data and model

⁵All ACT-R model fits reported in this article were repeated by fitting raw accuracy data (proportion correct) with a likelihood function based on the probability mass function for the binomial distribution (Lewandowsky & Farrell, 2011; see also Usher & McClelland, 2001). These fits were to all 32 cells in the 2 (probe) \times 2 (fan) \times 8 (lag) experimental design and involved making separate predictions for targets and for foils, as well as allowing *bias* to be a free parameter. The results of these fits are reported in Appendix B.

comparison involves determining whether the ACT-R model fits as well as the best SEF variant.

To facilitate model selection and comparison, we consider two measures based on the AIC and the BIC. The first measure is the rank of a model's fit to data. One can assign a rank from 1–9 to the ACT-R model and the eight SEF variants based on the AIC and the BIC values for their fits to each individual-subject data set, with lower ranks corresponding to lower AIC and BIC values (i.e., better fits). One can then compute the mean rank of a given model across subjects to get a sense of how well it fares when compared with the other models. The second measure is the model weight, which is the probability that a given model is the best model in the set being compared (Liu & Smith, 2009; Wagenmakers, 2007; Wagenmakers & Farrell, 2004). The weight for model i , w_i , is calculated as:

$$w_i = \frac{\exp[-0.5\Delta_i(IC)]}{\sum_{j=1}^J \exp[-0.5\Delta_j(IC)]}, \quad (13)$$

where IC is the value of either the AIC or the BIC, and $\Delta_i(IC)$ is the difference between the IC value of model i and the lowest IC value from the set of J models. One can compute the mean weight of a model across subjects to get sense of its overall probability of being the best model.

Finally, one can directly compare the fits of two models by calculating the likelihood ratio in favor of one model over the other, which is simply the ratio of their model weights (Glover & Dixon, 2004; Liu & Smith, 2009; Wagenmakers & Farrell, 2004). We use likelihood ratios to compare the fits of the ACT-R model with and without a threshold as a way of determining whether an internal deadline is necessary during retrieval. To provide a summary of the likelihood ratio across subjects, we report the group likelihood ratio, which is the product of the individual-subject likelihood ratios, and the average likelihood ratio, which is the geometric mean of the individual-subject likelihood ratios (Liu & Smith, 2009).

Simulation—The ACT-R model we fit to the data is completely specified by Equations 1–3 and 5–9. Although this analytic version of the model has the practical advantage of making model fitting fast and accurate, it has the disadvantage of underrepresenting the amount of variability in processing. More specifically, the model has just one source of variability—the distribution of retrieval times—but there is undoubtedly variability in the times associated with other processes. To determine the extent to which these sources of variability may be important, the analytic model fits are accompanied by simulation results for comparison.

Simulation of the model involved six sources of variability: retrieval time, stimulus encoding time, response signal encoding time, guessing time, guessing accuracy, and response execution time. The variability in retrieval time (and, by extension, the probability of retrieval) was simulated by sampling from a logistic distribution for activation noise and adding it to the activation of the fact retrieved on each trial (i.e., Equation 6 was not computed in the simulation). The variability in the times for stimulus encoding, response signal encoding, guessing, and response execution was simulated by sampling from a uniform distribution specific to each process, with the boundaries of the distribution being $t \pm t/2$, where t is the mean time for the process (see Anderson, Taatgen, & Byrne, 2005; Meyer & Kieras, 1997). The variability in guessing accuracy was simulated by sampling a random number between 0 and 1 from a uniform distribution, then determining whether it was above or below the *bias* parameter.

The model was simulated for each subject using that subject's best-fitting parameter values from the analytic fits. A total of 3.2 million trials were simulated per subject, representing 100,000 trials for each of the 32 cells in the 2 (probe) \times 2 (fan) \times 8 (lag) experimental design. The simulation fit indices are reported in parentheses beside the analytic fit indices in the tables of modeling results (see also Appendix B).

Modeling Results

Wickelgren and Corbett (1977) —The means of the individual-subject predictions for d' accuracy from the best-fitting ACT-R model (with a threshold) appear as solid lines in Figure 2. It is evident that the ACT-R model provides an excellent fit to the data, not only capturing the basic shape of each speed–accuracy tradeoff function, but also reproducing the large difference in asymptotic accuracy at the longest lags and the modest difference in the rate of rise to the asymptote across fan conditions. These fan effects were produced by the model because Fan 2 items are retrieved more slowly than are Fan 1 items due to less associative activation for the former than for the latter (see Equations 1–3). The rate difference predicted by the model is a direct reflection of the difference in retrieval time. The asymptote difference predicted by the model reflects a combined effect of the difference in retrieval time and the threshold. Recall that the threshold determines the failure time (Equation 7), which is an internal deadline for retrieval to finish. For a specific threshold, the probability that retrieval finishes before the internal deadline is lower for Fan 2 items than it is for Fan 1 items because retrieval times are longer for the former than for the latter. Given that retrieval accuracy is directly related to the probability of retrieval (Equation 8), it follows that asymptotic accuracy will be lower for Fan 2 items than for Fan 1 items.

The best-fitting parameter values (for F , s , t_{stim} , and τ) and fit indices (AIC, BIC, and R^2) for individual subjects are presented in Table 3. All the parameter values seem reasonable, with the sole exception of a very short stimulus encoding time of 32 ms for subject CM.⁶ The AIC and the BIC values are not particularly interpretable in isolation, so we discuss them only in the context of model comparisons below. The ACT-R model yielded good fits in terms of R^2 , with a mean $R^2 = .921$ for the individual-subject fits and $R^2 = .981$ for the group fit. The fit indices from simulations of the model (see Table 3) are similar to the analytic fit indices, indicating that the lack of extra variability in the analytic version of the model was not detrimental to its fits.

The predictions in Figure 2 are for the ACT-R model with a threshold, which fit better than did the model without a threshold (i.e., with $\tau = -\infty$). Individual-subject likelihood ratios favored the threshold model over the no-threshold model for all six subjects. The group likelihood ratio, expressed in terms of the threshold model over the no-threshold model, equaled 3.1×10^{27} for AIC and 4.1×10^{28} for BIC, and the average likelihood ratio equaled 3.8×10^4 for AIC and 5.9×10^4 for BIC. These results are not surprising because large empirical differences in asymptotic accuracy can only be captured by using an internal deadline set by a threshold. The presence of an internal deadline seems reasonable when facts are briefly studied and may never be retrieved due to their low activation, even when there is a long time available for retrieval.

The means of the individual-subject predictions for d' accuracy from the best SEF variant appear as dashed lines in Figure 2. The best variant across all subjects (based on the rank and model weight statistics presented below) was the one in which the asymptote and the intercept (but not the rate) differed between fan conditions, consistent with what Wickelgren

⁶If stimulus encoding time is constrained to be ≥ 100 ms, then there is a modest decrement in the quality of the fit to the data from subject CM ($AIC = 17.5$, $BIC = 16.9$, $R^2 = .947$; cf. Table 3).

and Corbett (1977) found with their fits. The mean $R^2 = .935$ for this variant and its mean parameter values were $\delta_1 = 455$ ms, $\delta_2 = 498$ ms, $1/\beta = 312$ ms, $\lambda_1 = 4.01$, and $\lambda_2 = 2.72$, where the subscripts 1 and 2 denote the Fan 1 and Fan 2 conditions, respectively. However, a similar fit was obtained with a variant similar to the ACT-R model, where the asymptote and the rate (but not the intercept) differed between fan conditions. Table 4 shows the best-fitting parameter values and fit indices for the best variant (based on BIC values) for each individual subject (which did not necessarily correspond to the best variant over all subjects). For four of the six subjects and for the group, the best variant involved different asymptotes between fan conditions, indicating a fan effect on memory strength. For all six subjects, the best variant involved either different intercepts or different rates (or both) between fan conditions, indicating a fan effect on retrieval dynamics. Interestingly, the best variant for the fit to the group data involved only different asymptotes (see Table 4), providing an example of a group fit not corresponding exactly with the individual-subject fits (Cohen, Sanborn, & Shiffrin, 2008) and supporting our decision to focus on modeling individual-subject data.

We assessed the adequacy of the ACT-R model (with a threshold) by comparing its fit with those of the eight SEF variants. The mean rank statistics appear in Figure 8A and the mean model weights appear in Figure 8B. In terms of rank, the ACT-R model compares favorably with the two SEF variants mentioned above (involving different asymptotes and either different intercepts or different rates). The ACT-R model has the lowest mean rank based on AIC, being ranked in the top 3 for four of the six subjects, and it has the third-lowest mean rank based on BIC, being ranked in the top 3 for two of the six subjects. In terms of model weight, the results are similar. The ACT-R model has the highest mean weight based on AIC and the third-highest mean weight based on BIC. These model comparison results indicate that the ACT-R model produced quantitative fits that were as good as those of the best SEF variants.

Our experiment—The means of the individual-subject predictions for d' accuracy from the best-fitting ACT-R model (without a threshold) appear as solid lines in Figure 4. The ACT-R model provides an excellent fit to the data, reproducing the very high accuracy at the longest lags and the difference in the rate of rise in accuracy between fan conditions. The very high accuracy was predicted because no threshold was used for these fits (justification is provided below), meaning that there was no internal deadline that prevented retrieval from finishing at long lags. Consequently, the probability that retrieval finished was near 1.0 at the long lags, resulting in accuracy that was at or near ceiling (with very small differences in proportion correct at long lags being magnified when expressed in terms of d') and no difference in asymptote between fan conditions. The fan effect was produced by the ACT-R model for the reasons discussed earlier.

The best-fitting parameter values (for F , s , and t_{stim}) and fit indices (AIC, BIC, and R^2) for individual subjects are presented in Table 5. All the parameter values seem reasonable.⁷ The ACT-R model yielded good fits in terms of R^2 , with a mean $R^2 = .939$ for the individual-subject fits and $R^2 = .992$ for the group fit. Once again, the fit indices from simulations of the model (provided in parentheses in Table 5) are similar to the analytic fit indices.

⁷Direct comparisons between the parameter values for the model fits to our experiment data and to Wickelgren and Corbett's (1977) data are complicated by the many methodological differences between the experiments. For example, the largest difference in parameter values is for stimulus encoding time (t_{stim}), with means of 178 and 386 ms for the fits to Wickelgren and Corbett's data and to our data, respectively. We suspect this difference in parameter values may reflect a difference in the complexity of the test probes. In Wickelgren and Corbett's experiment the probe was simply a pair of words, whereas in our experiment it was a question of the form *Is the person in the location?* It likely took subjects longer to read the probe in our experiment than in Wickelgren and Corbett's experiment, which would be reflected in a longer stimulus encoding time in the model.

The predictions in Figure 4 are for the ACT-R model without a threshold, which fit better than did the model with a threshold. Individual-subject likelihood ratios favored the no-threshold model over the threshold model for most of the subjects (eight subjects based on AIC; seven subjects based on BIC). The group likelihood ratio, expressed in terms of the no-threshold model over the threshold model, equaled 8180.9 for AIC and 108.9 for BIC, and the average likelihood ratio equaled 2.5 for AIC and 1.6 for BIC. These results are not surprising because the very high empirical accuracy at long lags suggests that retrieval was not subject to an internal deadline (or it was subject to an internal deadline that exceeded the longest external deadline). The lack of a role of an internal deadline seems reasonable when facts are well-learned and can ultimately be retrieved if there is sufficient time available for retrieval.

The means of the individual-subject predictions for d' accuracy from the best SEF variant appear as dashed lines in Figure 4. The best variant across all subjects (based on the rank and model weight statistics presented below) was the one in which only the rate (and neither the asymptote nor the intercept) differed between fan conditions, consistent with the ACT-R model. The mean $R^2 = .956$ for this variant and its mean parameter values were $\delta = 675$ ms, $1/\beta_1 = 387$ ms, $1/\beta_2 = 917$ ms, and $\lambda = 4.48$, where the subscripts 1 and 2 denote the Fan 1 and Fan 2 conditions, respectively. Table 6 shows the best-fitting parameter values and fit indices for the best variant (based on BIC values) for each individual subject (which did not necessarily correspond to the best variant over all subjects). For all ten subjects, the best variant involved either different intercepts or different rates (or both) between fan conditions, indicating a fan effect on retrieval dynamics. For six of the subjects and for the group, the best variant involved different rates rather than different intercepts. For none of the ten subjects did the best variant involve different asymptotes between fan conditions. This result may seem paradoxical given that fan is assumed to affect memory strength, which is reflected in the strength of association between items in memory in ACT-R and thought to be reflected in the asymptote of the speed-accuracy tradeoff function (as measured by the SEF). We contend that fan does affect memory strength, but there will be no fan effect on the asymptote when low and high fan items can both be retrieved accurately from memory, given sufficient time for retrieval. Such a situation exists when the materials are well-learned, as they were in our experiment, and subjects do not set an internal deadline for retrieval, as suggested by the results for the no-threshold ACT-R model.

We assessed the adequacy of the ACT-R model (without a threshold) by comparing its fit with those of the eight SEF variants. The mean rank statistics appear in Figure 9A and the mean model weights appear in Figure 9B. In terms of rank, the ACT-R model compares favorably with the best SEF variant mentioned above (involving different rates). The ACT-R model is tied with that SEF variant for the lowest mean rank based on AIC, being ranked in the top 2 for seven of the ten subjects, and it has the second-lowest mean rank based on BIC, being ranked in the top 2 for seven of the ten subjects. In terms of model weight, the results more strongly favor the ACT-R model, which has the highest mean weights based on both AIC and BIC. These model comparison results indicate that the ACT-R model produced quantitative fits that were as good as those of the best SEF variant.

Recall that the ACT-R model also makes predictions about RT (Equation 9), for which individual-subject data were available from our experiment. The means of the individual-subject predictions for RT from the best-fitting ACT-R model (without a threshold) appear as solid lines in Figure 5. The ACT-R model not only provides a good fit to the shape of the RT function in Figure 5A, but it also correctly predicts the change in the fan effect across lags in Figure 5B (although the model overestimates the magnitude of the fan effect at intermediate and long lags). What is remarkable about these predictions is that they are zero-parameter, quantitative predictions of the ACT-R model. The only free parameter in the fit

was for response execution time, which functioned merely as an intercept shift for overall RT.

The model explains the RT function in Figure 5A in terms of the contribution of guessing time to RT (see Equation 9). At short lags, retrieval rarely finishes in the time available, so the model has to guess (see Figure 6B). At long lags, retrieval almost always finishes in the time available, so the model rarely has to guess (see Figure 6A). Given that it takes some time to guess ($t_{guess} = 50$ ms), it follows from Equation 9 that RT will become shorter across lags as guessing becomes less frequent. Note that Equation 9 and the changing contribution of guessing time to RT reflect the fact that there is a discrete point in time when retrieval is finished and the assumption that guessing is contingent on unfinished retrieval. The model's prediction of the change in the fan effect across lags in Figure 5B arises from differences in the contribution of guessing time to RT between fan conditions. At the shortest lag, retrieval almost never finishes for Fan 1 or Fan 2 items, which means that guessing occurs approximately equally often in both conditions. As a result, there is a similar contribution of guessing time to RT for both conditions, producing almost no fan effect. However, at the intermediate lags, retrieval finishes more often for Fan 1 items than for Fan 2 items, which means that guessing occurs more often for Fan 2 items. As a result, there is a larger contribution of guessing time to RT for Fan 2 items, producing a fan effect. The fan effect decreases as the lag becomes longer because the difference in the probability of retrieval between Fan 1 and Fan 2 items diminishes. By the longest lag, retrieval almost always finishes for both Fan 1 and Fan 2 items, which means that guessing rarely occurs in both conditions. As a result, the lack of a contribution of guessing time to RT is similar for both conditions, producing almost no fan effect. Thus, the shape of the RT function and the change in the fan effect across lags are natural predictions of the model (as evidenced by the fact that they are zero-parameter predictions) based on the relative contributions of retrieval and guessing to performance over time.

The best-fitting parameter values (for t_{respx}) and fit indices (RMSD and R^2) for individual subjects are presented in Table 7. All the parameter values seem reasonable. The ACT-R model yielded good fits in terms of RMSD, with a mean RMSD = 5.8 ms, and in terms of R^2 , with a mean $R^2 = .879$. The fit indices from simulations of the model (provided in parentheses in Table 7) are similar to the analytic fit indices, indicating that the lack of extra variability in the analytic version of the model was not detrimental to its fits to RT data.

Summary—Our ACT-R model provided an excellent account of the d' accuracy data from Wickelgren and Corbett (1977) and from our experiment. It captured the shapes of the speed–accuracy tradeoff functions in both experiments, including fan effects on asymptotic accuracy and the rate of rise in accuracy across lags (see Figures 2 and 4). It produced these results with only three or four free parameters (depending on whether a threshold was involved) and no free parameters that distinguished between fan conditions. Its fits were as good as those of the best SEF variant(s) for each experiment (see Figures 8 and 9), indicating that its quantitative predictions were satisfactory. An important advantage of our ACT-R model over the SEF is that the former is a process model that provides a psychological explanation of the data based on a cognitive theory, whereas the latter is more of a descriptive model that specifies a speed–accuracy tradeoff function without explaining how it is realized in a complete information-processing stream. One example of the advantage of using a complete process model is that our ACT-R model could also provide an account of the RT data from our experiment. It captured the shape of the RT function and the changes in the fan effect across lags. These results were zero-parameter, quantitative predictions of the model. Collectively, the modeling results show that our ACT-R model is capable of accounting for fine-grained data from the response signal procedure.

General Discussion

The purpose of the present study was to investigate fan effects on the time course of associative recognition and to develop an ACT-R model that could explain them. We focused on speed–accuracy tradeoff functions obtained with the response signal procedure in a previous experiment involving a fan manipulation (Wickelgren & Corbett, 1977) and in a new experiment. When materials were studied briefly, fan had a large effect on asymptotic accuracy at long lags and a modest effect on retrieval dynamics. When materials were well-learned, fan predominantly affected the rate of rise in accuracy across lags. We found that both data patterns could be explained by applying the standard ACT-R model of the fan effect to the response signal procedure. Memory retrieval plays a central role in the model, with high fan resulting in weak associative activation that slows memory retrieval, thereby decreasing the probability that retrieval finishes in time and producing a speed–accuracy tradeoff function. We fit the model to the data from Wickelgren and Corbett and from our experiment, comparing its fits with those of the SEF that is commonly used to describe response signal data. Our ACT-R model provided an excellent account of the data from both experiments, yielding quantitative fits to the accuracy data that were as good as those of the best SEF variant, as well as good quantitative fits to the RT data. To our knowledge, the present study represents the first successful application of ACT-R to response signal data, thereby expanding the scope of the theory to account for fine-grained behavioral data concerning the time course of associative recognition.

Comparisons with Other Models of Response Signal Data

Several process models have been developed for or applied to response signal data. These models include the cascade model (McClelland, 1979), the continuous activation model (Doshier, 1982, 1984a), the diffusion model (Ratcliff, 1978, 1988, 2006, 2008), a dual-process integration model (Göthe & Oberauer, 2008), the exemplar-based random walk model (Nosofsky, Little, Donkin, & Fific, 2011), the leaky competing accumulator model (Usher & McClelland, 2001), REM models (Diller, Nobel, & Shiffrin, 2001; Malmberg, 2008; Wagenmakers, Steyvers, Raaijmakers, Shiffrin, van Rijn, & Zeelenberg, 2004), and the tandem random walk model (Nikolić & Gronlund, 2002). Doshier (1976, 1981, 1984a) has also simulated speed–accuracy tradeoff functions for a serial search model (based on Anderson, 1974; Anderson & Bower, 1973; Thorndyke & Bower, 1974), a spreading activation model (based on Collins & Loftus, 1975), and the random storage model of Landauer (1975). In this section we discuss some of the similarities and differences between these models and our ACT-R model.

Discrete versus continuous processing—Doshier (1981, 1982, 1984a) noted that empirical speed–accuracy tradeoff functions are compatible with memory retrieval being either a discrete or a continuous process. Most models of response signal data assume that retrieval is continuous, with activation or evidence accruing over time (e.g., Diller et al., 2001; Doshier, 1982, 1984a; McClelland, 1979; Nikolić & Gronlund, 2002; Nosofsky et al., 2011; Ratcliff, 1978, 1988, 2006, 2008; Usher & McClelland, 2001; Wagenmakers et al., 2004). For example, in Ratcliff's diffusion model, evidence is sampled continuously from memory and accumulated in favor of making either a “yes” or a “no” recognition response. When the accumulated evidence favors one response over the other by a relative difference criterion, the favored response has been selected and retrieval is finished. The speed–accuracy tradeoff function produced by the model depends on the distribution of finishing times for retrieval (the sampling process is stochastic) and whether partial information (i.e., the evidence accumulated at any point prior to the end of retrieval) is available for making a decision when retrieval has not finished.

In contrast, our ACT-R model and the aforementioned model simulations by Doshier (1976, 1981, 1984a) assume that retrieval is a discrete, all-or-none process: It is either finished or not finished in the time available for retrieval, with no intermediate stage of progress (e.g., a fact cannot be partly retrieved from memory). However, similar to a continuous model, a discrete model can produce a speed–accuracy tradeoff function by having a distribution of finishing times for retrieval. In our ACT-R model, retrieval times follow a log-logistic distribution controlled by the level of activation noise (see Equation 6 and Figure 7). Although we have presented a discrete version of our ACT-R model, it may be possible to develop a continuous version that yields similar results. Anderson (2007, pp. 131–134) explained how the standard equation for retrieval time in ACT-R (Equation 3) can be related to the finishing time of a continuous evidence accumulation process (viz., a leaky accumulator model).⁸ Van Maanen and Van Rijn (2007, 2010; Van Maanen, Van Rijn, & Taatgen, in press) developed a variant of the declarative memory retrieval mechanism in ACT-R that is based on a continuous sampling process. Thus, the discrete–continuous processing dichotomy may not be critical for modeling speed–accuracy tradeoff functions in general or for the formulation of our ACT-R model in particular.

The role of partial information—A related issue concerns the role of partial information in response selection when retrieval does not have time to finish. There is no partial information in our ACT-R model because retrieval is a discrete, all-or-none process. Consequently, when retrieval does not finish, the model makes an uninformed guess that is affected only by a response bias (i.e., a general tendency to guess “yes” or “no”). In contrast, continuous models of response signal data allow for the possibility of responding based on partial information (for relevant discussion, see Meyer et al., 1988). Returning to the example of the diffusion model, if retrieval has not finished (i.e., the criterion for response selection has not been reached) by the time of the response signal, then the model can respond based on the evidence it has accumulated up to that point, choosing to go with the response for which there is more evidence (e.g., Ratcliff, 2006, 2008). An alternative approach is to use partial information to make an informed or biased guess. For example, in Nikolić and Gronlund's (2002) tandem random walk model, if retrieval has not finished in the time available, then the accumulated evidence is used to set a starting-point bias in favor of the response with more partial information for a subsequent guessing process. However, the availability of partial information in these models need not imply its use. Partial information could be ignored and a continuous model could make an uninformed guess in the same way as our ACT-R model. In the few instances where partial information and guessing versions of continuous models have been compared, it has been found that both versions yield very similar fits to response signal data (e.g., Nosofsky et al., 2011; Ratcliff, 2006, 2008). Thus, the absence of partial information in our ACT-R model may not be crucial for modeling speed–accuracy tradeoff functions.

Explanations of RT data—Despite the interest in modeling speed–accuracy tradeoff functions, very little attention has been paid to explaining RT data from the response signal procedure. As noted earlier, RT typically becomes shorter as the lag becomes longer in response signal experiments (e.g., see Figure 3B). This RT function is generally thought to reflect either variation in response preparedness across lags (i.e., subjects are more prepared to respond at long lags than at short lags) or some sort of dual-task interference effect (i.e., task processing is impaired or postponed by having to process the response signal, especially at short lags when the two processing streams are likely to overlap; Pashler, 1994).

⁸The relationship can be interpreted in terms of a retrieval process whereby there is continuous accumulation of evidence for retrieving an item from memory, but the item is only delivered to the retrieval buffer in ACT-R when its accumulated evidence reaches a threshold. This process is analogous to models of eye movements in which evidence accumulates for making a saccade but the saccade itself is ballistic (e.g., Hanes & Schall, 1996).

However, as pointed out by Reed (1976), RT data may provide useful information for evaluating models (see also Nikolić & Gronlund, 2002; Ratcliff, 2006). Indeed, in our experiment we observed a fan effect on RT that varied across lags (see Figure 5) and our ACT-R model was able to predict it quantitatively with zero free parameters.

Our ACT-R model explains the RT function in terms of the changing contribution of guessing time to RT across lags, although it does not exclude the possibility that either response preparedness or dual-task interference may play a role.⁹ As discussed earlier, retrieval rarely finishes at short lags, forcing the model to guess (see Figure 6B), whereas retrieval almost always finishes at long lags, making guessing unnecessary (see Figure 6A). Given that it takes some time to guess (a point that is overlooked in many models), it follows from Equation 9 that RT will become shorter across lags as guessing becomes less frequent, with the difference in RT between the shortest and the longest lags being approximately equal to the guessing time. For modeling the RT data from our experiment, we set the guessing time, t_{guess} , equal to 50 ms, which is the default time for firing a single production in ACT-R (Anderson, 2007). This turned out to be nearly optimal for predicting the magnitude of the change in RT across lags (54 ms; see Figure 5A). While there are other experiments in which the differences between the longest and shortest RTs were in the same ballpark (e.g., Corbett, 1977; Hintzman et al., 1998; Hintzman & Curran, 1997; Ratcliff, 2006; Wickelgren & Corbett, 1977), there are others in which differences greater than 100 ms were found (e.g., Doshier, 1976, 1981, 1982, 1984a). We suspect the discrepancies between studies are primarily methodological, reflecting differences in the instructions, motivation, and feedback given to subjects about their RTs, coupled with varying criteria for excluding long RTs from data analyses. However, it is also possible that guessing may take longer than 50 ms in some situations. Thus, to accommodate larger RT differences, it might be preferable to allow guessing time to be a free parameter when fitting certain data sets.

Consideration of Response Signal Evidence for Dual Processes in Recognition

The ACT-R model presented in this article assumes that associative recognition reflects a single process: retrieval of an associative chunk from declarative memory. This single-process account was sufficient for explaining the time course of recognition in contexts where subjects had to discriminate between studied and rearranged pairs of items. However, there are response signal studies involving other types of discriminations that implicate more than one process in recognition. A representative example is a study by Rotello and Heit (2000). In their Experiment 1, subjects studied a list of word pairs, then they were tested with three kinds of probes: targets (studied word pairs), rearranged foils (alternative pairings of studied words), and new foils (completely new word pairs). The test phase involved the response signal procedure and the task was to recognize whether the probe words had been studied together. Rotello and Heit found that accuracy increased monotonically across lags for targets and for new foils, whereas it decreased across the early lags and then increased across the late lags for rearranged foils (as reflected by an initial increase and then a decrease in the false alarm rate). The nonmonotonic time-course function for rearranged foils was interpreted as evidence for two processes involved in recognition: an early process that discriminated between studied words (which composed the targets and the rearranged foils) and non-studied words (which composed the new foils), and a late process that discriminated between studied and non-studied word pairs (targets and rearranged foils, respectively).

⁹Nikolić and Gronlund (2002) modeled the RT function in a similar way, albeit with a more elaborate guessing process that was supplemented by a dual-task interference effect.

Nonmonotonic time-course functions of this sort have been observed in several response signal studies (Doshier, 1984b; Göthe & Oberauer, 2008; Gronlund & Ratcliff, 1989; Hintzman & Curran, 1994; McElree, Dolan, & Jacoby, 1999; McElree & Doshier, 1989; Ratcliff & McKoon, 1982, 1989; Rotello & Heit, 2000; for a review, see Rotello, 2000). One interpretation of these functions is that they reflect the retrieval of item and associative information at different points in time, with item information being retrieved more quickly than associative information. Applied to Rotello and Heit's experiment, the retrieval of item information was the early process that led to the increase in the false alarm rate for rearranged foils because they were recognized as being composed of studied words, whereas the retrieval of associative information was the late process that led to the subsequent decrease in the false alarm rate because the words were recognized as not having been studied together. The retrieval of associative information to reject the rearranged foils is the recall-to-reject strategy implemented in our ACT-R model. Rotello and Heit argued that their nonmonotonic time-course functions represented evidence that such a strategy is used in associative recognition (see also Rotello, 2000).

According to the preceding interpretation, the two processes underlying nonmonotonic time-course functions do not necessarily involve different retrieval mechanisms; rather, they involve the retrieval of different kinds of information (viz., item and associative). An alternative but related interpretation of nonmonotonic time-course functions is that they reflect two qualitatively different recognition processes: familiarity and recollection. According to many dual-process theories of recognition (e.g., Atkinson & Juola, 1974; Diana, Reder, Arndt, & Park, 2006; Jacoby, 1991; Mandler, 1980; Yonelinas, 1994; for a review, see Yonelinas, 2002), familiarity reflects a fast assessment of a probe's continuous memory strength (often modeled as a signal-detection process), whereas recollection reflects a slower, recall-based mechanism by which qualitative or associative information about the probe is retrieved. Applied to Rotello and Heit's (2000) experiment, familiarity was responsible for the early increase in the false alarm rate for rearranged foils because studied words were represented more strongly in memory than were non-studied words, whereas recollection was responsible for the subsequent decrease in the false alarm rate because the retrieval of associative information enabled rejection of rearranged foils. Thus, the main difference between the two dual-process interpretations concerns the first process (retrieval of item information versus familiarity), whereas the second process (retrieval of associative information versus recollection) is essentially the same in both cases.

In contrast with the modeling work cited earlier, computational models of nonmonotonic time-course functions are relatively rare. Malmberg (2008) presented simulation results for a dual-process REM model and showed that it produced a nonmonotonic time-course function, although he did not compare its predictions with experimental data. Göthe and Oberauer (2008) presented analytic results for variants of a dual-process model based on either integration of accumulated evidence from familiarity and recollection or dominance of an all-or-none recollection process over a continuous familiarity signal. Both model variants were able to produce the nonmonotonic time-course function observed in their data. However, their model provides only an abstract characterization of the underlying familiarity and recollection processes (e.g., evidence accumulation for each process is modeled with an SEF). Ratcliff (1980) derived an expression from the diffusion model for a nonmonotonic time-course function based on a change in the rate of evidence accumulation during retrieval (which might occur if new information becomes available later in processing) and it has been shown to do fairly well at capturing the shapes of empirical nonmonotonic time-course functions (e.g., Doshier, 1984b; Gronlund & Ratcliff, 1989; McElree & Doshier, 1989; Rotello & Heit, 2000). However, even though the diffusion model is a process model, its applications to nonmonotonic time-course functions have tended to be

more descriptive than explanatory (e.g., the diffusion model is neutral with respect to what kind of information changes the rate of evidence accumulation).

Our ACT-R model, in its present form, does not produce nonmonotonic time-course functions because there is only one retrieval process and only associative information is retrieved. It may be possible to address this limitation by creating a dual-process version of the model. Following the first dual-process interpretation described earlier, the retrieval of associative information (the process currently in our model) could be preceded by the retrieval of item information (for relevant empirical work concerning the fan effect, see Anderson, 1975). Applied to Rotello and Heit's (2000) experiment, item information would serve as the basis for rejecting new foils and falsely accepting rearranged foils at short to intermediate lags, whereas associative information would serve as the basis for rejecting rearranged foils at long lags. This extension to the model seems feasible in principle, although future work will be needed to determine the best way of implementing both retrieval processes and to assess the extended model's adequacy in accounting for empirical nonmonotonic time-course functions.

Prospects for Modeling RT Distributions

By showing how ACT-R can be applied successfully to speed–accuracy tradeoffs in response signal data, the present study provides a foundation for further applications of ACT-R to fine-grained temporal data. In particular, it allows one to make predictions for RT distributions in free-response paradigms, wherein response signals are absent and subjects are free to respond as soon as they have selected a response. Subjects tend to show considerable variability in their responding, as reflected in RT distributions that are often unimodal and positively skewed (e.g., see Luce, 1986).

Our ACT-R model for response signal data can be adapted to the free-response paradigm by having response execution immediately follow memory retrieval in the information-processing stream. Assuming there is activation noise in memory, retrieval times will follow a log-logistic distribution (see Equation 6), which under many parameterizations is unimodal and positively skewed (see Figure 7). However, the overall RT distribution produced by the model also depends on the distributions of times for stimulus encoding and response execution. To produce the simulation results mentioned earlier, we assumed each process had a uniform distribution of finishing times, with the boundaries of the distribution being $t \pm t/2$, where t is the mean time for the process (see Anderson, Taatgen, & Byrne, 2005; Meyer & Kieras, 1997). This assumption was made for simplicity and to limit the number of free parameters. However, it would be reasonable to consider alternative distributions (e.g., exponential or gamma) for these processes when modeling free-response RT distributions.

Another important consideration concerns the modeling of errors and error RT distributions. In our ACT-R model for response signal data, errors arise solely from guessing, which is contingent on whether retrieval beats the external deadline (t_{avail}) determined by the response signal and the internal deadline ($t_{failure}$) set for the retrieval process. Given that there is typically no external deadline in the free-response paradigm, guessing occurs only if retrieval takes longer than the internal deadline (i.e., $t_{retrieve} > t_{failure}$), resulting in errors that are slower on average than correct responses.

In empirical data from choice RT tasks in free-response paradigms, errors can be slower or faster than correct responses depending on experimental conditions such as task difficulty and speed–accuracy emphasis (e.g., Luce, 1986; Ratcliff & Rouder, 1998). For our ACT-R model to produce fast errors, one would have to change assumptions about the guessing process. One approach would be to remove the contingency of guessing on unfinished retrieval (e.g., by eliminating the internal deadline) and have a probabilistic guessing process

that either runs in isolation (i.e., the model guesses without attempting retrieval on some trials; e.g., Luce, 1986; Ollman, 1966; Yellott, 1967, 1971) or races in parallel against retrieval (in which case, guessing time serves as an internal deadline for retrieval; Yellott, 1971). If the guessing process is faster than retrieval, then errors (which come only from guessing) would be faster than correct responses (which come from guessing and retrieval) and the error RT distribution would reflect the distribution of guessing times.

For example, consider an alternative version of our ACT-R model for response signal data in which a guessing process is initiated early during a trial and runs in parallel with stimulus encoding and memory retrieval. Assume that the guessed response is used only if retrieval has not finished by the time the response signal is encoded; otherwise, the retrieved response is used, even if the guessed response is available first. In other words, use of the outcome of guessing—rather than the initiation of the guessing process—is contingent on unfinished retrieval. This alternative model would produce the same speed–accuracy tradeoff function in the response signal procedure as does our original model because in both cases the contribution of guessing to accuracy depends on the probability that retrieval finishes in time (see Equations 6 and 8). However, in the context of a free-response paradigm, wherein a response can be made as soon as either guessing or retrieval has finished, the alternative model would produce errors that are faster than correct responses because guessing would sometimes finish before retrieval. Whether this alternative or some other variant of our ACT-R model could capture the full pattern of behavioral data (viz., the probabilities of correct and error responses and their corresponding RT distributions) in a free-response paradigm involving associative recognition is an intriguing issue that warrants further investigation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by Office of Naval Research Grant N00014-03-1-0115 and National Institute of Mental Health Grant MH068243. We thank Abraham Anderson for running the subjects in the experiment, Charles Liu for clarifying methodological details from Liu and Smith (2009), and William Hockley, Caren Rotello, and an anonymous reviewer for their comments on this work.

Appendix A

Appendix A

Persons and locations used for the facts in our experiment

Person	Location				
actor	farmer	pirate	airport	garage	office
chef	gardener	queen	attic	hallway	park
coach	hippie	scientist	bank	hotel	prison
cowboy	inventor	sheriff	barn	kitchen	school
dancer	judge	soldier	castle	laboratory	stadium
detective	maid	spy	church	library	studio
doctor	musician	teacher	clinic	museum	temple
engineer	pilot	tourist	factory	nightclub	theater

Appendix B

Individual-subject data and modeling results are available as supplementary material. (See the Excel file linked to this manuscript.)

References

- Anderson JR. Retrieval of propositional information from long-term memory. *Cognitive Psychology*. 1974; 6:451–474.
- Anderson JR. Item-specific and relation-specific interference in sentence memory. *Journal of Experimental Psychology: Human Learning and Memory*. 1975; 104:249–260.
- Anderson, JR. Language, memory, and thought. Erlbaum; Hillsdale, NJ: 1976.
- Anderson JR. Interference: The relationship between response latency and response accuracy. *Journal of Experimental Psychology: Human Learning and Memory*. 1981; 7:326–343.
- Anderson, JR. The architecture of cognition. Harvard University Press; Cambridge, MA: 1983.
- Anderson, JR. The adaptive character of thought. Erlbaum; Hillsdale, NJ: 1990.
- Anderson JR. Is human cognition adaptive? *Behavioral and Brain Sciences*. 1991; 14:471–484.
- Anderson, JR. How can the human mind occur in the physical universe?. Oxford University Press; New York: 2007.
- Anderson JR, Bothell D, Byrne MD, Douglass S, Lebiere C, Qin Y. An integrated theory of the mind. *Psychological Review*. 2004; 111:1036–1060. [PubMed: 15482072]
- Anderson JR, Bothell D, Douglass S. Eye movements do not reflect retrieval processes: Limits of the eye-mind hypothesis. *Psychological Science*. 2004; 15:225–231. [PubMed: 15043638]
- Anderson JR, Bothell D, Lebiere C, Matessa M. An integrated theory of list memory. *Journal of Memory and Language*. 1998; 38:341–380.
- Anderson, JR.; Bower, GH. Human associative memory. Winston and Sons; Washington, D. C.: 1973.
- Anderson, JR.; Lebiere, C. The atomic components of thought. Erlbaum; Mahwah, NJ: 1998.
- Anderson JR, Matessa M. A production system theory of serial memory. *Psychological Review*. 1997; 104:728–748.
- Anderson JR, Paulson R. Interference in memory for pictorial information. *Cognitive Psychology*. 1978; 10:178–202. [PubMed: 668298]
- Anderson JR, Reder LM. The fan effect: New results and new theories. *Journal of Experimental Psychology: General*. 1999; 128:186–197.
- Anderson JR, Reder LM, Lebiere C. Working memory: Activation limitations on retrieval. *Cognitive Psychology*. 1996; 30:221–256. [PubMed: 8660785]
- Anderson JR, Schooler LJ. Reflections of the environment in memory. *Psychological Science*. 1991; 2:396–408.
- Anderson JR, Taatgen NA, Byrne MD. Learning to achieve perfect timesharing: Architectural implications of Hazeltine, Teague, and Ivry (2002). *Journal of Experimental Psychology: Human Perception and Performance*. 2005; 31:749–761. [PubMed: 16131247]
- Atkinson, RC.; Juola, JF. Search and decision processes in recognition memory.. In: Krantz, DH.; Atkinson, RC.; Luce, RD.; Suppes, P., editors. *Contemporary developments in mathematical psychology*. Vol. 1. W. H. Freeman; San Francisco, CA: 1974. p. 243-293.
- Cohen AL, Sanborn AN, Shiffrin RM. Model evaluation using grouped or individual data. *Psychonomic Bulletin & Review*. 2008; 15:692–712. [PubMed: 18792497]
- Collins AM, Loftus EF. A spreading-activation theory of semantic processing. *Psychological Review*. 1975; 82:407–428.
- Corbett AT. Retrieval dynamics for rote and visual image mnemonics. *Journal of Verbal Learning and Verbal Behavior*. 1977; 16:233–246.
- Danker JF, Gunn P, Anderson JR. A rational account of memory predicts left prefrontal activation during controlled retrieval. *Cerebral Cortex*. 2008; 18:2674–2685. [PubMed: 18321871]
- Diana RA, Reder LM, Arndt J, Park H. Models of recognition: A review of arguments in favor of a dual-process account. *Psychonomic Bulletin & Review*. 2006; 13:1–21. [PubMed: 16724763]

- Diller DE, Nobel PA, Shiffrin RM. An ARC-REM model for accuracy and response time in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2001; 27:414-435.
- Dosher BA. The retrieval of sentences from memory: A speed-accuracy study. *Cognitive Psychology*. 1976; 8:291-310.
- Dosher BA. Empirical approaches to information processing: Speed-accuracy tradeoff functions or reaction time—a reply. *Acta Psychologica*. 1979; 43:347-359.
- Dosher BA. The effects of delay and interference: A speed-accuracy study. *Cognitive Psychology*. 1981; 13:551-582.
- Dosher BA. Effect of sentence size and network distance on retrieval speed. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1982; 8:173-207.
- Dosher BA. Degree of learning and retrieval speed: Study time and multiple exposures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1984a; 10:541-574.
- Dosher BA. Discriminating preexperimental (semantic) from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology*. 1984b; 16:519-555.
- Dyne AM, Humphreys MS, Bain JD, Pike R. Associative interference effects in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1990; 16:813-824.
- Glover S, Dixon P. Likelihood ratios: A simple and flexible statistic for empirical psychologists. *Psychonomic Bulletin & Review*. 2004; 11:791-806. [PubMed: 15732688]
- Göthe K, Oberauer K. The integration of familiarity and recollection information in short-term recognition: Modeling speed-accuracy trade-off functions. *Psychological Research*. 2008; 72:289-303. [PubMed: 17357795]
- Gronlund SD, Ratcliff R. Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1989; 15:846-858.
- Hanes DP, Schall JD. Neural control of voluntary movement initiation. *Science*. 1996; 274:427-430. [PubMed: 8832893]
- Hintzman DL, Caulton DA, Levitin DJ. Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall. *Memory & Cognition*. 1998; 26:449-462.
- Hintzman DL, Curran T. Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language*. 1994; 33:1-18.
- Hintzman DL, Curran T. Comparing retrieval dynamics in recognition memory and lexical decision. *Journal of Experimental Psychology: General*. 1997; 126:228-247.
- Jacoby LL. A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*. 1991; 30:513-541.
- King DRW, Anderson JR. Long-term memory search: An intersecting activation process. *Journal of Verbal Learning and Verbal Behavior*. 1976; 15:587-605.
- Landauer TK. Memory without organization: Properties of a model with random storage and undirected retrieval. *Cognitive Psychology*. 1975; 7:495-531.
- Lewandowsky, S.; Farrell, S. *Computational modeling in cognition: Principles and practice*. Sage; Thousand Oaks, CA: 2011.
- Lewis CH, Anderson JR. Interference with real world knowledge. *Cognitive Psychology*. 1976; 8:311-335.
- Liu CC, Smith PL. Comparing time-accuracy curves: Beyond goodness-of-fit measures. *Psychonomic Bulletin & Review*. 2009; 16:190-203. [PubMed: 19145032]
- Luce, RD. *Response times*. Oxford University Press; New York: 1986.
- Malmberg KJ. Recognition memory: A review of the critical findings and an integrated theory for relating them. *Cognitive Psychology*. 2008; 57:335-384. [PubMed: 18485339]
- Mandler G. Recognizing: The judgment of previous occurrence. *Psychological Review*. 1980; 87:252-271.
- McClelland JL. On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*. 1979; 86:287-330.

- McElree B, Dolan PO, Jacoby LL. Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1999; 25:563–582.
- McElree B, Doshier BA. Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General*. 1989; 118:346–373.
- Meyer DE, Irwin DE, Osman AM, Kounios J. The dynamics of cognition and action: Mental processes inferred from speed–accuracy decomposition. *Psychological Review*. 1988; 95:183–237. [PubMed: 3375399]
- Meyer DE, Kieras DE. A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*. 1997; 104:3–65. [PubMed: 9009880]
- Myung IJ. Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*. 2003; 47:90–100.
- Nelder JA, Mead R. A simplex method for function minimization. *Computer Journal*. 1965; 7:308–313.
- Nikolić D, Gronlund SD. A tandem random walk model of the SAT paradigm: Response times and accumulation of evidence. *British Journal of Mathematical and Statistical Psychology*. 2002; 55:263–288. [PubMed: 12473228]
- Nosofsky RM, Little DR, Donkin C, Fific M. Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review*. 2011; 118:280–315. [PubMed: 21355662]
- Ollman R. Fast guesses in choice reaction time. *Psychonomic Science*. 1966; 6:155–156.
- Pachella, RG. The interpretation of reaction time in information-processing research.. In: Kantowitz, BH., editor. *Human information processing: Tutorials in performance and cognition*. Erlbaum; Hillsdale, NJ: 1974. p. 41-82.
- Pashler H. Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*. 1994; 116:220–244. [PubMed: 7972591]
- Pirolli PL, Anderson JR. The role of practice in fact retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1985; 11:136–153.
- Postman, L. Interference theory revisited.. In: Brown, J., editor. *Recall and recognition*. Wiley; New York: 1976. p. 157-181.
- Ratcliff R. A theory of memory retrieval. *Psychological Review*. 1978; 85:59–108.
- Ratcliff R. A note on modeling accumulation of information when the rate of accumulation changes over time. *Journal of Mathematical Psychology*. 1980; 21:178–184.
- Ratcliff R. Continuous versus discrete information processing: Modeling accumulation of partial information. *Psychological Review*. 1988; 95:238–255. [PubMed: 3375400]
- Ratcliff R. Modeling response signal and response time data. *Cognitive Psychology*. 2006; 53:195–237. [PubMed: 16890214]
- Ratcliff R. Modeling aging effects on two-choice tasks: Response signal and response time data. *Psychology and Aging*. 2008; 23:900–916. [PubMed: 19140659]
- Ratcliff R, McKoon G. Speed and accuracy in the processing of false statements about semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1982; 8:16–36.
- Ratcliff R, McKoon G. Similarity information versus relational information: Differences in the time course of retrieval. *Cognitive Psychology*. 1989; 21:139–155. [PubMed: 2706926]
- Ratcliff R, Rouder JN. Modeling response times for two-choice decisions. *Psychological Science*. 1998; 9:347–356.
- Rawson KA, Dunlosky J. Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General*. 2011; 140:283–302. [PubMed: 21707204]
- Reed AV. Speed-accuracy trade-off in recognition memory. *Science*. 1973; 181:574–576. [PubMed: 17777808]
- Reed AV. List length and the time course of recognition in immediate memory. *Memory & Cognition*. 1976; 4:16–30.

- Rotello, CM. Recall processes in recognition memory.. In: Medin, DL., editor. The psychology of learning and motivation. Vol. 40. Academic Press; San Diego, CA: 2000. p. 183-221.
- Rotello CM, Heit E. Associative recognition: A case of recall-to-reject processing. *Memory & Cognition*. 2000; 28:907–922.
- Rotello CM, Macmillan NA, Van Tassel G. Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory and Language*. 2000; 43:67–88.
- Schneider DW, Anderson JR. A memory-based model of Hick's law. *Cognitive Psychology*. 2011; 62:193–222. [PubMed: 21293788]
- Schouten JF, Bekker JAM. Reaction time and accuracy. *Acta Psychologica*. 1967; 27:143–153. [PubMed: 6062205]
- Snodgrass JG, Corwin J. Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*. 1988; 117:34–50. [PubMed: 2966230]
- Sohn M-H, Goode A, Stenger VA, Carter CS, Anderson JR. Competition and representation during memory retrieval: Roles of the prefrontal cortex and the posterior parietal cortex. *Proceedings of the National Academy of Sciences*. 2003; 100:7412–7417.
- Sohn M-H, Goode A, Stenger VA, Jung K-J, Carter CS, Anderson JR. An information-processing model of three cortical regions: Evidence in episodic memory retrieval. *NeuroImage*. 2005; 25:21–33. [PubMed: 15734340]
- Stevens M, Lammertyn J, Verbruggen F, Vandierendonck A. Tscope: A C library for programming cognitive experiments on the MS Windows platform. *Behavior Research Methods*. 2006; 38:280–286. [PubMed: 16956104]
- Surprenant, AM.; Neath, I. Principles of memory. Psychology Press; New York: 2009.
- Thorndyke PW, Bower GH. Storage and retrieval processes in sentence memory. *Cognitive Psychology*. 1974; 6:515–543.
- Usher M, McClelland JL. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*. 2001; 108:550–592. [PubMed: 11488378]
- Van Maanen L, Van Rijn H. An accumulator model of semantic interference. *Cognitive Systems Research*. 2007; 8:174–181.
- Van Maanen L, Van Rijn H. The locus of the Gratton effect in picture–word interference. *Topics in Cognitive Science*. 2010; 2:168–180.
- Van Maanen L, Van Rijn H, Taatgen NA. Accumulators in context: An integrated theory of context effects on memory retrieval. *Cognitive Science*. (in press).
- Verde MF. Associative interference in recognition memory: A dual-process account. *Memory & Cognition*. 2004; 32:1273–1283.
- Wagenmakers E-J. A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*. 2007; 14:779–804. [PubMed: 18087943]
- Wagenmakers E-J, Farrell S. AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*. 2004; 11:192–196. [PubMed: 15117008]
- Wagenmakers E-J, Steyvers M, Raaijmakers JGW, Shiffrin RM, van Rijn H, Zeelenberg R. A model for evidence accumulation in the lexical decision task. *Cognitive Psychology*. 2004; 48:332–367. [PubMed: 15020215]
- Watkins OC, Watkins MJ. Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory*. 1975; 104:442–452.
- White AS, Cerella J, Hoyer WJ. Strategy transitions during cognitive skill learning in younger and older adults: Effects of interitem confusability. *Memory & Cognition*. 2007; 35:2106–2117.
- Wickelgren WA. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*. 1977; 41:67–85.
- Wickelgren WA, Corbett AT. Associative interference and retrieval dynamics in yes–no recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory*. 1977; 3:189–202.
- Wickelgren WA, Corbett AT, Doshier BA. Priming and retrieval from short-term memory: A speed accuracy trade-off analysis. *Journal of Verbal Learning and Verbal Behavior*. 1980; 19:387–404.
- Yellott JI. Correction for guessing in choice reaction time. *Psychonomic Science*. 1967; 8:321–322.

- Yellott JI. Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology*. 1971; 8:159–199.
- Yonelinas AP. Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1994; 20:1341–1354.
- Yonelinas AP. The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*. 2002; 46:441–517.
- Zbrodoff NJ. Why is $9 + 7$ harder than $2 + 3$? Strength and interference as explanations of the problem-size effect. *Memory & Cognition*. 1995; 23:689–700.

Highlights

- We examine how associative fan affects the time course of recognition.
- High fan lowers asymptotic accuracy or the rate of rise in accuracy, or both.
- We develop an ACT-R model that produces speed–accuracy tradeoff functions.
- The model provides an excellent account of fan effects in response signal data.

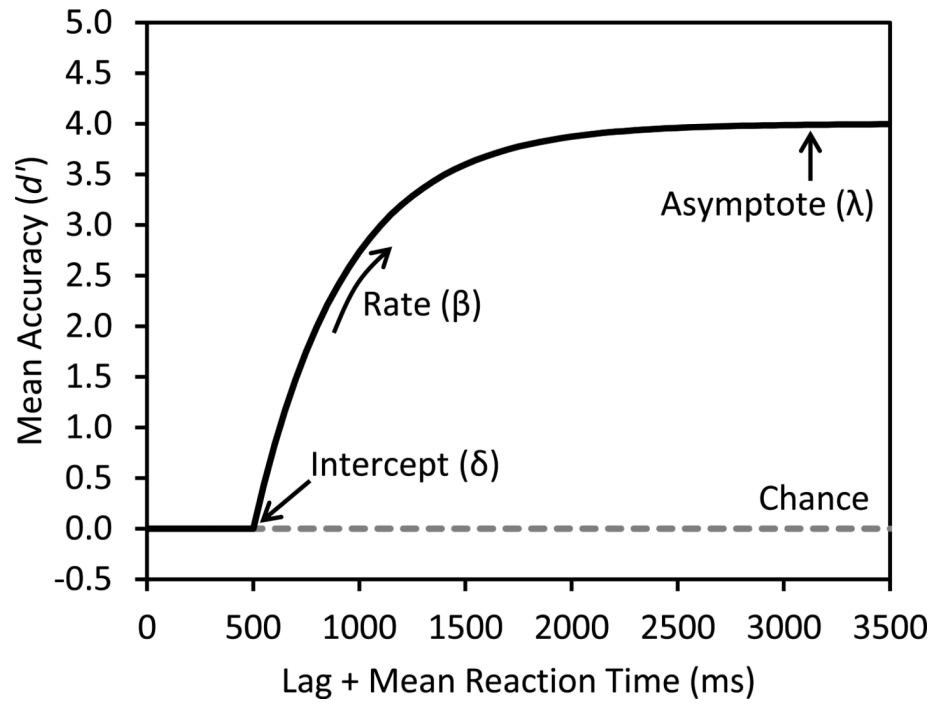


Figure 1. Idealized example of a speed-accuracy tradeoff function from the response signal procedure.

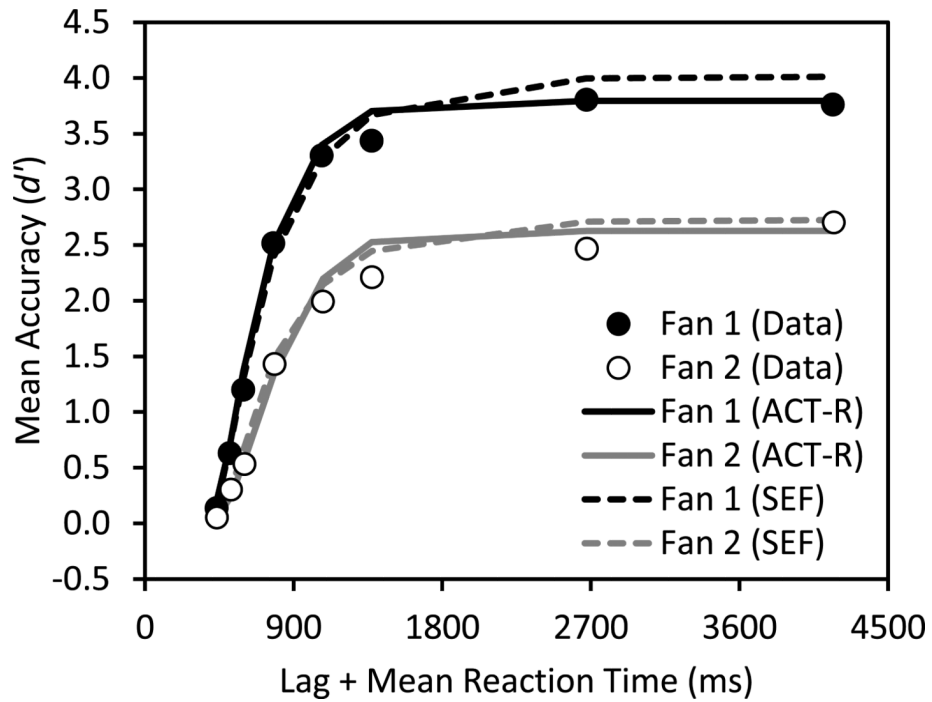


Figure 2. Speed-accuracy tradeoff functions for the Fan 1 and Fan 2 conditions in Wickelgren and Corbett (1977). Points denote data, solid lines denote ACT-R model predictions, and dashed lines denote shifted exponential function (SEF) predictions.

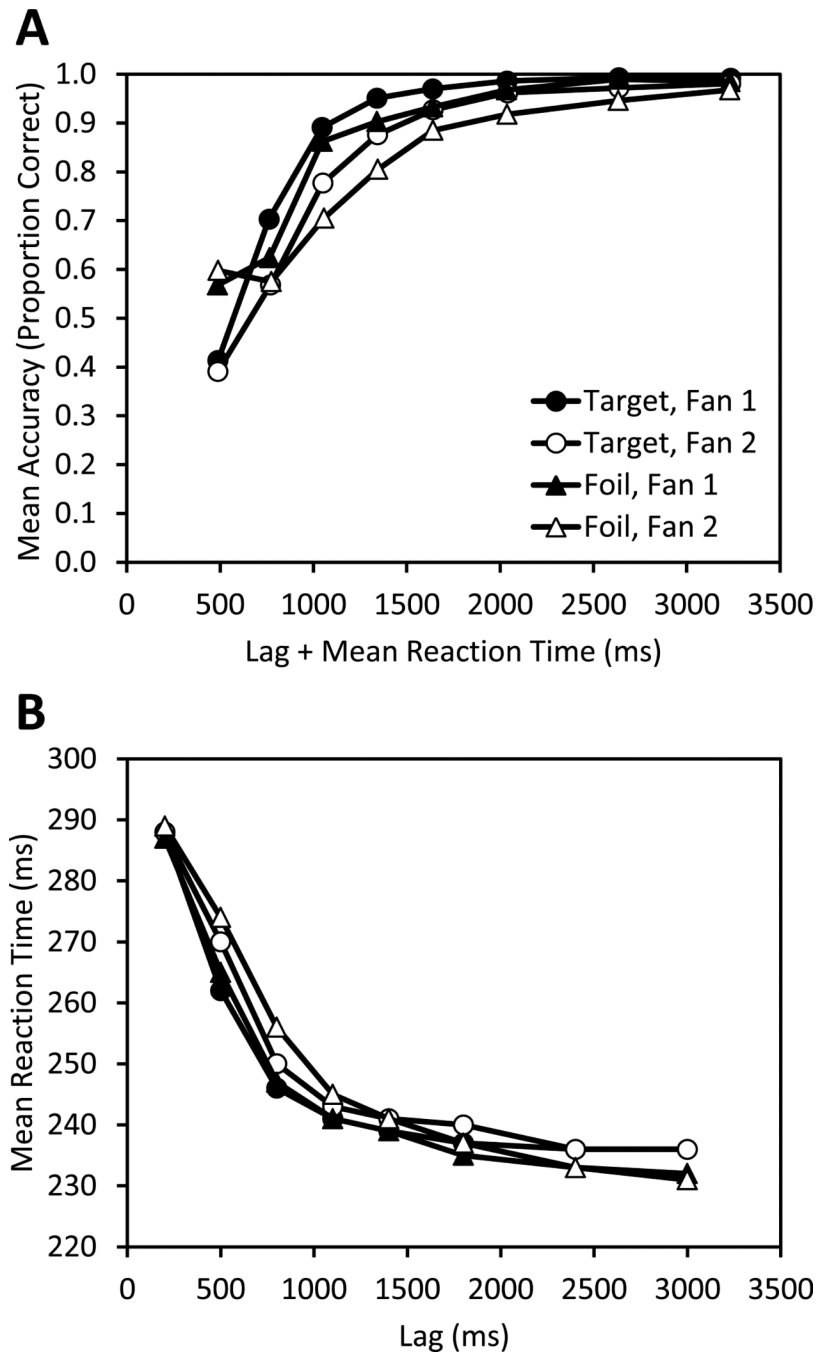


Figure 3. Data from our experiment as a function of probe (target or foil) and fan (1 or 2). A: Accuracy data. B: Reaction time data.

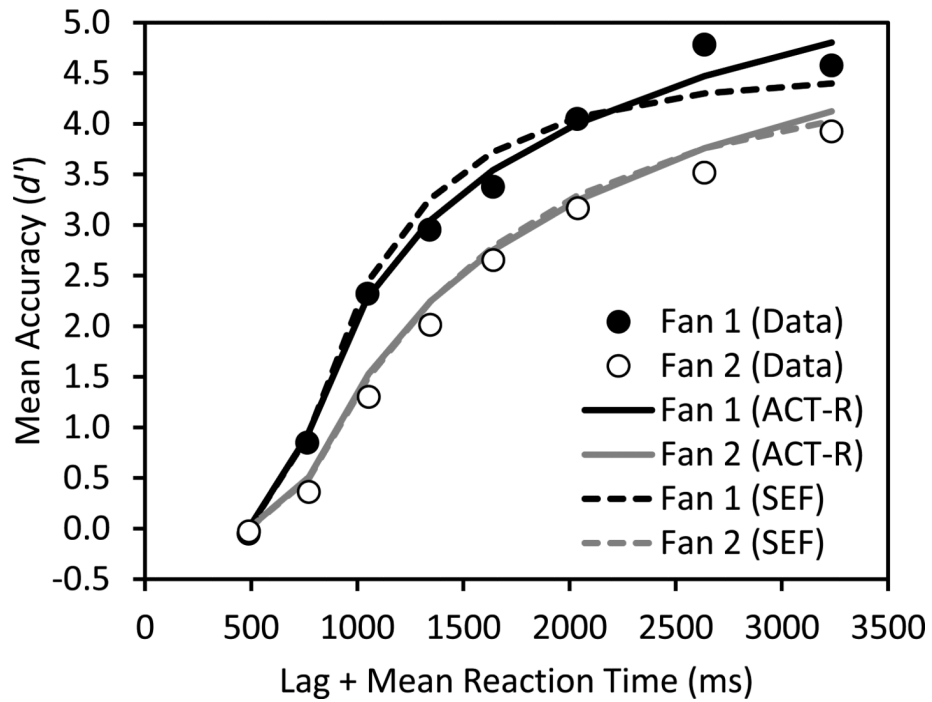


Figure 4. Speed-accuracy tradeoff functions for the Fan 1 and Fan 2 conditions in our experiment. Points denote data, solid lines denote ACT-R model predictions, and dashed lines denote shifted exponential function (SEF) predictions.

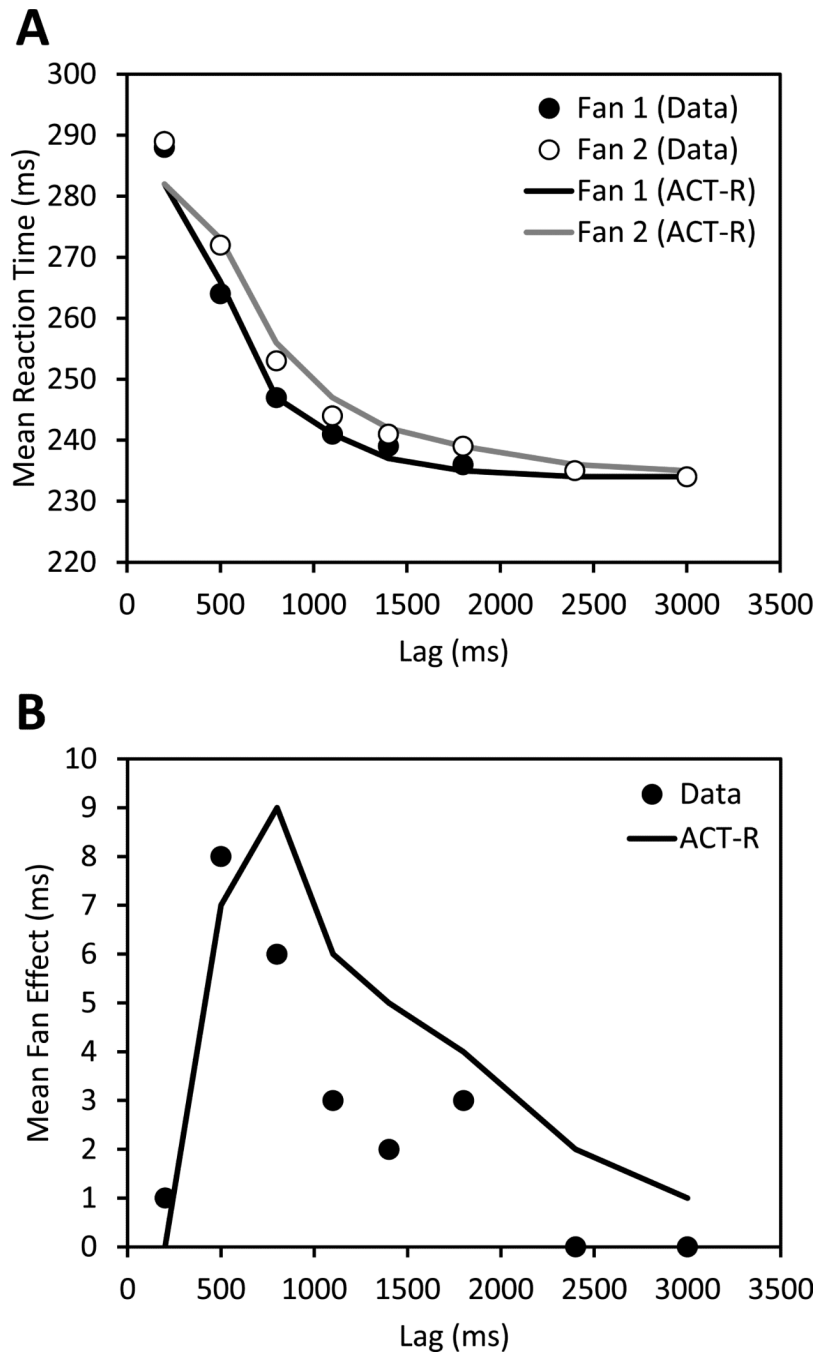
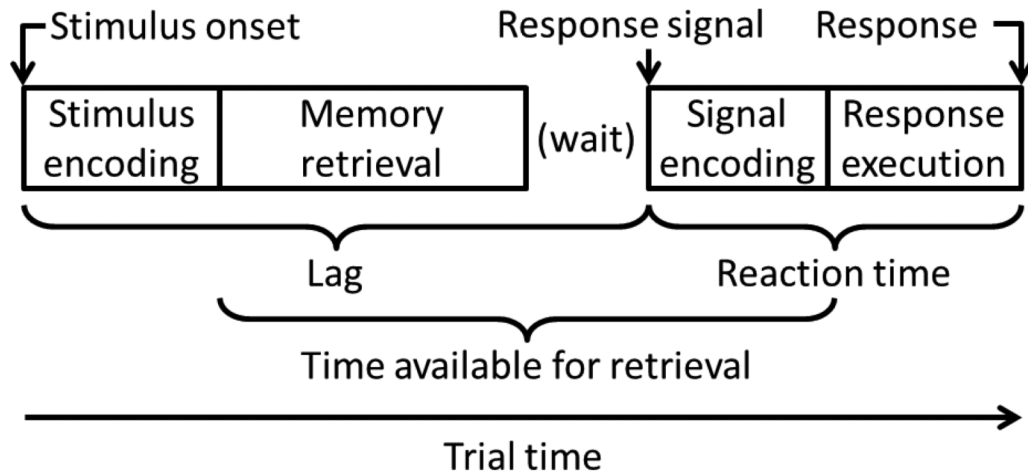


Figure 5. Reaction time data (points) and ACT-R model predictions (solid lines) for our experiment. A: Reaction time for the Fan 1 and Fan 2 conditions. B: Fan effect.

A: Long Lag



B: Short Lag

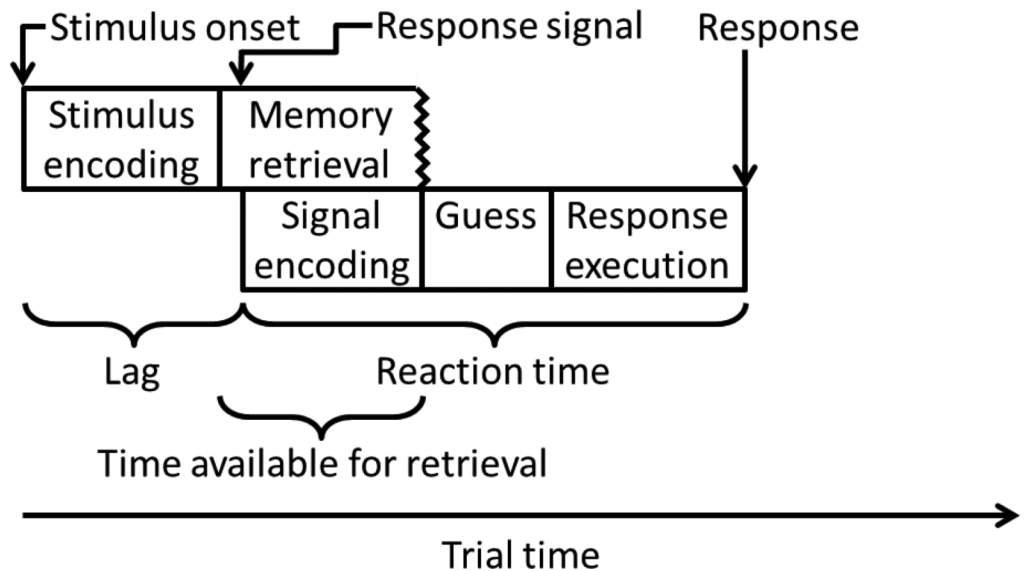


Figure 6. Schematic illustration of the organization of the ACT-R model's processing stages at long and short lags (panels A and B, respectively).

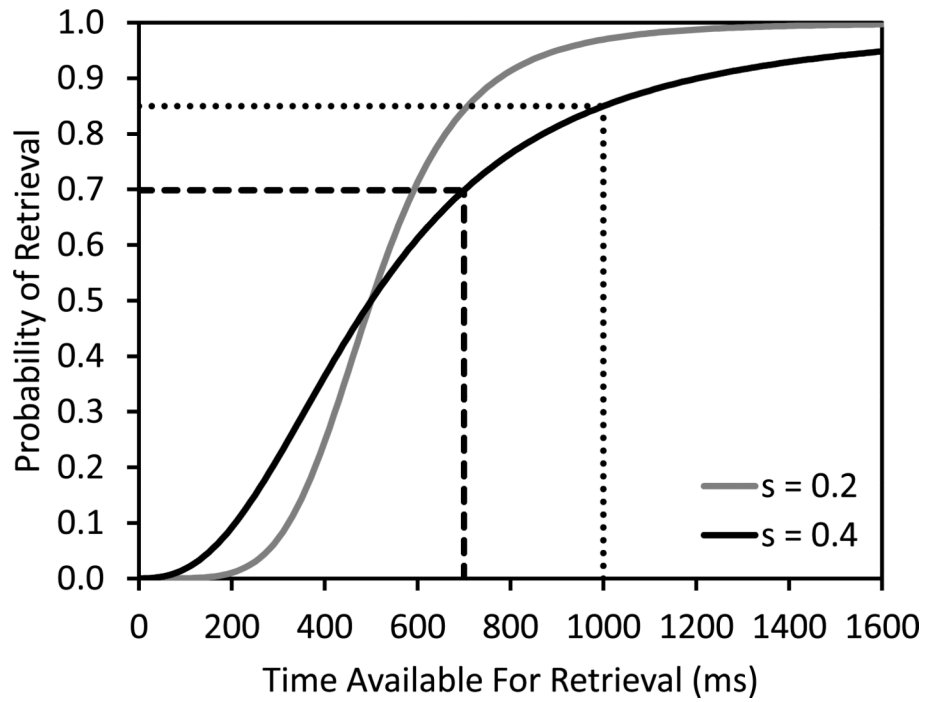


Figure 7. Examples of the log-logistic cumulative distribution function for two levels of activation noise (s).

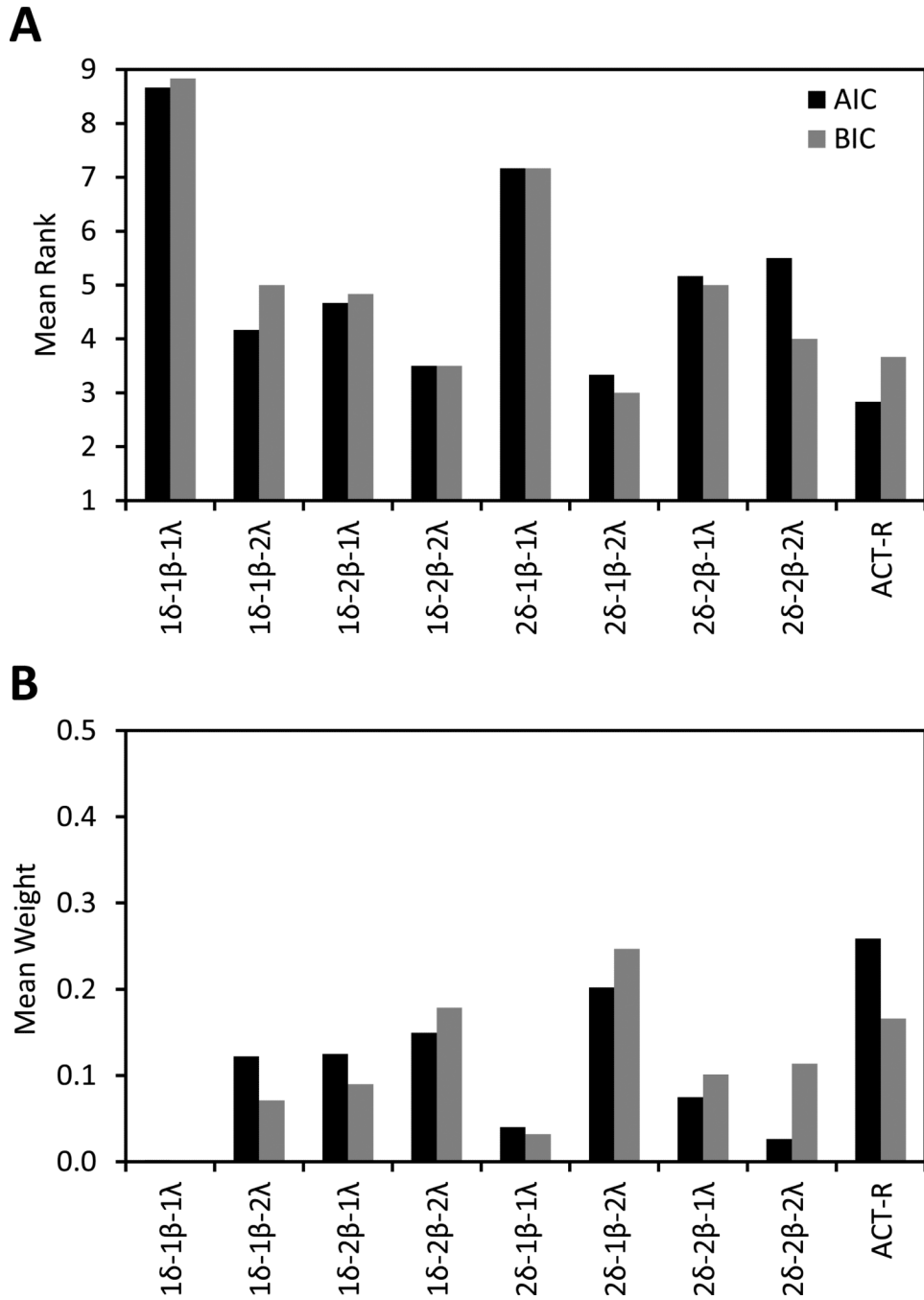


Figure 8. Model comparison results for the fits to data from Wickelgren and Corbett (1977). The first eight pairs of bars refer to different variants of the shifted exponential function involving either one or two intercepts (δ), rates (β), and asymptotes (λ) for the Fan 1 and the Fan 2 conditions. The last pair of bars refers to the ACT-R model. A: Rank statistics (lower is better). B: Model weights (higher is better).

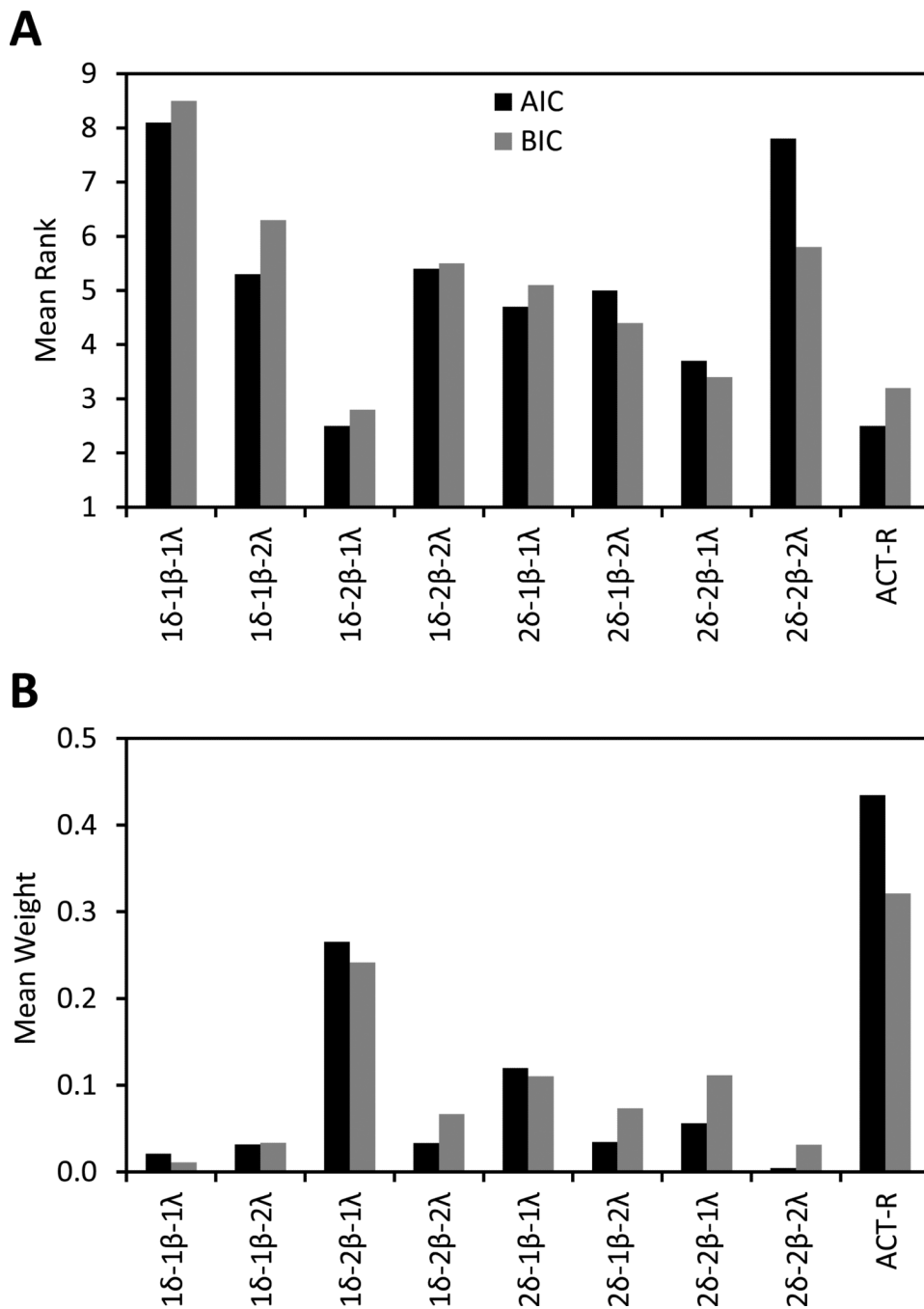


Figure 9. Model comparison results for the fits to data from our experiment. The first eight pairs of bars refer to different variants of the shifted exponential function involving either one or two intercepts (δ), rates (β), and asymptotes (λ) for the Fan 1 and the Fan 2 conditions. The last pair of bars refers to the ACT-R model. A: Rank statistics (lower is better). B: Model weights (higher is better).

Table 1

Mean question frequency in the study phase of our experiment

Question	Session 1		Session 2	Session 3	Session 4	Session 5
	Block 1	Block 2	Block 3			
Fan 1	1.8	1.0	1.0	1.0	1.0	1.0
Fan 2	3.3	1.2	1.1	1.5	1.2	1.0

Table 2

Summary of ACT-R model parameters

Process	Parameter	Value	Description
Encoding	t_{stim}	*	Stimulus encoding time
	t_{signal}	50 ms	Response signal encoding time
Retrieval	W	1.0	Total source activation
	S	1.5	Maximum associative strength
	F	*	Scales retrieval time
	s	*	Controls the variability of activation noise
	τ	* or $-\infty$	Activation threshold
Guessing	t_{guess}	50 ms	Guessing time
	$bias$	0.5	Bias to guess "yes"
Responding	$t_{resp\alpha}$	*	Response execution time

* denotes free parameters

Table 3

ACT-R model parameter values and fit indices for fits to d' accuracy data with a threshold in Wickelgren and Corbett (1977)

Subject	Parameter		Fit index				
	F	s	t_{stim}	τ	AIC	BIC	R^2
PB	589	0.41	211	-0.89	21.6 (18.6)	21.0 (18.0)	.959 (.958)
EC	967	0.25	240	0.43	15.4 (16.0)	14.9 (15.4)	.956 (.955)
SS	731	0.58	295	-0.10	21.9 (21.5)	21.4 (21.0)	.875 (.874)
CL	766	0.36	114	0.30	17.6 (17.5)	17.0 (17.0)	.892 (.893)
CM	1391	0.25	32	0.26	15.9 (16.0)	15.4 (15.4)	.954 (.954)
SJ	910	0.68	176	-0.13	18.7 (17.4)	18.1 (16.8)	.888 (.893)
Mean	892	0.42	178	-0.02	18.5 (17.8)	18.0 (17.3)	.921 (.921)
Group	945	0.42	166	0.06	4.0 (4.3)	3.5 (3.7)	.981 (.981)

Note. "Mean" indicates the mean of the individual-subject values and "Group" indicates the value from a fit to the group data. Values in parentheses are from simulation of the model using the parameters obtained from the analytic fits to the data.

Table 4

Shifted exponential function parameter values and fit indices for fits to d' accuracy data in Wickelgren and Corbett (1977)

Subject	Model	Parameter					Fit index			
		δ_1	δ_2	$1/\beta_1$	$1/\beta_2$	λ_1	λ_2	AIC	BIC	R^2
PB	1 δ -2 β -1 λ	460	460	350	596	5.05	5.05	17.6	17.1	.980
EC	2 δ -1 β -2 λ	543	604	292	292	4.02	2.09	18.5	16.3	.966
SS	2 δ -1 β -2 λ	542	451	364	364	3.90	2.36	11.1	9.0	.961
CL	1 δ -2 β -2 λ	417	417	155	381	3.35	2.36	16.4	14.3	.919
CM	2 δ -1 β -2 λ	441	567	369	369	4.46	2.56	15.7	13.6	.960
SJ	2 δ -2 β -1 λ	496	360	162	654	2.73	2.73	17.3	15.1	.902
Group	1 δ -1 β -2 λ	463	463	356	356	3.92	2.45	3.6	3.0	.986

Note. "Group" indicates the value from a fit to the group data. The parameters are the intercept (δ), rate (β), and asymptote (λ) for the Fan 1 and Fan 2 conditions (denoted by the subscripts 1 and 2, respectively). The parameter values and fit indices are for the model variant with the lowest BIC value.

Table 5

ACT-R model parameter values and fit indices for fits to d' accuracy data without a threshold in our experiment

Subject	Parameter		Fit index			
	<i>F</i>	<i>s</i>	<i>t_{stim}</i>	<i>AIC</i>	<i>BIC</i>	<i>R²</i>
1	268	0.75	407	17.0 (21.1)	17.3 (21.4)	.954 (.946)
2	494	0.54	249	30.2 (34.5)	30.5 (34.8)	.914 (.902)
3	1071	0.55	239	19.4 (19.9)	19.7 (20.3)	.923 (.921)
4	654	0.72	497	14.4 (17.8)	14.7 (18.2)	.952 (.945)
5	997	0.52	450	26.8 (27.1)	27.1 (27.4)	.918 (.916)
6	334	0.68	462	29.1 (29.5)	29.4 (29.8)	.904 (.901)
7	1562	0.42	189	13.5 (13.3)	13.8 (13.6)	.963 (.964)
8	1461	0.54	444	18.4 (20.1)	18.7 (20.4)	.941 (.937)
9	864	0.47	445	11.9 (12.8)	12.2 (13.1)	.977 (.975)
10	705	0.65	479	21.8 (24.9)	22.1 (25.2)	.945 (.934)
Mean	841	0.58	386	20.3 (22.1)	20.6 (22.4)	.939 (.934)
Group	937	0.53	351	5.0 (5.4)	5.3 (5.7)	.992 (.991)

Note. "Mean" indicates the mean of the individual-subject values and "Group" indicates the value from a fit to the group data. Values in parentheses are from simulation of the model using the parameters obtained from the analytic fits to the data.

Table 6

Shifted exponential function parameter values and fit indices for fits to d' accuracy data in our experiment

Subject	Model	Parameter						Fit index			
		δ_1	δ_2	$1/\beta_1$	$1/\beta_2$	λ_1	λ_2	AIC	BIC	R ²	
1	1 δ -2 β -1 λ	521	521	407	667	4.65	4.65	19.8	19.2	.962	
2	1 δ -2 β -1 λ	634	634	125	525	4.64	4.64	23.4	22.9	.979	
3	1 δ -2 β -1 λ	665	665	364	1151	4.06	4.06	13.7	13.2	.978	
4	2 δ -2 β -1 λ	663	763	555	845	4.00	4.00	22.0	19.9	.940	
5	2 δ -1 β -1 λ	668	908	838	838	4.54	4.54	26.2	25.6	.928	
6	2 δ -1 β -1 λ	604	701	532	532	4.84	4.84	23.0	22.4	.966	
7	1 δ -2 β -1 λ	642	642	806	1808	4.80	4.80	16.2	15.6	.968	
8	1 δ -2 β -1 λ	757	757	693	1672	3.96	3.96	19.2	18.6	.950	
9	2 δ -1 β -1 λ	730	958	589	589	4.48	4.48	15.3	14.7	.970	
10	1 δ -2 β -1 λ	717	717	304	913	3.90	3.90	22.6	22.0	.931	
Group	1 δ -2 β -1 λ	639	639	622	1206	4.52	4.52	9.2	8.7	.988	

Note. "Group" indicates the value from a fit to the group data. The parameters are the intercept (δ), rate (β), and asymptote (λ) for the Fan 1 and Fan 2 conditions (denoted by the subscripts 1 and 2, respectively). The parameter values and fit indices are for the model variant with the lowest BIC value.

Table 7

ACT-R model parameter values and fit indices for fits to reaction time data without a threshold in our experiment

Subject	Parameter	Fit index	
	t_{resp}	<i>RMSD</i>	R^2
1	172	3.5 (4.0)	.962 (.950)
2	191	7.0 (7.1)	.880 (.875)
3	207	7.1 (7.3)	.900 (.894)
4	165	5.6 (5.5)	.931 (.933)
5	174	7.6 (7.1)	.814 (.835)
6	180	6.8 (6.6)	.725 (.739)
7	201	6.7 (6.8)	.891 (.888)
8	169	6.9 (6.7)	.759 (.769)
9	205	2.2 (2.5)	.987 (.984)
10	160	4.6 (4.4)	.942 (.948)
Mean	182	5.8 (5.8)	.879 (.881)
Group	182	2.8 (2.9)	.976 (.974)

Note. “Mean” indicates the mean of the individual-subject values and “Group” indicates the value from a fit to the group data. Values in parentheses are from simulation of the model using the parameters obtained from the analytic fits to the data.