Psychological Review

Optimal Metacognitive Control of Memory Recall

Frederick Callaway, Thomas L. Griffiths, Kenneth A. Norman, and Qiong Zhang

Online First Publication, September 21, 2023. https://dx.doi.org/10.1037/rev0000441

AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Review

ISSN: 0033-295X

https://doi.org/10.1037/rev0000441

THEORETICAL NOTE

Optimal Metacognitive Control of Memory Recall

Frederick Callaway[1], Thomas L. Griffiths[1, 2], Kenneth A. Norman[1], and Qiong Zhang[3, 4]
[1] Department of Psychology, Princeton University
[2] Department of Computer Science, Princeton University
[3] Department of Psychology, Rutgers University–New Brunswick
[4] Department of Computer Science, Rutgers University–New Brunswick

Most of us have experienced moments when we could not recall some piece of information but felt that it was just out of reach. Research in metamemory has established that such judgments are often accurate; but what adaptive purpose do they serve? Here, we present an optimal model of how metacognitive monitoring (feeling of knowing) could dynamically inform metacognitive control of memory (the direction of retrieval efforts). In two experiments, we find that, consistent with the optimal model, people report having a stronger memory for targets they are likely to recall and direct their search efforts accordingly, cutting off the search when it is unlikely to succeed and prioritizing the search for stronger memories. Our results suggest that metamemory is indeed adaptive and motivate the development of process-level theories that account for the dynamic interplay between monitoring and control.

Keywords: metamemory, metacognition, resource-rational

Most of us have experienced moments when we could not recall some piece of information but felt that we knew it (feeling of knowing; Hart, 1965), perhaps even sensing that the answer was imminent and only momentarily blocked (tip-of-tongue; Brown & McNeill, 1966). These processes whereby people can examine and make judgments about the content of memory have been termed "metamemory." Different from memory itself, metamemory refers to the higher order processes that monitor and control basic memory processes (T. O. Nelson & Narens, 1990). In this article, we aim to characterize the functional role of these processes in supporting rapid memory recall.

Most empirical work in metamemory has focused on how people are able to monitor their memory states (Eakin, 2005; Miner & Reder, 1994; Reder & Ritter, 1992) and on the accuracy of metamemory judgments in predicting future recall (Dunlosky & Lipko, 2007; Dunlosky & Nelson, 1992; Hart, 1965; Vesonder & Voss, 1985). Recently, these phenomena have been understood through computational models of signal detection (Y. Jang et al., 2012) and probability theory (Hu et al., 2021). Less emphasis, however, has been placed on understanding the function of metamemory judgments (Schwartz & Metcalfe, 2017). In a highly influential article, T. O. Nelson and Narens (1990) proposed that the function of metacognitive systems is to allow effective control of ongoing cognition (Figure 1). For example, they outlined a theory in which a dynamically updated feeling of knowing is used to inform the decision of when to terminate an unsuccessful recall attempt (Figure 5 in T. O. Nelson & Narens, 1990), echoing an earlier proposal that people quickly terminate a memory search when no relevant information is found in an initial search (Glucksberg & McCloskey, 1981). However, despite this early progress, there is (to our knowledge) still no computational model of how these feeling-of-knowing estimates might be dynamically generated nor of how they could be used to control recall efforts. Consequently, despite intuitively suggestive findings such as longer search times for items with high feeling of knowing (Gruneberg et al., 1977; Lachman et al., 1979; T. O. Nelson, 1984; Nhouyvanisvong & Reder, 1998), it is unclear to what extent metamemory serves an adaptive function in people.

We believe two challenges have hindered progress in developing computationally explicit theories of metacognitive control of memory recall. First, on the empirical front, commonly used metamemory paradigms rely on self-report as the primary evidence of people's metamemory. However, the subjective nature of these reports makes it difficult to evaluate the objective utility of metamemory in guiding recall, as we seek to do. Moreover, because the judgments are most often made after retrieval is completed (or abandoned), the causal relationship between metamemory judgments and memory search behavior is unclear (Schwartz, 2001). For example, it is possible that participants report strong feeling of knowing because they spent a long time searching, rather than vice versa. Indeed, in perceptual decision making, manipulating response time (while holding accuracy constant) affects confidence judgments (Kiani et al., 2014). On the other hand, rapid feeling-of-knowing judgments made before recall (e.g., Reder, 1987) cannot capture knowledge that only becomes available in the course of recall (Koriat, 1993; Nhouyvanisvong & Reder, 1998). To address

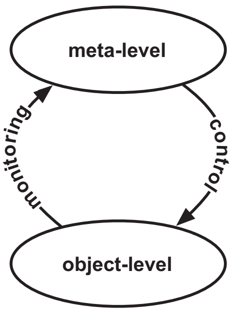
Frederick Callaway https://orcid.org/0000-0001-7687-5987

All data and code supporting this article can be found at https://github.com/fredcallaway/memory. The studies' design and analysis were preregistered at https://aspredicted.org/wr9ej.pdf and https://aspredicted.org/xq9nx.pdf. This research was supported by a grant from Facebook Reality Labs to Thomas L. Griffiths.

Correspondence concerning this article should be addressed to Frederick Callaway, Department of Psychology, Princeton University, Peretsman Scully Hall, Princeton, NJ 08540, United States. Email: fredcallaway@gmail.com

Figure 1

Illustration of Nelson and Narens' Theoretical Framework for Metamemory: A "Meta-Level" Process Monitors and Controls the Performance of a Basic "Object-Level" Process



Note. Adapted from "Metamemory: A theoretical framework and new findings," by T. O. Nelson and L. Narens, in *Psychology of learning and motivation* (Vol. 26, pp. 125–173), 1990, Academic Press. Copyright 1990 by the Academic Press. Adapted with permission.

this challenge, we developed a metamemory paradigm that allows us to establish a quantitative, objective measurement of memory strength before retrieval. An extension of this paradigm in Experiment 2 additionally allows us to see behavioral signatures of metacognitive control even before retrieval is completed or abandoned, revealing how the dynamic metamemory process unfolds over time. In this way, we can directly test our model's core predictions about how people will direct their recall efforts depending on the strength of the to-be-recalled memories.

The second challenge is a technical one. In many domains of cognitive science, theoretical progress has been spurred by the development of rational models that optimally solve the problem that the cognitive system is theorized to solve (Anderson, 1991; Knill & Richards, 1996; Marr, 1982; Savage, 1954; Tenenbaum & Griffiths, 2001). Indeed, Anderson and Milson (1989) famously applied this approach to shed light on basic properties of human memory. Metamemory, however, poses an especially thorny type of optimization problem, as it involves a cyclic, "closed-loop" interaction between two cognitive processes (Figure 1). It is not obvious how one should quantify the performance of such a system, let alone identify a system that maximizes this performance. To address this challenge, we draw on formal tools developed for meta-level control in artificial intelligence (Hay, 2016; Russell & Wefald, 1991). These tools have recently been applied to model dynamic metacognitive processes in decision-making contexts, revealing that people's behavior is remarkably consistent with models that optimally trade off utility with cognitive cost (Callaway et al., 2021; Callaway, van Opheusden, et al., 2022; cf. Chen et al., 2021; Drugowitsch et al., 2012; A. I. Jang et al., 2021; Tajima et al., 2019). By applying these tools to a simple model of memory recall, we can make concrete predictions about the behavior we would expect to see if people can adaptively control their memory processes.

The remainder of this article is organized as follows. We begin by reviewing empirical work on meta-level control of memory, focusing on the control of recall. Then, we define an optimal model of meta-level control in memory recall and characterize its predictions. Notably, the model predicts that unsuccessful memory searches will be longer when the target memory is (judged to be) stronger, consistent with the findings of Costermans et al. (1992). Next, we describe a cued-recall experiment that conceptually replicates and extends those findings. We confirm all key qualitative predictions of the model and establish moderate quantitative fit. Our second experiment extends the first by allowing participants to choose between two possible recall targets. This introduces a more complex meta-level control problem of selecting which memory to search for at each moment. Using a keypress-contingent display, we compare the time course of attention to each cue with the optimal model's search predictions and again achieve a strong qualitative and moderate quantitative fit. We conclude by discussing implications of the results for metamemory and metacognition research more generally and identifying interesting directions for further research.

Empirical Evidence for Control in Memory Recall

A number of behavioral studies have already suggested that people are capable of using their ability to monitor their memory in the service of controlling their memory processes. At the acquisition phase, a large body of work has investigated how people preferentially allocate study time depending on how well they have learned different pieces of information (Dunlosky & Hertzog, 1998; Gureckis & Markant, 2012; Metcalfe, 2009). Another substantial literature has addressed how people choose which memories to maintain or forget (Castel, 2007; Hu et al., 2019; Suchow & Griffiths, 2016; M. Williams et al., 2013). Here, however, we focus specifically on the control of recall.

Most work on metamemory for recall boils down to one essential question: How do people decide whether to (continue to) search for a memory? The initial decision of whether or not to search at all is often treated as part of a more general *strategy selection* process (Reder, 1988), with memory search being one of multiple possible strategies (along with, e.g., looking up the information in a dictionary). The choice of strategy appears to be driven by an initial feeling-of-knowing (Nhouyvanisvong & Reder, 1998), which is itself driven by surface-level properties of the question, such as familiarity with its terms (Reder & Ritter, 1992). A key finding from this line of work is that people can estimate the probability that they will be able to recall a target faster than they can actually recall it (Reder, 1987). This necessitates some form of metacognitive monitoring, as participants cannot be making judgments of recallability based on the outcome of recall if the former precedes the latter.

Once a memory search has been initiated, how long do people search before giving up? A key finding here is that participants spend longer before giving up on questions for which they have relevant information stored in memory (Glucksberg & McCloskey, 1981; Lachman et al., 1979), as well as those for which they report greater feeling of knowing (Gruneberg et al., 1977; Nhouyvanisvong & Reder, 1998) or being in a tip-of-the-tongue state (Schwartz, 2001). Feeling of knowing and tip-of-the-tongue states are themselves associated with greater subjective familiarity (Reder, 1988), partial recall of the target (Brown & McNeill, 1966; Koriat, 1993; Schacter & Worling, 1985), and the ability to recall given additional information (Gruneberg & Monks, 1974). Together, these results suggest that people are able to accurately identify targets they are likely to recall with further effort and allocate that effort accordingly.

A related finding, although not one central to control, is that participants give higher confidence judgments when they recall an answer more quickly (T. O. Nelson & Narens, 1990). In conjunction with the feeling-of-knowing effects, this produces a striking pattern.

Treating both feeling of knowing and confidence as judgments of memory strength, we see opposite relationships between judged strength and response time for successful versus failed recall. Costermans et al. (1992) demonstrated this pattern in a single study. On each trial, participants were given a general knowledge question. Then, if they provided an answer, they gave a confidence judgment; if they were unable to provide an answer, they instead gave a feeling-of-knowing judgment. Costermans et al. found that, on the recall trials, participants gave higher confidence judgments when they responded more quickly. But on the omission trials, participants gave higher feeling-of-knowing judgments when they responded more *slowly*. In the following section, we will show that both of these findings are consistent with a model in which memory recall follows an evidence accumulation process and search is terminated optimally based on metacognitive monitoring of the rate of progress.

An Optimal Model of Metamemory for Recall

Following classic theories of metamemory (T. O. Nelson & Narens, 1990; Figure 1), we specify our model as two interrelated processes operating at different levels. The *object-level* process includes the mechanisms supporting recall itself. Here, we abstract away from the details of memory search, modeling recall instead as a simple evidence accumulation process (Ratcliff & Tuerlinckx, 2002; Sederberg et al., 2008). The *meta-level* process supervises the object-level process; it *monitors* the rate of progress toward recall and *controls* how long the search process is allowed to continue. Here, we assume that the meta-level process is optimal in the sense that it terminates search when the expected costs of search outweigh the expected benefits.

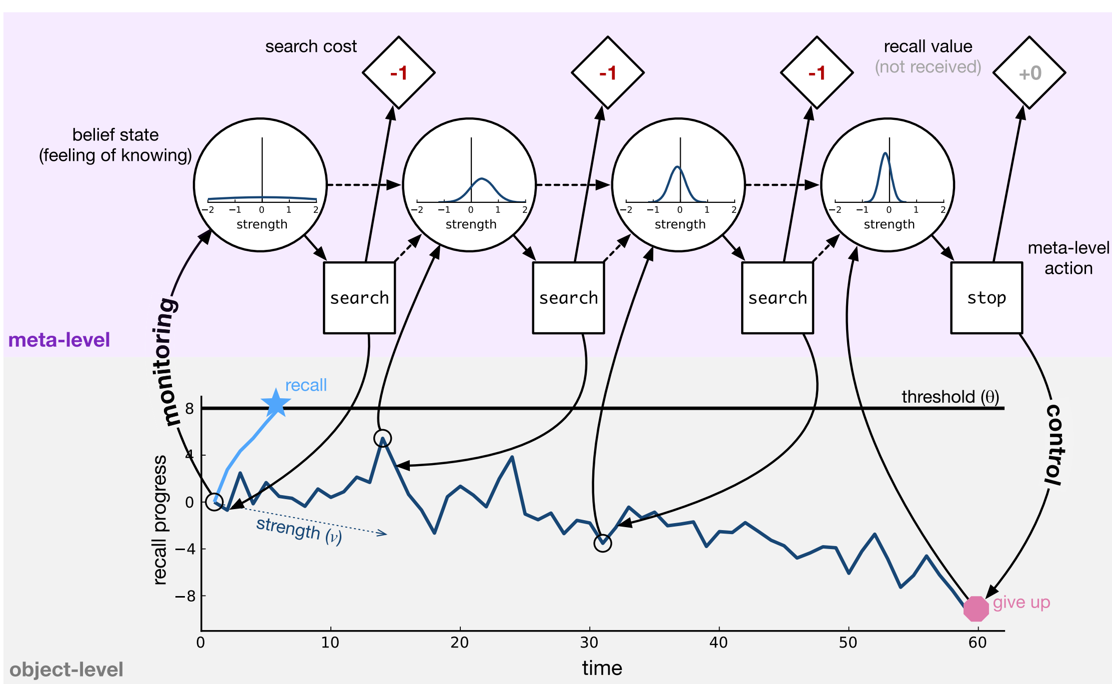
Importantly, we do not intend this model as a precise characterization of the mechanisms underlying human metamemory nor do we hope to achieve a close quantitative fit to data (although we do fit the model using maximum likelihood estimation). Instead, our goals are to (a) characterize the problem that a metamemory system must solve and (b) identify qualitative behavioral signatures of a system that solves this problem well. This will allow us to determine in which ways people can or cannot deploy metamemory adaptively, potentially generating clues about the nature of the true underlying cognitive processes. The model is illustrated in Figure 2; we describe its components below.

Object-Level Process

We model recall as a process of evidence accumulation. Evidence accumulation (or "sequential sampling") models assume that decisions are made by accumulating noisy information over time until a threshold level of evidence is reached. They have been widely applied in the decision-making (Busemeyer & Townsend, 1993;

Figure 2

A Dynamic Model of Metamemory



Note. Bottom: The object-level recall process is modeled as evidence accumulation. A memory is recalled when a threshold level of evidence is accumulated (blue star). The average rate of accumulation corresponds to the strength of the memory. Top: The meta-level process monitors and controls the object-level recall process. This is modeled as a Markov decision process (MDP) where the states (circles) correspond to beliefs about the memory's strength and the actions (squares) determine whether the search continues; the rewards (diamonds) capture the cost of the search and the utility of recalling a memory. Solving this MDP yields an optimal policy for determining when to stop searching memory based on partial recall progress (pink octagon). See the online article for the color version of this figure.

Ditterich, 2006; Krajbich et al., 2010; Usher & McClelland, 2001) and memory (Ratcliff, 1978; Sederberg et al., 2008) literatures and are successful in accounting for the effects of various experimental manipulations on accuracy and response times during recognition and recall tasks (Ratcliff & Tuerlinckx, 2002; Sederberg et al., 2008; Yonelinas et al., 2010). In our model, the "evidence" captures progress toward recalling a target. Thus, when a threshold level of evidence is reached, the target is recalled (blue star in Figure 2).

Concretely, at each time point t, the current recall progress z_t is incremented by a sample from a Gaussian distribution,[1]

$$z_t = z_{t-1} + x_t \text{ where } x_t \sim \mathcal{N}(v, \sigma_x^2). \tag{1}$$

The mean of this distribution, v, controls the rate of accumulation; it is often called the *drift rate* (illustrated as a thin dashed blue arrow in Figure 2). In our model, it captures the strength of the memory. The noise σ_x^2 captures the consistency of that progress. The target is recalled when the total progress exceeds a threshold θ.

Meta-Level Process

The problem of deciding when to cut off an unsuccessful memory search is addressed by the meta-level process. That is, the meta-level process *controls* how long the object-level process is allowed to continue. How should it do so? From a rational perspective, one should keep searching as long as the probability of recall multiplied by the utility of recall is greater than the expected cost of search (Anderson & Milson, 1989). Putting this logic into notation, we can define the optimal meta-level action as

$$a^* = \begin{cases} \text{SEARCH} & \text{if } p(\text{recall}) \cdot U(\text{recall}) > E[\text{cost}(\text{search})] \\ \text{STOP} & \text{otherwise} \end{cases}, \tag{2}$$

where U stands for utility. The challenge lies in estimating p(recall) and E[cost(search)]. In our evidence accumulation model, these values correspond respectively to the probability that the evidence will eventually cross the threshold and the time point at which this occurs.

Intuitively, one could accurately estimate the probability and cost of future recall if one knew the strength of the target memory, v. However, a key assumption of our model—and the metamemory literature more broadly—is that the meta-level process does not have direct access to this information. Instead, we assume that the meta-level process must infer the memory's strength by *monitoring* the object-level process. The existence of such a monitoring process is widely agreed on; however, its precise nature is controversial. In particular, it is unclear to what extent monitoring tracks the underlying memory strength (Hart, 1965), partial recall progress (Koriat, 1993), or superficial cues that happen to be predictive of recall (Reder & Ritter, 1992; Schwartz & Metcalfe, 1992). Resolving this debate is beyond the scope of this article. Thus, for simplicity and tractability, we assume that the meta-level process directly observes the state of the object-level process. We emphasize that this is purely a simplifying assumption and not a claim about how people actually monitor their memory. We return to this point in the discussion.

Concretely, we assume that the meta-level process observes the current recall progress z_t and the time spent so far t, which provides a complete summary statistic for the entire sequence up to time t. Given this information, the meta-level process then infers a posterior distribution over the strength of the memory,

$$p(v|t, z_t) = \mathcal{N}(v; \mu_t, \sigma_t^2)$$
$$\mu_t = \frac{z_t\sigma_x^{-2} + \mu_0\sigma_0^{-2}}{\sigma_t^{-2}} \qquad \sigma_t^2 = \frac{1}{t\sigma_x^{-2} + \sigma_0^{-2}}, \tag{3}$$

where μ_0 and σ_0^2 encode the agent's prior, $\mathcal{N}(\mu_0, \sigma_0^2)$. To build intuition, note that with a very weak prior (large σ_0^2), μ_t reduces to z_t/t, the average rate of recall progress. This time-varying belief about the strength of a memory formalizes the concept of *feeling of knowing*.

Given this estimate of memory strength, how should the meta-level process determine whether to continue searching? That is, how should *monitoring* inform *control*? One intuitively appealing strategy is to estimate future recall progress by repeatedly applying Equation 1, compute p(recall) and E[cost(search)] from those estimates, and then choose the optimal action by Equation 2. However, this strategy fails because it implicitly assumes that the object-level process will be allowed to continue indefinitely—precisely what the meta-level process is intended to prevent. The decision of whether to stop is not made just once; it is continually remade at each time step. As a result, the probability and cost of recall—and therefore the optimal stopping decision—depend on one's future stopping decisions.

Thus, we see that metamemory poses a particularly thorny type of decision problem, one where one's current choice depends on one's future choices. That is, metamemory poses a *sequential decision problem*. Fortunately, a great deal of work in artificial intelligence has focused on solving exactly this sort of problem, typically using the formalism of *Markov decision processes* (MDPs). An MDP is defined by a set of states the environment can be in S, a set of actions the agent can take A, a transition function specifying how actions change state T, and a reward function specifying the agent's goals r. MDPs are the key formalism underlying reinforcement learning (RL; Sutton & Barto, 2018), supporting recent advances in artificial intelligence (Mnih et al., 2015; Silver et al., 2017), as well as providing a foundation for the psychology and neuroscience of decision making (Dayan & Daw, 2008; Glimcher, 2011; Niv, 2009).

The insight that meta-level control poses a sequential decision problem has been formalized by the field of rational metareasoning (Russell & Wefald, 1991), which has the goal of building artificial intelligence that makes efficient use of their limited computational hardware. In particular, we apply the framework of *meta-level Markov decision processes* (Hay et al., 2012), which models the meta-level control problem as an MDP. In a meta-level MDP, the states correspond to beliefs about the world and the actions correspond to computations (or cognitive operations) the agent can execute. The transition function describes how computations update beliefs, and the reward function encodes the cost of computation as well as the benefits of acting according to a more refined belief.

In our meta-level MDP model, a state $s \in S$ captures both the current recall progress as well as the belief about memory strength (feeling of knowing). Because the belief depends only on the recall progress and time spent so far (Equation 3), the state can be

[1] The exact choice of distribution is arbitrary. We use a Gaussian for mathematical convenience. Anecdotally, we found similar qualitative predictions with a Bernoulli distribution, but this model had a worse quantitative fit.

compactly represented as $s_t = (t, z_t)$. There are two possible actions $a \in A$: SEARCH continues searching for the memory and STOP terminates recall.

The reward function r encodes the benefit of recall and the cost of search:

$$r(s_t, a_t) = \begin{cases} U(\text{recall}) & \text{if } z_t \geq \theta \\ -\gamma_{\text{SEARCH}} & \text{if } a_t = \text{SEARCH} \\ 0 & \text{if } a_t = \text{STOP}, \end{cases} \tag{4}$$

where U(recall) specifies the utility of correct recall (capturing, in our case, experimental incentives) and γ_{SEARCH} is a free parameter that specifies the cost of searching for one-time step (capturing any experimentally imposed costs as well as implicit costs such as the opportunity cost of the time spent searching).

Finally, the transition function T captures the evidence accumulation dynamics of the object-level process and the fact that STOP terminates search. Note, however, that the object-level dynamics depend on the true memory strength (v in Equation 1), which the agent does not have access to. Thus, the transition function must marginalize over the strength according to the current belief state:

$$\begin{aligned} T(s_{t+1}|s_t, a) &= p(z_{t+1}|t, z_t) \\ &= \int p(z_{t+1}|z_t, v)p(v|t, z_t)dv \\ &= \int \mathcal{N}(z_{t+1} - z_t|v, \sigma_x^2)\mathcal{N}(v|\mu_t, \sigma_t^2)dv \\ &= \mathcal{N}(z_{t+1} - z_t|\mu_t, \sigma_x^2 + \sigma_t^2). \end{aligned} \tag{5}$$

The two substitutions in the third line follow from Equations 1 and 3, respectively. The final line is a standard property of Gaussian distributions (Murphy, 2007).

Optimal Policy

We have now defined a meta-level MDP that formalizes the problem of deciding when to give up on recalling a memory. The final step is to specify a strategy for solving that problem. In MDP terms, we need to specify a policy π that chooses which action to take in each state. Here, we focus on the optimal policy, which is the one that maximizes the total expected reward. As detailed in the methods, we can identify this policy by computing the optimal value function V^*, which specifies the maximal total reward one could expect to gain starting from any given state. To build intuition, we can factorize V^* for the current model into two components, capturing the utility of recall and the cost of search,

$$V^*(s_t) = p(\text{recall} \mid s_t)U(\text{recall}) - (\text{E}[t_{\max} \mid s_t] - t)\gamma_{\text{SEARCH}}, \tag{6}$$

where $t_{\max}$ is the time step on which the item is recalled or the search is terminated. The optimal policy is then defined as

$$\pi^*(s_t) = \begin{cases} \text{SEARCH} & \text{if } \text{E}_{s_{t+1} \sim T(\cdot|s_t, \text{SEARCH})}[V^*(s_{t+1})] > \gamma_{\text{SEARCH}} \\ \text{STOP} & \text{otherwise} \end{cases}. \tag{7}$$

To understand this equation in comparison to Equation 2, note that p(recall) and cost(search) have each been split into two components, capturing immediate versus future outcomes. The immediate recall probability is encoded in the transition function, $T(\cdot \mid s_t, \text{SEARCH})$; the immediate search cost is encoded in the reward, $-\gamma_{\text{SEARCH}}$. The expected future outcomes are both integrated into $V^*(s_t + 1)$, as shown in Equation 6. A key advantage of specifying our model as an MDP is that we can apply standard techniques (in particular, backward induction) to compute V^*, allowing us to identify an optimal policy for metamemory. See the Method section below for details.

Predictions

As illustrated in Figure 3, the model makes two key predictions regarding the relationship between memory strength and response time. First, stronger memories should be recalled more quickly because stronger memories accumulate progress faster and hit the threshold sooner. Note that this prediction is a simple consequence of the object-level process and does not depend on metacognition. Second, stronger memories should be abandoned less quickly. In particular, while the meta-level process can quickly identify very weak memories as such (leading it to terminate the search), marginal-strength memories produce ambiguous evidence and it takes more time for the meta-level process to determine that the memory is too weak to justify further search.

Figure 3 also highlights that the optimal policy can be represented as a time-varying threshold, such that the search is terminated if the progress ever falls below the threshold (cf. Drugowitsch et al., 2012). In the language of Marr (1982), this can be understood as an algorithmic-level implementation of the computational-level theory outlined above. We return to this point in the General Discussion section. Note that the threshold is nonmonotonic because a fixed amount of negative progress provides stronger evidence that the memory has low strength if the negative progress was generated more quickly.

Experiment 1

In our first experiment, we sought to replicate and extend the findings of Costermans et al. (1992) in a cued-recall setting (Figure 4). The key finding from the original study was that participants reported higher confidence judgments on trials where they more quickly recalled the answer to a question but lower feeling-of-knowing judgments when they more quickly reported being unable to recall the answer. Our model can capture both of these effects under the assumption that the metamemory judgments are based on the inferred memory strength at response time (explained below). However, it is also possible that the metamemory judgments reflect a purely post hoc rationalization of the longer response time, not influencing the decision to stop at all. To avoid this reverse-causality concern, we modified the task such that we could obtain objective measures of memory strength before the critical trials. Specifically, we used a cued-recall paradigm that allowed us to query the same target multiple times. This allowed us to test whether people's stopping decisions depended on their true memory strength, as the optimal policy predicts they should.

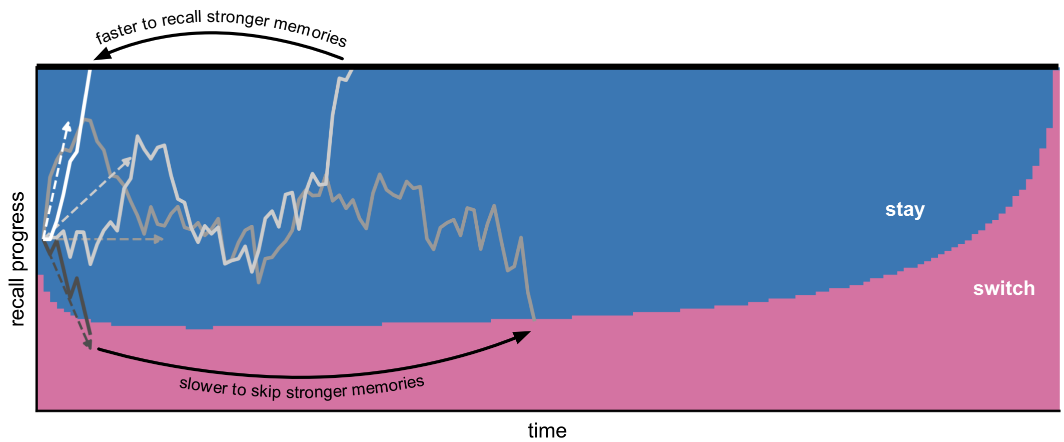
Method

Participants

We recruited 612 participants through Prolific with the restriction that they reported current U.S. or U.K. residence, had at least a 95%

Figure 3

Experiment 1: Optimal Policy and Predictions



Note. The optimal policy partitions the state space into two sections, one (blue) in which the policy continues searching and another (pink) in which it terminates search. The response time for each trial is thus given by the first time point at which the recall progress either exceeds the threshold (recall) or enters the pink region (no recall). In the former case, stronger memories (lighter lines) will result in faster responses because such memories accumulate progress and hit the threshold faster. In the latter case, stronger memories will result in slower responses because such memories can hover in the search region before ultimately hitting the stop region. This plot was generated with parameters fit to the data in Experiment 1. See the online article for the color version of this figure.

approval rating, and had not participated in any pilot studies. As preregistered, we excluded 106 (17%) participants who did not provide a response on more than 90% of critical trials.[2] This yielded 506 participants in our final analysis. The target sample size of 500 participants had over 95% power with $\alpha = .05$ for all our preregistered hypotheses based on a boot-strapping power analysis conducted on pilot data. This experiment was approved by the institutional review board of Princeton University (Protocol Number 10859).

Stimuli

Each participant was randomly assigned 40 images and words, which were arbitrarily paired. The images were randomly sampled from 40 common scene categories (one image per category), selected from the Scene UNderstanding database (Xiao et al., 2010). We manually removed photos that contained a person. All images were resized and then cropped to 300 by 300 pixels. The words were selected randomly from those used in Madan (2021), which were themselves selected from the University of South Florida free association norms word database (D. L. Nelson et al., 2004).

Procedure

The experiment consisted of four phases: exposure, distractor, pretest, and critical. After learning the mapping between images and words through a single round of passive exposure, participants solved simple arithmetic problems to clear working memory. They then completed the pretest and critical trials, both of which involved cued recall. In the pretest trials, participants were given an image and asked to type in the corresponding word; they were incentivized to be both accurate and fast. These trials provide an objective measure of how well each participant had learned each pair. In the critical trials, we increased the speed incentive and added an error penalty. However, we also allowed participants to skip a trial without penalty, still earning the speed bonus. This creates an incentive to quickly identify trials in which the target is unlikely to be correctly recalled. We provide further details on the procedure for each of these components below.

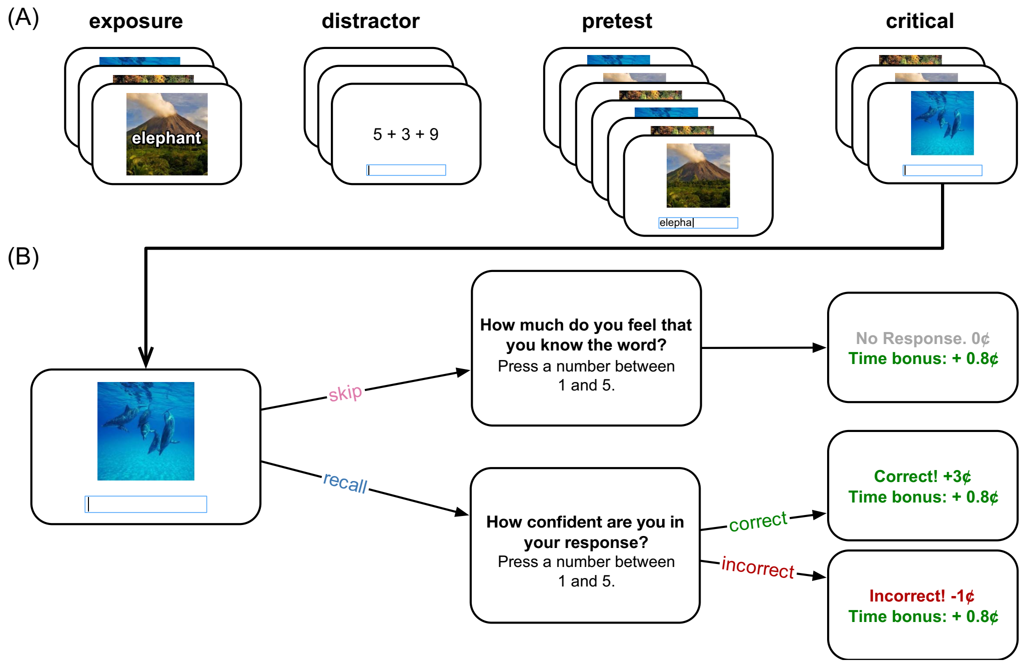
Exposure. On each exposure trial, participants viewed a word superimposed on the center of an image; the word was printed in white font with black outlines such that it would be clearly legible on any image. The word–image pair was shown for 2 s, with a half-second intertrial interval. Each of the 40 pairs was shown once.

Distractor. On each distractor trial, a simple arithmetic problem was presented and participants had 3 s to enter the correct answer. Each problem was an addition of three single-digit numbers. After a response (or time-out), feedback was presented for at least 1 s. If a response was made before the deadline, the feedback phase was extended such that all trials lasted exactly 4 s. Participants were informed that they would earn one cent for each correct answer. Due to a programming error, participants were incorrectly instructed that they would have 5 s to enter a response; however, no participant reported noticing this discrepancy in the debriefing survey.

Pretest. Each pretest trial began with a blank screen and the text “press space when ready.” When the participant pressed space, an image, text box, and timer appeared. The timer immediately began counting down from 15 s. The trial ended when the participant entered a word (by typing it into the text box and pressing enter) or when the timer expired. If the timer expired, “time-out” appeared in large red letters. No other trial-by-trial feedback was provided.

[2] This slightly high exclusion rate may reflect the difficulty of learning from a single round of exposure. Pilot studies indicated that longer learning phases resulted in very few skips.

Figure 4
Experiment 1: Procedure



Note. (A) Participants viewed 40 image–word pairs for 2 s each (exposure). They then completed simple math problems for 60 s (distractor). Next, they attempted to recall the word associated with each image, two trials per image (pretest). Finally, they completed one critical trial for each image. (B) Critical trials were similar to pretest trials, except that incorrect responses were penalized. The penalty could be avoided by providing an empty response, "skipping" the trial. After giving a response, participants made a metamemory judgment: confidence if they had entered a word, feeling of knowing if they had not. A speed bonus was given regardless of the response. See the online article for the color version of this figure.

Participants were instructed that they would receive one cent for each correct answer, as well as a small extra bonus for answering quickly and correctly. (The time bonus was a quarter of a cent multiplied by the proportion of time left when a response was given.) At the end of each block, participants received a summary of their performance, separately indicating the amount of bonus money they made from correct responses and response speed. There were two blocks, and each image–word pair was shown once in each block, for a total of 80 trials (two presentations of each pair). The first trial in the first block was a practice trial and was excluded from the analysis.

Critical Trials. The critical trials were similar to the pretest trials but with a different incentive scheme. The bonus for correct responses was increased to three cents, but a one-cent penalty for incorrect responses was introduced. Additionally, participants could skip a trial by pressing enter without typing a word; this did not incur a penalty. Finally, the speed incentive was raised to a tenth of a cent for each second left on the timer (i.e., up to 1.5 cents per trial). The speed bonus was given on all trials, including skip and error trials. To ensure that participants understood the incentives, they were required to pass a quiz, affirming that there was a penalty for mistakes, no penalty for skipping, and a time bonus regardless of response type. Participants were additionally encouraged to quickly skip trials for which they did not know the word.

After a response was given, a metamemory judgment was elicited. When participants gave a response, they were asked "How confident are you in your response?" They then pressed a number between 1 and 5 to indicate that they were "not at all sure," "not so sure," "more or less sure," "nearly sure," or "absolutely sure" that their response was correct. If they did not give a response (i.e., they skipped the trial), they were asked "How much do you feel that you know the word?" again pressing a number between 1 and 5. The responses were described as "I am absolutely sure I do not know the word," "I am rather sure I do not know the word," "I have a vague impression I know the word," "I am rather sure I know the word," and "I am absolutely sure I know the word."

Each image–word pair was shown once. The first three trials were practice trials that did not count toward the participant's bonus and were not analyzed. This leaves 37 analyzed critical trials per participant.

Modeling

Computing the Optimal Policy. We compute the optimal policy by backward induction. See Puterman (2014, p. 92) for a general description of the method; here, we report the details necessary to apply the method to our model.

Recall that a belief state in the model is a tuple (t, z_t). Because backward induction can only be applied in finite state spaces, we begin by discretizing the progress dimension of the belief into 100 equally sized bins, ranging from $-\theta$ to θ. Note that θ is the maximum possible value z_t can take. The lower bound of $-\theta$ is an arbitrary choice; we found that the optimal policy for well-fitting parameters always terminated well before this value was reached (e.g., Figure 3), suggesting that this imposed lower bound did not meaningfully affect the solution.

We first computed the transition function. To account for the discretization, we computed the probability of transitioning from (t, z_t) to $(t + 1, z_t + 1)$ as $\Pr(b^{\text{bot}} < z_{t+1} < b^{\text{top}} | t, z_t)$ where b^{bot} and b^{top} are the boundaries of the bin with $z_t + 1$ in the center. Because $z_{t+1} | t, z_t$ is Gaussian (Equation 5), we could compute this quantity with standard statistical library functions (the normal cumulative distribution function). For most bins, the boundaries were $z_{t+1} \pm \theta/100$. The top bin was clipped at $b^{\text{top}} = \theta$, and the bottom bin was unbounded, with $b^{\text{bot}} = -\infty$. This ensures that the transition probability from each state sums to one.

Next, we initialized the value function for all terminal belief states. The value of states with $z_t = \theta$ is U(recall) and the value of states with $t = 150$ (the maximum trial duration) but $z_t < \theta$ is 0. Then, we iterated backward in time, computing the value of all states with $t = 149$ as the maximum of the expected value of each possible action

$$V^*(s) = \max_{a \in \{\text{SEARCH, STOP}\}} Q^*(s, a), \tag{8}$$

where

$$Q^*(s, \text{SEARCH}) = \sum_{z_{t+1}} p(z_{t+1} | t, z_t) \, V^*(t + 1, z_{t+1}) - \gamma_{\text{SEARCH}}, \tag{9}$$

and $Q^*(s, \text{STOP}) = 0$. The iteration continues with $t = 148$ down to $t = 1$. After computing Q^* for all states and actions, the optimal action in each state can be quickly identified as

$$\pi^*(s) = \operatorname*{argmax}_{a \in \{\text{SEARCH, STOP}\}} Q^*(s, a). \tag{10}$$

Simulation Procedure. In order to compare the behavior of the model with that of our participants, we simulated experimental data. Simulating a trial corresponds to executing one "episode" of the meta-level MDP. That is, we initialized the state at $s_0 = (t = 0, z_t = 0)$ and then repeatedly applied Equation 1 to generate a sequence of states.[3] At each time step, we first checked if the recall threshold has been exceeded, that is, if $z_t < \theta$. If so, the episode ended and the trial was classified as a recall trial. Otherwise, we determined the optimal action $\pi^*(s_t)$, defined in Equation 7. If the optimal action was STOP, then the episode ended and the trial was classified as a skip trial. Otherwise, we repeated the process, unless the maximum time step had been reached, in which case the episode ended as a skip trial.

The simulated response time was determined based on the final value of t. We assumed that response times reflected both time spent actively searching memory as well as "nondecision time" (NDT) spent on, for example, perceptually encoding the cue and preparing the motor response. For search time, we assumed that each time step took a fixed amount of time, a value we arbitrarily set to 100 ms (the predictions of the model do not depend critically on this parameter; we chose 100 ms to balance prediction fidelity with model runtime). For NDT, we assumed that it was drawn separately for each trial from a gamma distribution, with parameters fit to data as described below. The simulated response time was the sum of the two components. Note that, for computational reasons, we did not factor the NDT into the time-out condition (the maximum time step of 150 is the maximum trial duration of 15 s divided by 100 ms). This has a negligible effect on model predictions because time-outs were rare (less than 0.01% of trials) with well-fitting parameter values.

The simulated metamemory judgments (confidence and feeling of knowing) were determined based on the posterior mean μ_t at the final time step, that is, when the target was recalled or the policy terminated the search. To account for factors contributing to the judgment besides those captured by our model (e.g., individual differences in scale usage), we first corrupted μ_t with Gaussian noise, arbitrarily setting the variance to $\sigma_x/2$. We then binned the continuous measure into five bins, corresponding to the 1–5 response scale. We set the bin boundaries separately for each judgment type in order to match the proportion of each response in the human data.

In order to capture the relationship between performance in the pretest and critical trials, we simulated both phases using the following procedure. For each simulated word/image pair, we sampled its memory strength v from the prior distribution $\mathcal{N}(\mu_0, \sigma_0^2)$. The parameters of the prior are free parameters of the model. Next, we simulated the two pretest trials for that pair by rolling out two episodes of the meta-level MDP. For these trials, we set U(recall) to the experimentally imposed value of one cent. The search cost γ_{SEARCH} is a free parameter. We then simulated the critical trial for the pair, setting U(recall) to the new value of three cents and increasing γ_{SEARCH} by the experimentally imposed value of 0.01 cents per sample (0.1 cents per second and 100 ms per sample).

Parameter Estimation. The model's behavior is governed by six free parameters: the prior mean and standard deviation, μ_0 and σ_0, the progress noise σ_x, the search cost γ_{SEARCH}, and the mean and shape of the NDT distribution, μ_{NDT} and α_{NDT}.[4] We arbitrarily fixed the threshold $\theta = 1$ as it is not identifiable along with the other parameters. We set these parameters by maximizing the likelihood of the critical trials at the group level. For fitting, we disregarded the metamemory judgment. Each trial (human or simulated) was thus defined by a pretest accuracy rate (0%, 50%, or 100%), a response type (skip or recall), and a response time (discretized into 100 ms bins from 0 to 15,000 ms).

Because the optimal policy does not depend on the NDT parameters, we treated these separately (described below). For the remaining four parameters, we considered 50,000 configurations sampled pseudorandomly according to the Sobol sequence (Bergstra & Bengio, 2012; Sobol', 1967) within the range, $\mu \in (-0.5, 0.5)$, $\sigma_0 \in (0, 1)$, $\sigma_x \in (0, 1)$, $\gamma_{\text{SEARCH}} \in (0, 0.05)$. For each configuration, we computed the optimal policy by backward induction and then simulated 100,000 critical trials. For each simulated data set, we constructed a $3 \times 2 \times 151$ histogram over possible trials (three accuracy rates, two response types, and 151 response time bins). To apply the NDT model, we convolved this histogram with a Gamma distribution parameterized by μ_{NDT} and α_{NDT}. Finally, to ensure nonzero probability was assigned to all trials, we mixed the model-predicted distribution with a uniform distribution with weight 10^{-6}.

[3] Note that we simulate data conditional on a known strength v rather than with the transition function that marginalizes over v. This allows us to model multiple trials for a single cue/target pair, as described below.

[4] We use this parameterization rather than the traditional shape-scale parameterization to aid interpretability. The mean is the scale parameter multiplied by the shape parameter.

The likelihood of each trial in the human data set is simply the corresponding entry of this histogram. The total log-likelihood is the sum of the log-likelihood for each trial.

Note that the NDT-convolution step does not require simulating the model. We thus optimized the NDT parameters for each configuration of the other four parameters using the Nelder–Mead algorithm (Nelder & Mead, 1965). This provides the best possible likelihood for each configuration of the four primary parameters. We then selected the top 5,000 such configurations and approximated the likelihood more precisely, using 1,000,000 simulated trials. The best-performing configuration from this smaller set was then identified as the maximum likelihood estimate (MLE). The MLE was $\mu_0 = -0.002$, $\sigma_0 = 0.186$, $\sigma_x = 0.139$, $\gamma_{\text{SEARCH}} = 0.014$, $\mu_{\text{NDT}} = 717$, $\alpha_{\text{NDT}} = 8.70$ with negative log-likelihood 69,698. Note, however, that this exact number is arbitrary because it depends on the coarseness of the response time discretization.

Lesioned Model Without Meta-Level Control. In order to distinguish between model predictions that depend only on the object-level process versus those that reflect adaptive meta-level control, we implemented a lesioned version of the model that lacks the control component. In this model, the decision to stop searching is no longer made by the optimal policy; instead, it is made randomly. Concretely, when simulating a trial from this model, we begin by sampling the stopping time from a Gamma distribution. We then apply Equation 1 until either (a) the threshold is crossed, resulting in a correct trial as in the optimal model, or (b) the predetermined stopping time is reached, resulting in a skip trial. This model has all the parameters of the main model except γ_{SEARCH}, which only influences the optimal policy. It has two additional parameters for the stopping time distribution.

The parameters are fit by maximum likelihood estimation using the procedure described above. The MLE was $\mu_0 = 0.278$, $\sigma_0 = 0.156$, $\sigma_x = 0.019$, $\mu_{\text{stop}} = 250$, $\alpha_{\text{stop}} = 56.56$, $\mu_{\text{NDT}} = 1{,}261$, $\alpha_{\text{NDT}} = 2.17$ with negative log-likelihood 73,840. We also considered a version of the lesioned model that sampled stopping times directly from the empirical distribution. This model fit the data worse (negative log-likelihood of 76,485; see Appendix B).

Statistical Analyses

All reported regressions are linear mixed-effects models with random slopes and intercepts for each participant. We use logistic regression for binary outcome variables and linear regression otherwise. We report nonstandardized regression coefficients throughout, with time in units of seconds, accuracy as a proportion (0–1), and judgments in the original 1–5 scale. Degrees of freedom are estimated using the Satterthwaite method. As mentioned above, we excluded 106 (17%) participants who skipped more than 90% of nonpractice critical trials. We also exclude all trials with incorrect responses: 905 (5%) intrusions and 1,153 (6%) unclassified errors, as well as 207 (1%) time-out trials. This leaves only the correct recall and skip trials. Finally, we excluded one trial with a response time under 30 ms (as planned but not explicitly preregistered).

In the experiment itself, responses were considered correct only if they matched the target word exactly (ignoring white space and capitalization). When analyzing the data, however, we additionally marked as correct any response for which a spellchecker (implemented in the pyspellchecker package) identified the correct word as a possible intent of the given response. Similarly, responses with a single character were treated as skip trials.

In order to eliminate typing-related variance in response time, we defined response time as the time between stimulus presentation and the first key press initiating the response. If the input box was ever cleared (presumably because the participant changed their mind about which word to enter), response initiation time is defined as the last key press when the text box was empty (i.e., the beginning of typing the final response). For skip trials (indicated by submitting an empty response), we use the time the final response was made, ignoring any earlier typing (of which there was usually none).

Transparency and Openness

All sample sizes, exclusion criteria, statistical analyses, modeling procedures, and plotting decisions were preregistered (https://aspredicted.org/wr9ej.pdf). After preregistering, we discovered a conceptual error in our specification of a null model. This led us to remove one plot that we discovered the more flexible null model could capture (indicating that the plot was not actually a good test of rational metamemory). See Appendix B for details, including the removed plot. Additionally, we previously ran a preregistered version of this experiment with a smaller sample size and a slightly different analysis (which produced a marginally significant result). See Appendix C for details, including full results with the previous data set.

Results

Preliminary Analyses

Before addressing the model's key predictions, we first describe participants' basic task performance.[5] Overall, participants correctly recalled the target in 35.9% of trials and skipped 52.0% of trials. Of the remaining trials, 4.8% were intrusions, 6.2% were unclassified errors, and 1.1% were time-outs. This second group of trials cannot be explained by the model because we do not model competition between items in memory nor the attentional lapses that lead to timing out (one should always skip before hitting the time limit to avoid the penalty). Thus, as preregistered, we focus all remaining analyses on the correct recall and skip trials, excluding errors and time-outs.

As shown in Table 1, recall rate in the critical trials varied strongly with recall rate in the pretest trials, from 1.5% for targets that were never correctly recalled in the pretest to 94.4% for targets that were correctly recalled in both pretest trials. Targets that were recalled correctly on only one pretest trial yielded an intermediate rate of 55.9%.[6] All three levels of pretest accuracy are represented fairly well, with the 50% case accounting for the smallest proportion of trials (5.4%), and all three models capture this distribution well. However, the optimal model systematically overpredicts recall rates, especially for the 50% trials. One plausible explanation for this

[5] The analyses presented in this section were added based on reviewer suggestions and were thus not preregistered.

[6] This case can be further divided into cases in which the correct response was provided on the first or the second pretest trial, with recall rates of 35.0% and 73.4%, respectively. This difference could be captured in the model by including a decay or drift in memory strength between trials. For simplicity, however, we do not consider such sequential effects and collapse across the two cases in the remaining analyses.

Table 1
Recall Rates in Critical Trials by Pretest Accuracy

Model/Data	0%	50%	100%
Optimal metamemory	.039 (.512)	.668 (.071)	.974 (.417)
No meta-level control	.033 (.600)	.518 (.082)	.934 (.319)
Human	.015 (.554)	.559 (.054)	.944 (.391)

Note. Each cell shows the recall proportion conditional on pretest accuracy. The proportion of trials at the given pretest accuracy level is shown in parentheses. The human values exclude error and time-out trials. See Table A1 in Appendix A for the proportions of all response types.

discrepancy is that the model invests more effort into recall on the critical trials because the payoff is three times higher than on pretest trials. People may be undersensitive to this change in incentives (cf. van den Berg et al., 2023).

Table 2 shows the distribution of confidence and feeling-of-knowing judgments. As expected, we see that people tend to give high confidence judgments (after providing an answer) and low feeling-of-knowing judgments (after declining to provide an answer). However, the full range of responses is covered in both cases. Both types of judgments were accurate in the sense that they reflected performance in the pretest, confidence: $B = 0.575$, 95% CI [0.451, 0.698], $t(287.6) = 9.14$, $p < .001$; feeling-of-knowing: $B = 1.209$, 95% CI [1.023, 1.394], $t(176.6) = 12.77$, $p < .001$. Additionally, for all trials where a response was given, the confidence judgments were predictive of accuracy ($B = 1.360$, 95% CI [1.258, 1.462], $z = 26.18$, $p < .001$), from 15.9% at the lowest level to 94.9% at the highest.

Having established the expected basic relationships between pretest accuracy, recall in the critical trials, and metamemory judgments, we now turn to our primary response time analyses.

Metamemory Judgments and Response Time

As illustrated in Figure 5A (right), we replicated the basic pattern reported by Costermans et al. (1992). Participants were faster to correctly recall targets that they reported greater confidence in ($B = -0.244$, 95% CI [−0.294, −0.194], $t(516.1) = -9.54$, $p < .001$), but slower to skip targets that they reported higher feeling-of-knowing for ($B = 0.385$, 95% CI [0.309, 0.460], $t(189.1) = 10.02$, $p < .001$). The confidence effect indicates that people's metamemory judgments correlate with true memory strength (assuming that stronger memories are recalled more quickly), suggestive of metacognitive monitoring. In contrast, the feeling-of-knowing effect suggests that participants spent longer trying to recall targets that they thought they were more likely to recall, a form of metacognitive control.

Table 2
Distribution of Metamemory Judgments

Judgment	1	2	3	4	5
Confidence (recalled)	.023	.039	.118	.211	.609
Feeling of knowing (skipped)	.820	.107	.048	.009	.016

Note. Each cell shows the proportion of a given level of confidence/feeling-of-knowing judgment. Note that the models are fixed to have the same proportions (see the Method section).

To capture these metacognitive judgments in the model, we assume that the judgment (confidence or feeling of knowing) is made based on the inferred evidence accumulation rate at the end of the trial (see the Simulation Procedure section). Unsurprisingly, the optimal model infers a higher accumulation rate when a word is recalled faster and thus produces higher judgments. More importantly, it gives higher feeling-of-knowing judgments for cues that it took longer to skip. A lesioned model with the same object-level process but no metalevel control (sampling skipping times randomly; see the Method section) failed to capture either effect. Interestingly, with some parameter values, the lesioned model can capture either effect in isolation. However, it was not able to predict both effects at once with any parameter configuration (see Appendix D).

Pretest Accuracy and Response Time

While the above results are suggestive, the direction of causation is not clear. The metamemory judgment temporally follows the response; thus, it is entirely possible that participants are actually reporting higher feeling-of-knowing judgments because they spent longer searching. To test whether participants stopping times are truly influenced by an awareness of the memory's strength, we can replace the metamemory judgment with an objective measurement of memory strength, concretely, the proportion of pretest trials in which the target was recalled correctly. As shown in Figure 5B, the model predicts a similar pattern: faster recalls and slower skips with increasing pretest accuracy.[7] People were likewise faster to recall targets that they had previously recalled correctly ($B = -0.978$, 95% CI [−1.215, −0.742], $t(253.3) = -8.11$, $p < .001$). More importantly, they were also slower to skip such targets ($B = 0.404$, 95% CI [0.251, 0.558], $t(193.7) = 5.16$, $p < .001$). The lesioned model could not produce this effect with any parameter values. These results suggest that participants' decisions to stop searching depended on a metacognitive awareness of how likely they were to recall the target.

Discussion

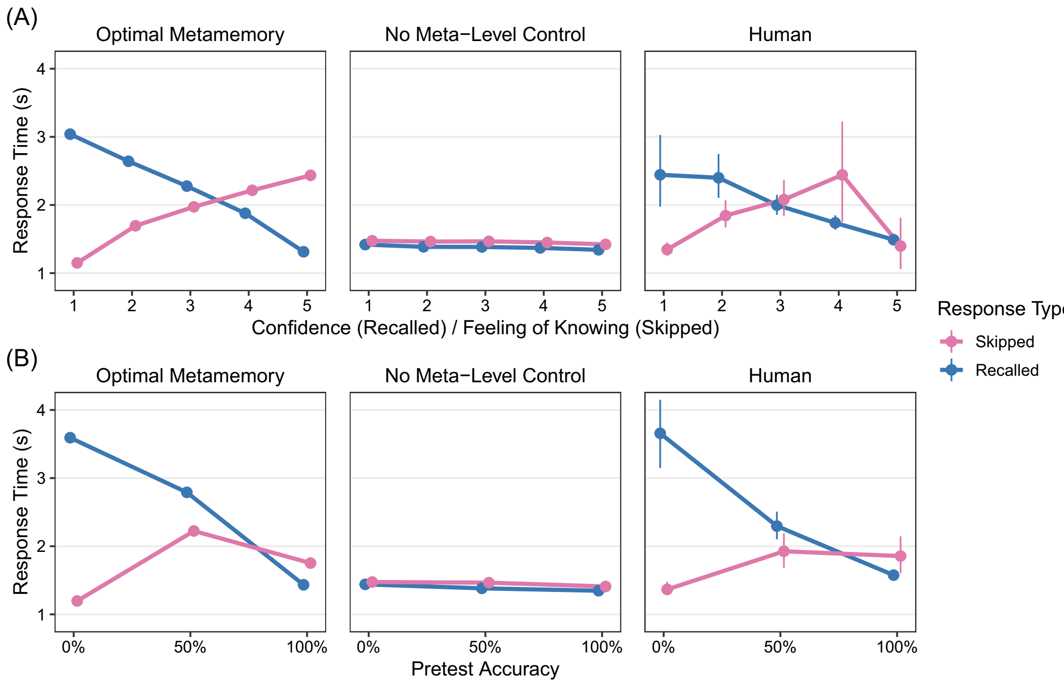
In this experiment, we found that participants were faster to recall targets with higher strength but slower to give up on targets with higher strength. Importantly, this pattern held for both a subjective measure of strength (replicating Costermans et al., 1992) as well as an objective measurement of strength (accuracy on the pretest trials). The latter is critical because it shows that people spent longer searching for memories that they were actually more likely to recall, thus demonstrating the objective utility of metamemory in guiding recall. Furthermore, because pretest accuracy is defined before the critical trials, this measure is not subject to the reverse-causality concern that response times are driving feeling of knowing rather than vice versa (Schwartz, 2001).

The full pattern of results was qualitatively consistent with the optimal model, which terminates search when the expected value of search falls below the expected cost, and reports a Bayesian estimate of

[7] The nonmonotonic prediction for skip trials is due to a selection effect: the optimal model only skips high-strength memories when it greatly underestimates the strength. This becomes increasingly unlikely as the trial progresses and more evidence is collected. Thus, these "erroneous" skips generally occur quickly (cf. the "fast errors" phenomenon in decision making Ratcliff & Rouder, 1998).

Figure 5

Opposing Effects of Memory Strength on Time to Recall Versus Skip a Target



Note. (A) Reaction time as a function of metamemory judgment (feeling of knowing for skip trials, confidence for recall trials), separately for trials in which participants *correctly* recalled the target versus skipped without responding (errors are excluded). The left panel shows the predictions of the proposed optimal metamemory model, the center panel shows the predictions of a model with the same recall process but no meta-level control (sampling stopping times randomly), and the right panel shows human data. The models' metamemory judgments are made based on the inferred memory strength at the end of the trial. (B) Response time as a function of the accuracy rate for the presented cue in the pretest phase. Points show means of participant medians, and error bars show 95% bootstrapped confidence intervals over participant medians. Each model is treated as one participant (with 1 million simulated trials). All plotting decisions (including which effects to show, the aggregation method, and axis limits) were preregistered. See the online article for the color version of this figure.

strength as feeling of knowing or confidence. Perhaps more importantly, the results could *not* be captured by a model without meta-level control. This provides computational support for the intuition that the correlation between search time and feeling of knowing is a distinctive signature of an adaptive metamemory process.

Experiment 2

In our first experiment, we considered a very simple form of metamemory, the decision of how long to search memory before giving up. Control of memory is not limited, however, to such a simple kind of decision. Instead, successful recall often requires deciding between multiple strategies for finding an answer (Reder, 1988). Going further, Koriat (2000) had characterized recall as a form of problem solving, with a meta-level process "coordinating between different operations directed toward the recovery of the elusive memory" (p. 334). A careful characterization of the operations underlying recall (let alone how they are chosen) is beyond the scope of this article. Nevertheless, in our second experiment, we sought to characterize a core aspect of this richer form of metamemory: monitoring multiple target memories and allocating retrieval efforts between them.

On each trial, participants were presented with two cues and could recall the target for either one of them (Figure 6). They thus had to make a metacognitive decision about which of the two possible targets to search for at each moment. In order to observe how this selection process unfolded over time, we used a keypress-contingent display, such that only one cue was visible at any moment. This provides process-tracing data similar to eye tracking, but in a format amenable to online presentation.

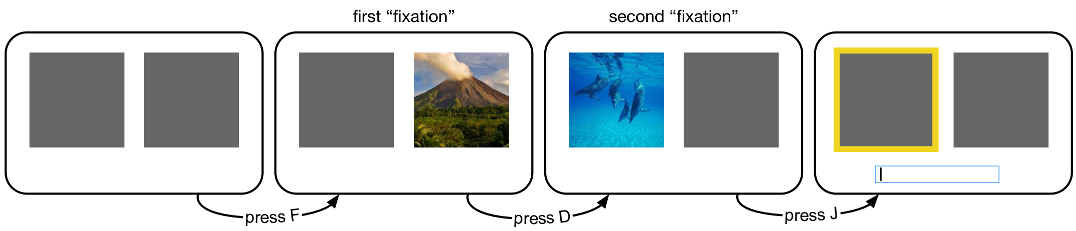
Predictions

As detailed below, we extended the model to the multiple-memory case by creating a separate evidence accumulation process for each target memory. The meta-level process decides which accumulator is allowed to progress at each time point. The optimal policy (illustrated in Figure 7) can thus be characterized by when it decides to switch between the two memories (and also when it decides to terminate search, as before).

In general, the optimal policy attends to the memory that it believes can be recalled soonest, as this will incur the least cost. In our experiment, attending to a memory is operationalized by looking at (or "fixating") the associated cue. Thus, the basic prediction is that

Figure 6

Experiment 2: Critical Trials



Note. Participants were presented with two images on each trial and were instructed to recall the word associated with either of them. Only one cue was visible at a time and participants could flip between them with the D and F keys. At any point, they could press J or K to select an image for recall, at which point they had 5 s to enter the associated word. We refer to the periods of time in which one cue was contiguously visible as "fixations." See the online article for the color version of this figure.

the cues for stronger memories will receive a greater share of the total fixation time. More specifically, the model will be slower (and less likely) to switch away from a strong memory but faster to switch when the other memory is strong.

Inspecting model simulations, we also discovered a surprising feature of the optimal policy. Its final fixations are longer than its nonfinal fixations. This prediction is surprising because we see the opposite pattern in decision-making tasks (e.g., Krajbich et al., 2010, discussed further in the General Discussion section). Why does the optimal policy make this prediction? We can understand long final fixations as a sort of "rational commitment" behavior, in which the model effectively commits to recalling one memory before it is actually recalled. After committing to a memory, the model continues to attend to it until it is either recalled or its inferred strength drops well below the competitor. The latter occurs only rarely. Thus, commitment tends to happen on final fixations, and final fixations are therefore longer. To see why this is rational, note that constant switching between the two memories is wasteful, as it can take up to twice as long as if one had immediately committed to one memory. On the other hand, immediately committing to the first cue could lead to getting stuck on an out-of-reach memory. Thus, the model only commits to a cue after becoming reasonably confident that the cue is strong.

We now describe an experiment that tests these predictions.

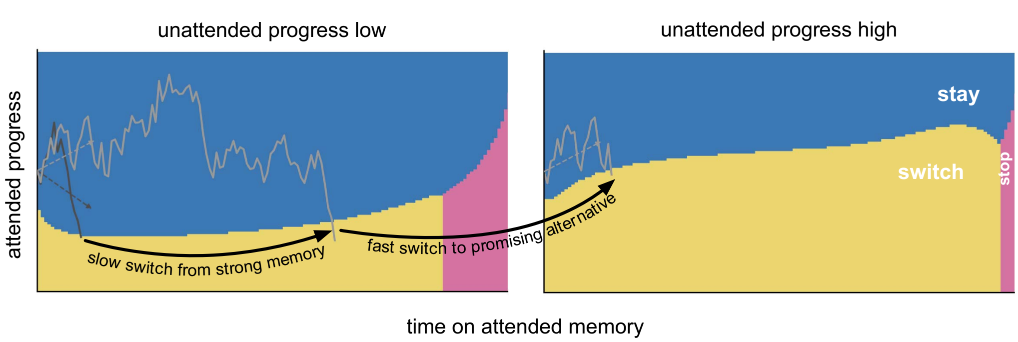
Method

Participants

We recruited 685 participants through Prolific with the restriction that they reported current U.S. or U.K. residence, had at least a 95% approval rating, and had not participated in any related studies (including pilots and Experiment 1). As preregistered, we excluded 184 (27%) participants who failed to correctly recall a target on more than 50% of critical trials. This yielded 501 participants in our final analysis. The target sample size of 500 participants had over 95%

Figure 7

Experiment 2: Optimal Policy and Predictions



Note. With two possible memories to recall, the optimal policy partitions the state space into three sections where it is optimal to either: continue searching for the currently attended memory (blue), switch to the other memory (yellow), or give up (pink). The optimal policy depends on the recall progress and time spent on both memories; here, we show two slices of the full four-dimensional state space, setting the time spent on the unattended memory to 30 time steps and its progress to either −0.3 (left) or 0.1 (right). The gray lines show example progress traces for a weak (dark gray) and moderately strong (light gray) attended memory. The arrows highlight two key features of the optimal policy: It is slower to switch when the currently attended memory is strong (vs. weak) but faster to switch if the unattended memory has already shown promising recall progress (vs. if it has not). This plot was generated with parameters fit to the data in Experiment 1 (the same parameters used to generate the behavioral predictions). See the online article for the color version of this figure.

power with $\alpha = .05$ for all our preregistered hypotheses based on a boot-strapping power analysis conducted on pilot data. This experiment was approved by the institutional review board of Princeton University (Protocol Number 10859).

Stimuli

The stimuli were identical to those used in Experiment 1.

Procedure

The procedure was identical to Experiment 1 with two exceptions. First, we lengthened the training phase to include two blocks of exposure (with each cue/target pair shown once) and one intervening block of cued recall that was identical to the pretest block except that each pair was shown only once. Second, the critical trials followed an entirely different design described below.

Critical Trials. The critical trials employed a modified form of cued recall in which two cues were presented on each trial. At the beginning of each trial, two gray occluders were displayed. Participants could temporarily remove the occluders, revealing the cue image underneath, by pressing the J and K keys. However, revealing one image would hide the other one. The 15-s timer appeared and began counting down when the first image was revealed. At any point, participants could press D or F to select one of the two cues for recall. At this point, both images were hidden, a yellow ring was drawn around the occluder for the selected image, and a text box appeared where they could enter the word associated with the selected image. Correct/incorrect feedback was provided after each response. Unlike Experiment 1, there was no penalty for errors (and hence, no skipping mechanism), no additional time incentive (besides the small incentive for fast correct answers that was also present in the pretest trials), and no collection of metamemory judgments. Note that the lack of a skipping mechanism means that we cannot distinguish between genuine recall errors and the decision to give up on recall and enter a random guess (i.e., performing the STOP action). For this reason, we only analyze trials in which a target was correctly recalled. We decided to omit the skipping mechanism despite this drawback because we were specifically interested in the switching decisions (not the stopping decisions), and we wished to minimize task complexity.

Modeling

To generalize the model to the case with multiple memories that could be recalled, we assume that each memory is associated with its own independent object-level recall process. Furthermore, we assume that progress can be made on only one memory at a time, with the progress of the nonattended memory being held fixed.[8] The state is defined $s_t = (t^L, t^R, z_t^L, z_t^R, f)$, with t^L and t^R denoting the number of time steps the "left" and "right" memory have each been attended, z_t^L and z_t^R denoting their respective progress levels, and $f \in \{L, R\}$ indicating which memory is currently attended. When the progress for either memory hits the threshold, the corresponding target is recalled.

At the meta-level, the agent now has three actions: STAY, SWITCH, and STOP. STAY and SWITCH are both similar to SWITCH, but the latter additionally flips the value of f. We assume that there is some reconfiguration associated with switching. Thus, the new reward function is:

$$r(s_t, a_t) = \begin{cases} U(\text{recall}) & \text{if } \max\{z_t^L, z_t^R\} \geq \theta \\ -\gamma_{\text{SEARCH}} & \text{if } a_t = \text{STAY} \\ -(\gamma_{\text{SEARCH}} + \gamma_{\text{SWITCH}}) & \text{if } a_t = \text{SWITCH} \\ 0 & \text{if } a_t = \text{STOP}. \end{cases} \tag{11}$$

The new transition function has two parts. First, if the SWITCH action is taken, f is flipped. Then the original one-memory transition function is applied to the attended cue, updating the corresponding t and z_t variables. If the left memory is attended, we have

$$T(s_{t+1}|s_t, a) = p(z_{t+1}^L | t^L, z_t^L), \tag{12}$$

and similarly, if the right cue is attended.

Computing the Optimal Policy. We again computed the optimal policy by backward induction. We applied the same discretization and computed the transition function in the same way. Note that, because recall progresses for only one memory at a time, it is not necessary to represent the transition function over the full state space.

To compute the value functions, we began by initializing the value of terminal states to U(recall) if either recall progress exceeded the threshold and 0 if the combined time exceeded 150. We then computed the value at previous time steps by iterating backward. In this case, for each time step, we must consider all combinations of time spent on each item that sum to the time step under consideration. Additionally, we must consider three possible actions. Assuming the left cue is currently attended, the action values are

$$\begin{aligned} &Q^*(s, \text{SEARCH}) = \\ &\quad \sum_{z_{t+1}^L} p(z_{t+1}^L | t^L, z_t^L) V^*(t^L + 1, t^R, z_{t+1}^L, z_t^R, L) - \gamma_{\text{SEARCH}} \\ &Q^*(s, \text{SWITCH}) = \\ &\quad \sum_{z_{t+1}^R} p(z_{t+1}^R | t^R, z_t^R) V^*(t^L, t^R + 1, z_t^L, z_{t+1}^R, R) - \gamma_{\text{SEARCH}} - \gamma_{\text{SWITCH}} \end{aligned}, \tag{13}$$

with $Q^*(s, \text{STOP}) = 0$ as before. Besides these differences, the procedure is identical to the one-memory case.

Parameter Estimation. Due to the high dimensionality of the data in this experiment (sequences of fixation durations), maximum likelihood estimation is computationally prohibitive. Although approximate fitting schemes are possible, given that we are not interested in the quantitative fit of the model, we instead used this as an opportunity to test the generalization capabilities of the model (cf. Krajbich & Rangel, 2011). That is, we simply used the best-fitting

[8] This assumption is plausible given that the stimuli, images, are difficult to hold in working memory. Nevertheless, it is possible that recall would continue to progress for the unattended memory, albeit at a slower rate. We did not consider this possibility for reasons of simplicity and computational efficiency (allowing both memories to progress at once squares the MDP's branching factor, making backward induction far less efficient). Allowing for this would encourage the policy to quickly check both cues to allow for such parallel processing; however, the main qualitative predictions would likely remain the same.

parameters from Experiment 1. For the switch cost parameter, which was not present in the Experiment 1 model, we arbitrarily set $\gamma_{\text{SWITCH}} = \gamma_{\text{SEARCH}}$, noting that the predictions do not depend greatly on the exact value of this parameter. However, because the model must predict the duration of each fixation, the original NDT model is no longer appropriate. Instead, we assumed that NDT was added to each fixation independently. We fit the parameters of this model by maximizing the likelihood of all nonfinal fixation durations, assuming (for tractability) that they were independent and identically distributed. We excluded final fixations from this fitting procedure because they have different distributional properties (discussed further in the Rational Commitment section below). The fitted NDT parameters were $\mu_{\text{NDT}} = 615$, $\alpha_{\text{NDT}} = 2.75$.

Lesioned Model Without Meta-Level Control. The lesioned model is an extension of the lesioned model from Experiment 1, with an additional mechanism to determine fixation durations. As before, the stopping time was sampled from a Gamma distribution at the beginning of each trial. Similarly, at the beginning of each fixation, the switching time was sampled from a second Gamma distribution. If this time was reached before the memory was recalled or the stopping time was reached, then the model switched to attending the other cue and sampled a new switching time.

To give the lesioned model the best chance of capturing the qualitative effects, we fit all of its parameters to the behavioral data (in contrast to the optimal model, which uses parameters fit to Experiment 1). Computing an exact likelihood is intractable in this case; thus, we approximated the likelihood by assuming that the duration of each fixation depends only on the pretest accuracy of the fixated and nonfixated cues, and whether or not the fixation is final. Given this assumption, we estimated the likelihood in the same way as for Experiment 1, with the exceptions that the histogram had size $3 \times 3 \times 2 \times 151$ (three accuracy rates for each cue, final vs. nonfinal, and 151 response time bins) and that the likelihood was computed per fixation rather than per trial. Note that we constructed the histogram using only correct simulated trials, as we exclude error trials from the human data. The MLE was($\mu_0 = 0.083$, $\sigma_0 = 0.103$, $\sigma_x = 0.148$, $\mu_{\text{stop}} = 4{,}527$, $\alpha_{\text{stop}} = 76.16$, $\mu_{\text{switch}} = 4{,}943$, $\alpha_{\text{switch}} = 0.18$, $\mu_{\text{NDT}} = 814$, $\alpha_{\text{NDT}} = 3.56$.

Statistical Analyses

As in Experiment 1, all reported regressions are linear mixed-effects models with nonstandardized regression coefficients (see the Method section of Experiment 1 for details). As mentioned above, we excluded 184 (27%) participants who failed to correctly recall a target on more than 50% of critical trials. This yielded 501 participants in our final analysis. We also exclude all trials with incorrect responses: 497 (5%) intrusions and 552 (6%) unclassified errors, as well as 309 (3%) time-out trials. This leaves only the correct trials (as there was no option to skip a trial). As before, we marked a response as correct if a spellchecker suggested the correct word as a possible intent of the given response.

Transparency and Openness

All sample sizes, exclusion criteria, statistical analyses, modeling procedures, and plotting decisions were preregistered (https://aspredicted.org/xq9nx.pdf). As in Experiment 1, we used a more flexible lesioned model than we originally intended. Due to this change, we elected to fit the parameters of the lesioned model to data rather than using parameters from Experiment 1 as preregistered. Furthermore, we introduced a new plot and accompanying regression to better distinguish between the models regarding the "rational commitment" prediction. See Appendix B for details. As in Experiment 1, we previously ran a preregistered version of this experiment; in this case, we reran the experiment because we discovered a conceptual flaw in the original analysis plan. See Appendix C for details, including full results with the previous data set.

Results

Attention Is Drawn to Stronger Cues

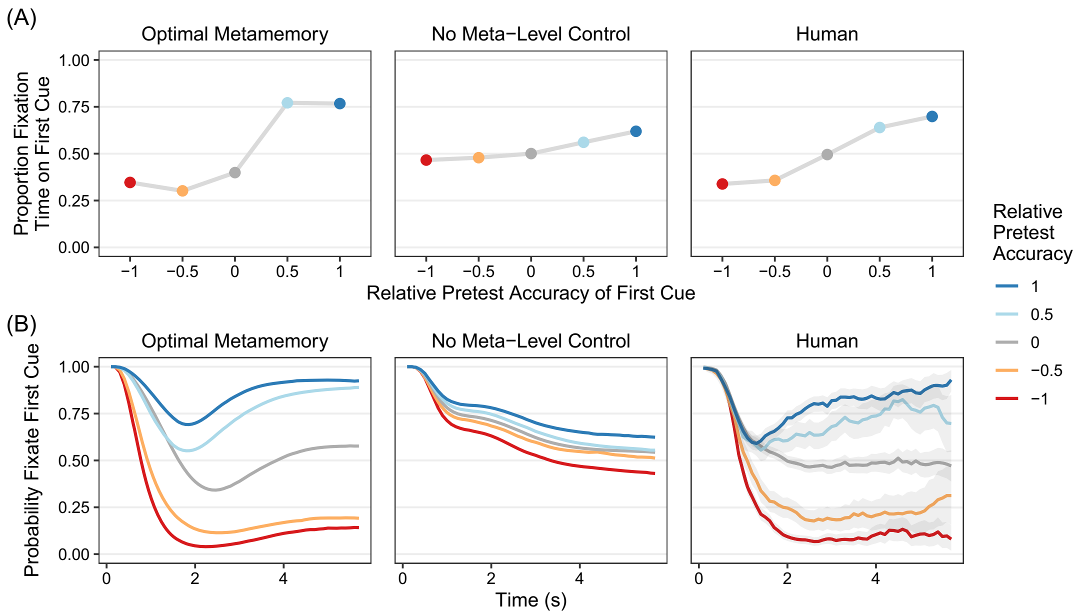
The critical model predictions concern participants' "fixation" behavior in the double cued-recall trials, that is, the sequence of key presses they made to alternately display the two images. The most basic prediction of the optimal model is that participants should attend more to the cue assocaited with a stronger memory. Intuitively, the target for the stronger cue can be recalled faster, and so time spent looking at this cue is more productive. Indeed, as illustrated in Figure 8A, participants spent substantially more time looking at cues that were stronger than the other available cue ($B = 0.187$, 95% CI [0.178, 0.196], $t(326.9) = 41.46$, $p < .001$). However, this pattern is also shown (to a lesser extent) by a lesioned model that randomly switches between the cues. This is due to two properties of the object-level recall process. First, the last fixation is always on the cue whose target is recalled. Second, stronger cues are more likely to be recalled. Together, this implies that stronger cues are more likely to be fixated last and thus receive more fixations (and more fixation time) on average.

Inspecting the time course of attention across the trial (Figure 8B) reveals a more nuanced picture. People tend to quickly check both cues (as indicated by the initial dip in the probability of fixating on the first cue). Then, if the second cue is stronger (red lines), they continue fixating it. But if the first cue is stronger (blue lines), they switch back to it. From about 1–3 s, participants show an increasingly strong tendency to fixate on the stronger cue, and this tendency remains stable for the remainder of the trial. The optimal model shows a similar pattern, although its tendency to fixate the stronger cue emerges faster (as indicated by the earlier divergence of fixation probability for different relative strengths). In the lesioned model, a slight tendency to fixate the stronger cue emerges in the first second and it remains small for the remainder of the trial. This is due to the last fixation confound discussed above.

Nonfinal Fixation Durations

Although the lesioned model was not able to capture the strength or time course of the tendency to fixate stronger cues, the fact that it can produce the effect at all casts some doubt on the interpretation of the pattern in human data. Thus, for our next analysis, we inspected the duration of individual nonfinal fixations, which are clearly not subject to the last fixation confound. As illustrated in Figure 7, the optimal policy's decision to terminate a fixation by switching to the other cue depends on the recall progress of both targets; it is slower to switch when the currently attended cue is generating rapid progress but faster

Figure 8
Attention Is Drawn to Stronger Cues



Note. (A) The proportion of total viewing time allocated to the first-seen cue image as a function of the difference in pretest accuracy of the first- and second-seen cues. Trials for which the second cue was never shown are excluded. Note that the 95% confidence intervals are too small to be distinguishable. (B) The probability that the first-seen cue is currently displayed over the course of the trial, split by relative pretest accuracy. See Figure A2 for a version of this figure split by total pretest accuracy. See the online article for the color version of this figure.

to switch when the unattended cue has already generated substantial progress. The model thus predicts that nonfinal fixation durations will increase[9] with the pretest accuracy of the fixated cue but decrease with the pretest accuracy of the nonfixated cue. As illustrated in Figure 9, both of these predictions were confirmed: Participants' nonfinal fixations increased with the pretest accuracy of the fixated cue ($B = 0.101$, 95% CI [0.069, 0.133], $t(219.9) = 6.22$, $p < .001$), and decreased with the pretest accuracy of the nonfixated cue ($B = -0.432$, 95% CI [−0.564, −0.300], $t(71.4) = -6.41$, $p < .001$; first fixations excluded). The lesioned model predicts no such effect. In fact, it is incapable of predicting either effect under any parameter setting that achieves accuracy levels comparable to our participants (with very low accuracy, it can capture the effect through a selection mechanism; see Appendix D).[10]

Rational Commitment

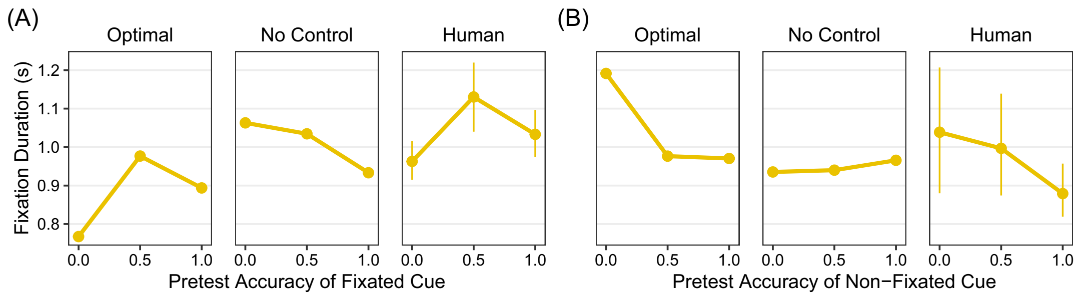
For our final analysis, we tested the model's "rational commitment" behavior, in which the model effectively commits to recalling one memory before it is actually recalled. One observable consequence of this is that the model's final fixations are longer than their nonfinal fixations, as the commitment decision can occur when the memory is still well below the threshold. Consistent with this, Figure 10A shows that our participants' final fixations were indeed longer than their nonfinal fixations ($B = 0.869$, 95% CI [0.810, 0.929], $t(429.4) = 28.86$, $p < .001$). However, this figure also shows that the lesioned model can capture this effect as well. It is able to do this through a "random commitment" mechanism: by assuming a high-variance switching-time distribution, it occasionally samples a very long fixation duration, which is likely to end in recall.

Did participants' long final fixations reflect rational commitment or random commitment? The key distinction between these types of commitment is that it is only rational to commit to a memory that is at least as strong as the alternative. Therefore, only with a random commitment strategy will one allocate long fixations to cues that are weaker than the alternative. Figure 10B thus shows the frequency with which the weaker memory is fixated, separately for fixations of different lengths. Importantly, we do not limit this analysis to final fixations as this would exclude the cases where the lesioned model sampled a long fixation duration on a low-strength cue, thus selecting for the cases where the lesioned model was "accidentally"

[9] To be exact, the model predicts a nonmonotonic effect such that fixations are longest for intermediate strength cues. This is due to the same selection effect that produces a nonmonotonic prediction in Figure 5B. Switching away from a strong cue suggests that the model incorrectly estimated its strength (because progress happened to be slow at first). This becomes increasingly unlikely as the fixation progresses and more evidence is collected. Such "erroneous" switches are thus more likely to occur quickly.

[10] Again, the lesioned model predicts a reverse effect due to selection effects similar to those discussed in Experiment 1 and the previous footnote. For the fixated cue's strength, it is exactly the same logic as in Experiment 1. A strong cue is likely to be recalled quickly; it will only be switched away from when a short switching time is sampled. For the nonfixated cue's strength, the logic is more complex. Because we condition one of the memories being recalled, the nonfixated cue being weak implies that the fixated cue is strong. Following the same logic as before, this in turn implies that a short switching time was sampled (as otherwise the cue would have been recalled).

Figure 9
Nonfinal Fixation Durations



Note. (A) The duration of nonfinal fixations is a function of the pretest accuracy of the currently fixated cue. (B) The same, but for the pretest accuracy of the cue that is *not* currently fixated (first fixations excluded). See the online article for the color version of this figure.

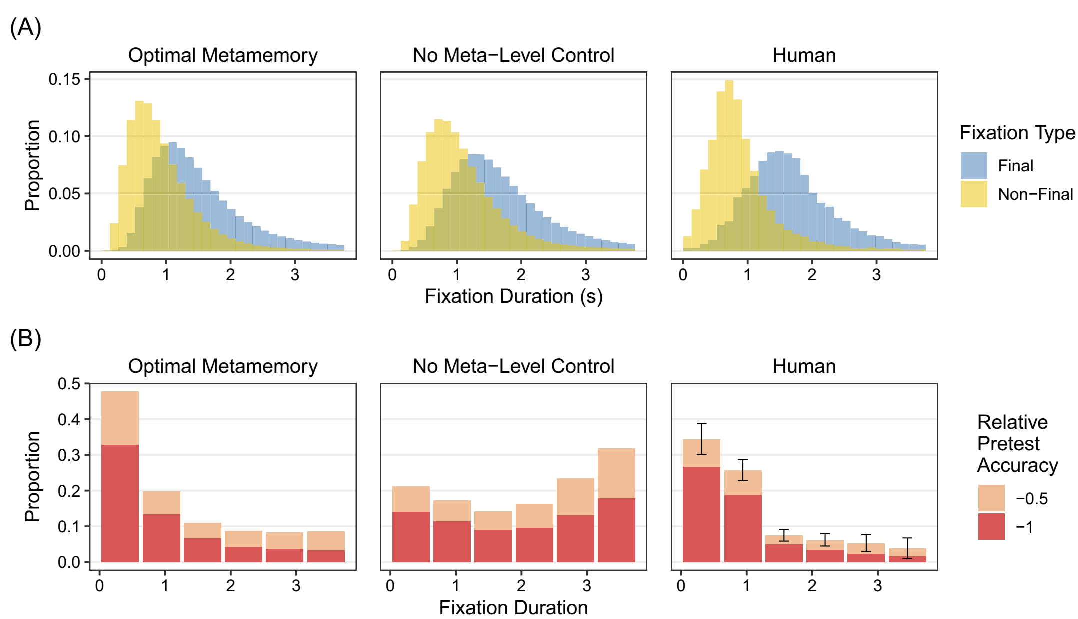
rational. In both the optimal model and the human data, we see that long fixations are unlikely to be directed to weaker memories. In particular, the probability of fixating the weaker memory significantly decreased with fixation duration ($B = -1.376$, 95% CI [−1.537, −1.215], $z = -16.78$, $p < .001$; logistic regression, excluding trials where the cues had equal pretest accuracy). In contrast, under the lesioned model, the weaker cues are actually more likely to receive long fixations (because strong cues are likely to be recalled quickly, cutting off long fixations). The fact that people show a strong tendency not to direct long fixations to weak cues despite this selection effect suggests that their commitment decisions were indeed rational.

Note that we did not preregister this final analysis because our initial, less flexible, implementation of the lesioned model that sampled switching times from the empirical distribution could not capture the long final fixation effect (see Appendix B, Figure D5A). However, the effect also holds in a previously collected data set, which we did not use when developing Figure 10B or the accompanying statistical test. This provides a quasiconfirmatory test of the exploratory analysis in the previous paragraph. See Appendix C for details.

Discussion

In this experiment, we found evidence of a richer form of meta-level control of memory. Specifically, when presented with two cues associated with different target memories, our participants directed their attention toward the cue associated with the stronger memory.

Figure 10
Rational Commitment to Stronger Cues



Note. (A) The distribution of final and nonfinal fixation durations. (B) The proportion of fixations on items with lower pretest accuracy than the alternative, separately for fixations of different durations (first fixations excluded). The error bars indicate 95% confidence intervals computed over the total proportion (including −0.5 and −1) for each participant. See the online article for the color version of this figure.

This was reflected in the overall proportion of fixation time, the time course of fixations, and the duration of individual nonfinal fixations. The latter two results are especially important because they are behavioral signatures of metacognitive control during the retrieval process itself (before a recall is completed or abandoned). To our knowledge, this is the first empirical demonstration of a dynamic metamemory process unfolding over time.

General Discussion

In this article, we presented an optimal model of meta-level control for cued recall. The model consists of a metacognitive process that monitors the progress of a basic recall process and optimally controls how long the process is allowed to continue (either terminating recall or switching to recall of a different memory). In two experiments, we showed that human behavior is qualitatively consistent with the predictions of this model. In Experiment 1, we replicated and extended the findings of Costermans et al. (1992), showing that people were faster to recall targets with higher strength (as indicated by both subjective judgments and objective performance) but slower to give up on targets with higher strength. In Experiment 2, we showed that people also attend to stronger memories when they can choose between two possible targets. Together, our results suggest that people can estimate the strength of a memory that they are trying to recall and use this information to adaptively control their retrieval efforts.

An Explicit Instantiation of a Classic Theory

Our results contribute to the metamemory literature by providing (to our knowledge) the first computationally explicit instantiation of the classic theory of metamemory proposed by T. O. Nelson and Narens (1990) in the context of memory recall. According to this theory, metacognition involves the interaction between a meta-level process and an object-level process, where the meta-level process monitors the state of the object-level process and controls it accordingly. In the context of memory recall, they proposed a verbal model in which a feeling of knowing is generated by “an evaluation in terms of whether there has been sufficient progress to continue,” with the process terminating in an omission error when this feeling of knowing no longer “exceeds the FOK threshold for claiming to know the answer” (p. 137).

Here, we have formalized this classic model as a sequential decision problem in which the meta-level process executes a sequence of actions (continuing or terminating memory search) to maximize reward (utility of recall minus cost of search) given the observed state of the object-level process (recall progress). By formalizing this problem as an MDP, we could identify the optimal metacognitive control policy using standard dynamic programming techniques. This allowed us to generate quantitative predictions about the observable behavior we would expect to see if people were indeed using a rational metamemory system to control how long they search for an elusive memory. By confirming these predictions in an experiment, we have contributed quantitative support for this classic theory, which was previously supported only by intuitive qualitative predictions.

Formalizing Nelson and Naren’s model as an MDP also allows us to create a conceptual link between metacognition and RL (Sutton & Barto, 2018). Specifically, we map the meta-level and object-level processes onto the concepts of agent and environment, respectively. That is, we model metacognition as a meta-level agent interacting with an object-level environment in much the same way as one would model, for example, a mouse (the agent) searching for food in a maze (the environment). RL has become a major theoretical foundation in the psychology and neuroscience of decision making (Dayan & Daw, 2008; Glimcher, 2011; Niv, 2009). However, for the most part, it has been used to model the interaction between agents and *external* environments. Applying RL to model the metacognitive interaction between an agent and its *internal* environment (cf. Simon, 1955) creates the opportunity to transfer much of what we know about how people learn to act effectively in the world to understand how people learn to think effectively in their own minds (Lieder et al., 2018).

Rational Analysis for Metamemory

Outside of metamemory, the intellectual roots of our model lie in Anderson’s (1990) rational analysis, with early work demonstrating that the forgetting behavior commonly observed in lab settings is not a weakness or peculiarity of the memory system, but instead a reflection of rational adaptation to the statistics of the environment (Anderson & Milson, 1989). The key idea underlying both this model and ours is that memory (and cognition more generally) can be treated as an optimization problem. More recently, researchers have emphasized that this optimization problem must also account for the constraints imposed by our limited computational resources (Gershman et al., 2015; Griffiths et al., 2015; Howes et al., 2009; Lewis et al., 2014). This approach, sometimes called resource-rational analysis, has generated insight into a wide variety of cognitive processes (see Lieder & Griffiths, 2020, for a review).

Focusing on memory, a large body of work has shown that many apparent memory biases actually reflect optimal statistical reasoning under the constraint of noise or capacity limitations (Gershman, 2021). For example, in reconstruction tasks, people draw on prior knowledge of a stimulus category to adjust for memory imprecision (Huttenlocher et al., 2000), using more abstract categories for less familiar stimuli (Hemmer & Steyvers, 2009). In working memory tasks, people are sensitive to cost–benefit trade-offs when choosing how many items to encode (Howes et al., 2016), how to allocate encoding resources across items (Yoo et al., 2018), and how much total resource to allocate (van den Berg & Ma, 2018). Researchers have also begun to explore the implications of the constraints that emerge from more detailed models of memory. For example, Zhang et al. (2022) characterized the optimal order in which to recall items from a list, assuming that items are stored and recalled with the context maintenance and retrieval model (Polyn et al., 2009, discussed further below). They find that the optimal policy is to start from the beginning of the list and then sequentially recall forwards, providing a rational account of the often-observed primacy and forward asymmetry effects (Zhang et al., 2022).

Focusing on metamemory in particular, two recent models of judgments of learning are based on signal detection theory (Y. Jang et al., 2012) and Bayesian inference (Hu et al., 2021), both of which have rational bases in probability theory. However, these models only attempt to explain how metamemory judgments are produced, not how they are used—a critical component in a complete rational analysis. In one of the earliest computational models of metamemory, Metcalfe (1993) proposed that feeling-of-knowing judgments are used to adaptively control the weight with which new memories are encoded in a holographic memory representation, thus controlling variance in the representation. In the context of recall, Koriat and Goldsmith (1996) proposed a model of how people avoid giving inaccurate answers, withholding a response when a confidence signal is below a fixed threshold. Bennett et al. (2017) extended this model, providing a more detailed account of how the opt-out decisions relate to recall performance. All of these models are motivated by functional concerns; however, none of them take an explicitly optimal approach (e.g., identifying the confidence threshold that maximizes task performance).

Recent work has begun to fill this gap by developing task-optimized models of metamemory. For example, Hu et al. (2019) proposed a model of cognitive offloading in which an agent decides whether or not to use an external memory aid by comparing the expected increase in recall probability with the cost of using the external aid. Most similar to our own work (from a methodological perspective), Suchow and Griffiths (2016) proposed an MDP model of working memory maintenance in which an agent selects actions that increase the strengths of different memories, as encoded in the state. This model could account for the findings of M. Williams et al. (2013), in which directing participants to forget certain items reduced recall accuracy for the cued item and increased accuracy for uncued items.

Our model draws on several specific ideas from the work described above, in addition to the core principle of rationality. Like Anderson and Milson (1989), we frame the decision to terminate search as a cost–benefit comparison (Equations 2 and 7). Like van den Berg and Ma (2018), we jointly consider the problem of how much resource to allocate (here, the resource being time) as well as how to split that resource between items. Like Hu et al. (2021), we model metamemory judgments as the product of Bayesian inference about the strength of a memory. And like Suchow and Griffiths (2016), we model the resource allocation problem as an MDP. Our work contributes to this literature by synthesizing these previous insights and showing how they can be applied in a new domain, cued recall.

Importantly, like the other resource-rational models reviewed above, our model aims only to characterize the optimal solution to the problem of cognition under limited resources. We have not attempted to explain how the mind might actually approximate that solution. We do note, however, that executing the optimal policy for our model does not require one to continually perform Bayesian inference over the memory strength. Indeed, as illustrated in Figure 3, the optimal policy in the one-cue case is defined by a single boundary, such that recall is terminated if the evidence ever falls below it. Although we identified the shape of this optimal boundary with computationally intensive, model-based methods, it could be well approximated by simple model-free learning mechanisms. Understanding how people learn effective metacognitive strategies is a subject of ongoing research (Binz et al., 2022; Callaway, Jain, et al., 2022; He & Lieder, 2022; Jain et al., 2019; Lieder et al., 2018).

Dynamic Models of Metacognition

Our results also contribute to the literature on metacognition more broadly. Beyond memory, a wide variety of functional roles for metacognition have been proposed, including the regulation of perception (Deroy et al., 2016), judgment (Lebreton et al., 2015; Polanía et al., 2019), decision making (De Martino et al., 2013; Yeung & Summerfield, 2012), learning (Frömer et al., 2021; Nassar et al., 2012), information seeking (Boldt et al., 2019; Desender et al., 2018), and social interaction (Frith, 2012). Some of these roles have been incorporated into computational models that formally describe how confidence might inform decisions about, for example, when to opt out of a difficult trial (Kiani & Shadlen, 2009), when to change one's mind given additional evidence (Folke et al., 2016), how to interpret error feedback (Frömer et al., 2021), or where to fixate in visual search (Stewart et al., 2022). However, these models often assume that the functional role of metacognition is *static*. Although the mechanism underlying confidence judgments may be dynamic (typically being based on evidence accumulation Moreno-Bote, 2010; Pleskac & Busemeyer, 2010; Vickers, 1970), any resulting metacognitive control occurs in a separate stage, typically as a single action, and not feeding back onto the object-level process in a dynamic, interleaved way. In an influential review, Yeung and Summerfield (2012) identified this as a critical gap to be explored in future research.

In the intervening decade, this gap has already begun to be filled. In particular, a substantial body of work has explored the within-trial dynamics of metacognition in controlling a decision-making process. These models track posterior distributions over an evidence accumulation rate and use this information to make optimal decisions about when to stop accumulating evidence (Bitzer et al., 2014; Drugowitsch et al., 2012; Fudenberg et al., 2018; Tajima et al., 2019; Woodford, 2014) or how to allocate attention between multiple options (Callaway et al., 2021; A. I. Jang et al., 2021). However, to our knowledge, this type of model has been applied exclusively in decision-making contexts. Here, we have shown how a model with a very similar structure can be applied to memory recall.

The key structural difference between the model proposed here and these previous models of dynamic metacognition for decision making is the assumption of an *exogenous* threshold associated with the recall. That is, the amount of "evidence" necessary to recall memory is not under the agent's control. This contrasts with decision-making models, where both thresholds (for choosing each option) are *endogenous*. The agent can choose an option based on very little evidence if they so desire. This simple change has a profound effect on the role of attention in the model. In the decision-making context, the sole purpose of fixating an option is to collect information about its utility. In the memory context, fixating the cue similarly provides information about its strength; but it also contributes to recalling the associated memory.

This additional functional role for attention (stimulating recall as well as estimating strength) has two important consequences for the model's predictions in the two-memory case. First, the model predicts—and our results confirm—that cues for stronger memories receive more fixation time. In contrast, optimal models of attention in binary choice predict equal attention to high- and low-value items (Callaway et al., 2021; Fudenberg et al., 2018; A. I. Jang et al., 2021). Second, the model predicts—and again, our results confirm—that final fixations are longer than nonfinal fixations. In contrast, evidence accumulation models of attention-guided decision making predict that final fixations will be shorter than nonfinal fixations, a prediction that is consistently confirmed in data (Krajbich et al., 2010; Krajbich & Rangel, 2011; Tavares et al., 2017). The fact that such a simple structural change can account for these major qualitative differences in the allocation of attention in value-based choice versus cued recall suggests the promise of our MDP-based approach as a general framework for modeling metacognition.

Limitations

While our results make a strong case for the existence of an adaptive metamemory control system in the human mind, we do not claim to have characterized the precise nature of this system. In particular, our results do not speak to the source or nature of metacognitive monitoring. This is reflected in both our model and our experiments; we address these in turn.

Assumption of Perfect Monitoring

In our model, we made the substantial (and unrealistic) simplifying assumption that the meta-level process has direct access to the state of the object-level process, that is, the amount of recall progress. This form of monitoring can be seen as a simplified form of Koriat's (1993) accessibility model. Specifically, our simplification does not allow for the possibility that incorrect partial information is recalled, which eliminates any divergence between the feeling of knowing and recall progress. While this assumption greatly simplified the implementation of the model and has precedence in previous metamemory models (Hu et al., 2019; Suchow & Griffiths, 2016), perfect monitoring is intuitively implausible and it is inconsistent with the sensitivity of feeling-of-knowing judgments to spurious information generated during recall (Koriat, 1993) and manipulations that do not affect accuracy (discussed below).

Beyond imperfect monitoring of recall progress, the meta-level process could rely on other sources of information about a memory's strength. Indeed, this possibility has been formalized and supported by static models in which a meta-level and object-level signal are imperfectly correlated (Fleming & Daw, 2017; Hu et al., 2021; Y. Jang et al., 2012). Intuitively, the meta-level signal in these models contains some information about recall progress (or "processing experience") and some information about the underlying memory strength from other sources (e.g., cue familiarity, see below), with both components corrupted by independent noise. Our model can also be seen as a special case of this class of models, where the amount of information from other sources and the noise are both fixed to zero. Allowing the noise to be nonzero would result in a formalization of Koriat's (1993) accessibility model, in which metamemory is a noisy readout of actual recall progress. Further allowing additional sources of information besides recall would yield a complete model in which control is jointly informed by recall progress and ancillary cues.

Unfortunately, incorporating these generalizations into an optimal dynamic model presents a major technical challenge. Specifically, decoupling the recall progress and monitoring signals makes the meta-level belief state depend on the full trajectory of signals (as opposed to just the sum). This makes the optimal policy intractable to solve.[11] Nevertheless, exploring the implications of imperfect monitoring and cue familiarity for meta-level control (with suitable approximations to the optimal policy) is a critical direction for future research.

Multiple Potential Sources of Monitoring Signal

As detailed above, we assumed that the meta-level process has direct access to the object-level process. Critically, however, we made this assumption for purely technical reasons, not based on the conviction that people actually derive all (or any) of their metacognitive monitoring from direct access to recall progress. There are several possible sources for this information, and our experiment is incapable of distinguishing between them.

It is of course possible that people's monitoring truly does depend on (indirect) access to recall progress. Supporting this hypothesis is the observation that people in tip-of-tongue states are able to correctly recall some aspects of the memory target, such as the number of syllables (Brown & McNeill, 1966) or semantic associations (Koriat, 1993; Schacter & Worling, 1985). Importantly, it is unlikely that people can perfectly monitor the exact amount of progress (as our model can); instead, they may have access to a noisy or corrupted signal of true progress (Koriat, 1993; Koriat & Goldsmith, 1996).

Another likely source of monitoring information is *cue familiarity*, that is, the degree to which one feels that they recognize the prompt or question that triggered the memory search (Metcalfe et al., 1993). In our experiment, participants could find some images more familiar than others, and this could correlate with their ability to recall the associated words. Although all images were displayed the same number of times, participants could have processed them at different depths (e.g., due to lapses in attention or properties of the images themselves); this would likely increase both familiarity and recall rates. Much evidence supports the idea that metacognitive judgments are based on cue familiarity. For example, participants give higher feeling-of-knowing judgments to arithmetic problems that visually resemble previously seen problems (Reder & Ritter, 1992). Similarly, priming words in a question increases feeling-of-knowing without increasing recall (Reder, 1987, 1988; Schwartz & Metcalfe, 1992). Indeed, this fact, along with the observation that priming the memory target does not increase the feeling of knowing, has led some researchers to suggest that the feeling of knowing does not depend at all on partial recall (Reder & Ritter, 1992; Schwartz & Metcalfe, 1992; but cf. Jameson et al., 1990; Narens et al., 1994).

A third possible source of monitoring—which we have not seen discussed in the literature, but is likely present in our task—is memory itself. That is, people could remember their previous experience trying to remember the target for the displayed cue. Importantly, however, we did not provide feedback in the pretest trials. Thus, this monitoring strategy would require participants to remember an earlier metacognitive evaluation of the accuracy of that previous response (or, less likely, to generate such an evaluation on the fly).[12] This strategy is thus not any "less metacognitive" than the others (indeed, it may be the most sophisticated of the three we have considered). We think that such memory-based monitoring is likely to be present in many naturalistic settings outside of our task (consider, e.g., the feeling when you see someone whose name you

[11] To see why this is, note that when the recall progress is not directly observed, the agent must infer a posterior distribution over it. This posterior must account for the continual observation that the boundary has not been crossed, and the likelihood of that event depends on the order of the signals. For example, if a large positive signal is followed by a large negative signal, the fact that the threshold was not crossed after the first signal indicates that the large positive signal was not accompanied by a similarly large change in progress. If the order is reversed, no such inference is drawn, and so the estimated recall progress across the two-time steps will be higher.

[12] On the other hand, such an evaluation would be trivial (and perhaps not metacognitive) in a case when they were unable to provide any reasonable answer. To address the possibility that the memory of such events was driving meta-level control, we recomputed pretest accuracy using only correct and intrusion trials (cases that would be difficult to distinguish without metacognition). We found that this revised pretest accuracy still predicted response times on skip trials in Experiment 1 ($B = 0.301$, 95% CI [0.167, 0.435], $t(228.0) = 4.39$, $p < .001$), albeit slightly less strongly ($B = .404$ for the main analysis).

can never remember). Future work should attempt to identify empirical evidence distinguishing it from other kinds of monitoring.

While the possible sources of monitoring information are often presented as competing theories (Jameson et al., 1990; Narens et al., 1994; Reder & Ritter, 1992; Schwartz & Metcalfe, 1992), we think it is likely that our metamemory monitoring system draws on multiple sources of information. The relative importance of each may depend on the context or indeed on the stage of the recall process (e.g., cue familiarity in early stages and recall progress in later stages; Nhouyvanisvong & Reder, 1998). This possibility further supports the development of models that can account for multiple sources of monitoring information, as discussed above.

Critically, although our results do not help to differentiate between these different possible sources of monitoring, this in no way undermines our core claim, that people are capable of monitoring the strength of memory and using this information to adaptively control their recall efforts.

Extensions

Beyond accounting for imperfect monitoring, as discussed above, there are several ways our model could be extended in future work. We highlight three possibilities below.

Free Recall

Although we have focused on cued recall, the decision about when to terminate a search is present in almost all naturalistic recall tasks, including, for example, free recall. The basic principle of our model is that this decision should be made by balancing the benefit and probability of recall with the cost of search, a principle we inherit from Anderson and Milson (1989). This is in contrast to existing models of free recall such as the context maintenance and retrieval model, which assumes that the search is terminated either when a target memory is retrieved or after a fixed amount of time has passed (Lohnas et al., 2015; Polyn et al., 2009). Alternatively, according to the search of associative memory model, the search is terminated after a fixed number of retrieval attempts that do not result in the recall of a new word (Raaijmakers & Shiffrin, 1981). In both models, the decision to terminate recall is based on a general stopping rule independent of the state of the memory system. Integrating a rational metacognitive stopping rule into these models is an interesting direction for future research. Empirically validating such a model will likely require additional data, as participants in free recall experiments are typically given a fixed amount of time to recall as many items as possible; the decision to terminate the search cannot be observed in this setting.

Recognition Memory

Our model could also be applied in the context of recognition memory (Anderson & Bower, 1972; Murdock, 2006; Ratcliff, 1978; Yonelinas, 2002). In the simplest recognition tasks, a person must simply state whether they recognize a target or not. The critical difference from cued recall is that giving a response does not require actually recalling a target. We modeled this constraint as a fixed threshold associated with the recall. Thus, to model recognition memory, one could simply remove that fixed threshold, replacing it with an always-available action to indicate recognition (analogous to the STOP action). This would make the model structurally identical to evidence accumulation models of perceptual decision making (as discussed above). A more complex form of recognition memory has been studied in the context of eyewitness identification from lineups (Wixted et al., 2018). Here, one must both identify the most likely target from a set of distractors and also determine that the identified target is indeed recognized. This requires one to simultaneously evaluate relative memory strength (as in Experiment 2) and absolute memory strength (as described above). Such a task could thus be modeled using a multiple target model to the one we used in Experiment 2, with the addition of the recognition action described above.

Self-Generated Cues

In Experiment 2, we considered the problem of arbitrating between multiple externally provided memory cues. In more naturalistic settings, however, people may need to generate their own cues. Indeed, in some cases, people generate information that is incidental to the information they are searching memory for, for example, recalling where people live in an attempt to recall their names (M. D. Williams & Hollan, 1981). These incidental pieces of information can then be used as “stepping stones on the way to the sought-after target” (Koriat, 2000, p. 334). This metaphor highlights the sense in which memory recall is a sequential decision problem. In our framework, each generated cue would correspond to a meta-level action, with some actions serving only to bring one to a mental state where one can generate a more useful cue. This suggests a fascinating direction for further research: How do people know which “stones” are going in the right direction?

Conclusion

In this article, we have developed and experimentally validated a model of optimal meta-level control for memory recall. This model can be seen as a union of three influential frameworks in cognitive science: rational analysis (Anderson, 1990), the two-process model of metacognition (T. O. Nelson & Narens, 1990), and reinforcement learning (Dayan & Daw, 2008). Concretely, we characterized a rational metamemory system as the optimal policy for a MDP in which a meta-level agent monitors and controls its object-level environment in order to maximize reward. Although here we have focused on metamemory, we are optimistic that this approach could be applied to model metacognition in other domains as well. We believe that dynamic metacognitive processes such as the one studied here are a critical but understudied feature of human cognition. We thus hope that our work will encourage other researchers to further develop a rich, computational understanding of these important processes.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*(3), 409–429. https://doi.org/10.1037/0033-295X.98.3.409

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, *79*, 97–123. https://doi.org/10.1037/h0033773

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*(4) 703–719. https://doi.org/10.1037/0033-295X.96.4.703

Bennett, S. T., Benjamin, A. S., & Steyvers, M. (2017). A Bayesian model of knowledge and metacognitive control: Applications to opt-in tasks. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the annual conference of the Cognitive Science Society* (pp. 1623–1628). The Cognitive Science Society.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*, 281–305. https://doi.org/10.5555/2188385.2188395

Binz, M., Gershman, S. J., Schulz, E., & Endres, D. (2022). Heuristics from bounded meta-learned inference. *Psychological Review*, *129*(5), 1042–1077. https://doi.org/10.1037/rev0000330

Bitzer, S., Park, H., Blankenburg, F., & Kiebel, S. J. (2014). Perceptual decision making: Drift-diffusion model is equivalent to a Bayesian model. *Frontiers in Human Neuroscience*, *8*, Article 102. https://doi.org/10.3389/fnhum.2014.00102

Boldt, A., Blundell, C., & De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. *Neuroscience of Consciousness*, *2019*(1), Article niz004. https://doi.org/10.1093/nc/niz004

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 325–337. https://doi.org/10.1016/S0022-5371(66)80040-3

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–459. https://doi.org/10.1037/0033-295X.100.3.432

Callaway, F., Jain, Y. R., van Opheusden, B., Das, P., Iwama, G., Gul, S., Krueger, P. M., Becker, F., Griffiths, T. L., & Lieder, F. (2022). Leveraging artificial intelligence to improve people's planning strategies. *Proceedings of the National Academy of Sciences*, *119*(12), Article e2117432119. https://doi.org/10.1073/pnas.2117432119

Callaway, F., Rangel, A., & Griffiths, T. L. (2021). Fixation patterns in simple choice reflect optimal information sampling. *PLOS Computational Biology*, *17*(3), Article e1008863. https://doi.org/10.1371/journal.pcbi.1008863

Callaway, F., van Opheusden, B., Gul, S., Das, P., Krueger, P. M., Griffiths, T. L., & Lieder, F. (2022). Rational use of cognitive resources in human planning. *Nature Human Behaviour*, *6*(8), 1112–1125. https://doi.org/10.1038/s41562-022-01332-8

Castel, A. D. (2007). The adaptive and strategic use of memory by older adults: Evaluative processing and value-directed remembering. In A. S. Benjamin & B. H. Ross (Eds.), *Psychology of learning and motivation* (Vol. 48, pp. 225–270). Academic Press.

Chen, H., Chang, H. J., & Howes, A. (2021). Apparently irrational choice as optimal sequential decision making. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(1), 792–800. https://doi.org/10.1609/aaai.v35i1.16161

Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(1), 142–150. https://doi.org/10.1037/0278-7393.18.1.142

Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective and Behavioral Neuroscience*, *8*(4), 429–453. https://doi.org/10.3758/CABN.8.4.429

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), 105–110. https://doi.org/10.1038/nn.3279

Deroy, O., Spence, C., & Noppeney, U. (2016). Metacognition in multisensory perception. *Trends in Cognitive Sciences*, *20*(10), 736–747. https://doi.org/10.1016/j.tics.2016.08.006

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science*, *29*(5), 761–778. https://doi.org/10.1177/0956797617744771

Ditterich, J. (2006). Stochastic models of decisions about motion direction: Behavior and physiology. *Neural Networks*, *19*(8), 981–1012. https://doi.org/10.1016/j.neunet.2006.05.042

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, *32*(11), 3612–3628. https://doi.org/10.1523/JNEUROSCI.4010-11.2012

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–275). Lawrence Erlbaum.

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, *16*(4), 228–232. https://doi.org/10.1111/j.1467-8721.2007.00509.x

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition*, *20*(4), 374–380. https://doi.org/10.3758/BF03210921

Eakin, D. K. (2005). Illusions of knowing: Metamemory and memory under conditions of retroactive interference. *Journal of Memory and Language*, *52*(4), 526–534. https://doi.org/10.1016/j.jml.2005.01.009

Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological Review*, *124*(1), 91–114. https://doi.org/10.1037/rev0000045

Folke, T., Jacobsen, C., Fleming, S. M., & De Martino, B. (2016). Explicit representation of confidence informs future value-based decisions. *Nature Human Behaviour*, *1*(1)Article 0002. https://doi.org/10.1038/s41562-016-0002

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2213–2223. https://doi.org/10.1098/rstb.2012.0123

Frömer, R., Lin, H., Dean Wolf, C. K., Inzlicht, M., & Shenhav, A. (2021). Expectations of reward and efficacy guide cognitive control allocation. *Nature Communications*, *12*(1), Article 1030. https://doi.org/10.1038/s41467-021-21315-z

Fudenberg, D., Strack, P., & Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, *108*(12), 3651–3684. https://doi.org/10.1257/aer.20150742

Gershman, S. J. (2021). The rational analysis of memory. In A. Beckermann, B. P. McLaughlin, & S. Walter (Eds.), *Oxford handbook of human memory* (pp. 663–677). Oxford Academic.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278. https://doi.org/10.1126/science.aac6076

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, *108*(Suppl. 3), 15647–15654. https://doi.org/10.1073/pnas.1014269108

Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 311–325. https://doi.org/10.1037/0278-7393.7.5.311

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229. https://doi.org/10.1111/tops.12142

Gruneberg, M. M., & Monks, J. (1974). 'Feeling of knowing' and cued recall. *Acta Psychologica*, *38*(4), 257–265. https://doi.org/10.1016/0001-6918(74)90010-9

Gruneberg, M. M., Monks, J., & Sykes, R. N. (1977). Some methodological problems with feeling of knowing studies. *Acta Psychologica*, *41*(5), 365–371. https://doi.org/10.1016/0001-6918(77)90014-2

Gureckis, T. M., & Markant, D. B. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives on Psychological Science*, *7*(5), 464–481. https://doi.org/10.1177/1745691612454304

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*(4), 208–216. https://doi.org/10.1037/h0022263

Hay, N. (2016). *Principles of metalevel control*. University of California.

Hay, N., Russell, S., Tolpin, D., & Shimony, S. E. (2012). Selecting computations: Theory and applications. In N. de Freitas & K. Murphy (Eds.), *Proceedings of the twenty-eighth conference on uncertainty in artificial intelligence* (pp. 346–355). AUAI Press.

He, R., & Lieder, F. (2022). *Where do adaptive planning strategies come from?* https://doi.org/10.13140/RG.2.2.28966.60487

Hemmer, P., & Steyvers, M. (2009). A Bayesian account of reconstructive memory. *Topics in Cognitive Science*, *1*(1), 189–202. https://doi.org/10.1111/j.1756-8765.2008.01010.x

Howes, A., Duggan, G. B., Kalidindi, K., Tseng, Y.-C., & Lewis, R. L. (2016). Predicting short-term remembering as boundedly optimal strategy choice. *Cognitive Science*, *40*(5), 1192–1223. https://doi.org/10.1111/cogs.12271

Howes, A., Lewis, R. L., & Vera, A. (2009). Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review*, *116*(4), 717–751. https://doi.org/10.1037/a0017187

Hu, X., Luo, L., & Fleming, S. M. (2019). A role for metamemory in cognitive offloading. *Cognition*, *193*, Article 104012. https://doi.org/10.1016/j.cognition.2019.104012

Hu, X., Zheng, J., Su, N., Fan, T., Yang, C., Yin, Y., Fleming, S. M., & Luo, L. (2021). A Bayesian inference model for metamemory. *Psychological Review*, *128*(5), 824–855. https://doi.org/10.1037/rev0000270

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, *129*(2), 220–241. https://doi.org/10.1037//0096-3445.129.2.220

Jain, Y. R., Gupta, S., Rakesh, V., Dayan, P., Callaway, F., & Lieder, F. (2019). How do people learn how to plan? *Proceedings of the Annual Conference on Cognitive Computational Neuroscience*, *21*(9), 1148–1160. https://doi.org/10.1038/s41593-018-0210-5

Jameson, K. A., Narens, L., Goldfarb, K., & Nelson, T. O. (1990). The influence of near-threshold priming on metamemory and recall. *Acta Psychologica*, *73*(1), 55–68. https://doi.org/10.1016/0001-6918(90)90058-N

Jang, A. I., Sharma, R., & Drugowitsch, J. (2021). Optimal policy for attention-modulated decisions explains human fixation behavior. *ELife*, *10*, Article e63436. https://doi.org/10.7554/eLife.63436

Jang, Y., Wallsten, T. S., & Huber, D. E. (2012). A stochastic detection and retrieval model for the study of metacognition. *Psychological Review*, *119*(1), 186–200. https://doi.org/10.1037/a0025960

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, *84*(6), 1329–1342. https://doi.org/10.1016/j.neuron.2014.12.015

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, *324*(5928), 759–764. https://doi.org/10.1126/science.1169405

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*(4), 609–639. https://doi.org/10.1037/0033-295x.100.4.609

Koriat, A. (2000). Control processes in remembering. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 333–346). Oxford University Press.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517. https://doi.org/10.1037/0033-295X.103.3.490

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. https://doi.org/10.1038/nn.2635

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108*(33), 13852–13857. https://doi.org/10.1073/pnas.1101328108

Lachman, J. L., Lachman, R., & Thronesbery, C. (1979). Metamemory through the adult life span. *Developmental Psychology*, *15*(5), 543–551. https://doi.org/10.1037/0012-1649.15.5.543

Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nature Neuroscience*, *18*(8), 1159–1167. https://doi.org/10.1038/nn.4064

Lewis, R. L., Howes, A., & Singh, S. (2014). Computational rationality: Linking mechanism and behavior through bounded utility maximization. *Topics in Cognitive Science*, *6*(2), 279–311. https://doi.org/10.1111/tops.12086

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, Article e1. https://doi.org/10.1017/S0140525X1900061X

Lieder, F., Shenhav, A., Musslick, S., & Griffiths, T. L. (2018). Rational metareasoning and the plasticity of cognitive control. *PLOS Computational Biology*, *14*(4), Article e1006043. https://doi.org/10.1371/journal.pcbi.1006043

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015). Expanding the scope of memory search: Modeling intralist and interlist effects in free recall. *Psychological Review*, *122*(2), 337–363. https://doi.org/10.1037/a0039036

Madan, C. R. (2021). Exploring word memorability: How well do different word properties explain item free-recall probability? *Psychonomic Bulletin & Review*, *28*(2), 583–595. https://doi.org/10.3758/s13423-020-01820-w

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. WH Freeman.

Metcalfe, J. (1993). Novelty monitoring, metacognition, and control in a composite holographic associative recall model: Implications for Korsakoff amnesia. *Psychological Review*, *100*(1), 3–22. https://doi.org/10.1037/0033-295X.100.1.3

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, *18*(3), 159–163. https://doi.org/10.1111/j.1467-8721.2009.01628.x

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 851–861. https://doi.org/10.1037/0278-7393.19.4.851

Miner, A. C., & Reder, L. M. (1994). A new look at feeling of knowing: Its metacognitive role in regulating question answering. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 47–70). The MIT Press.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

Moreno-Bote, R. (2010). Decision confidence and uncertainty in diffusion models with partially correlated neuronal integrators. *Neural Computation*, *22*(7), 1786–1811. https://doi.org/10.1162/neco.2010.12-08-930

Murdock, B. (2006). Decision-making models of remember-know judgments: Comment on Rotello, Macmillan, and Reeder (2004). *Psychological Review*, *113*, 648–655. https://doi.org/10.1037/0033-295X.113.3.648

Murphy, K. P. (2007). *Conjugate Bayesian analysis of the Gaussian distribution* (Tech. Rep.). University of British Columbia.

Narens, L., Jameson, K. A., & Lee, V. A. (1994). Subthreshold priming and memory monitoring. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 71–92). The MIT Press.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*(7), 1040–1046. https://doi.org/10.1038/nn.3130

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, *7*(4), 308–313. https://doi.org/10.1093/comjnl/7.4.308

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The university of south florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*(1), 109–133. https://doi.org/10.1037/0033-2909.95.1.109

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 26, pp. 125–173). Academic Press.

Nhouyvanisvong, A., & Reder, L. M. (1998). Rapid feeling-of-knowing: A strategy selection mechanism. In V. Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Cognitive and social dimensions* (pp. 35–52). Sage Publications.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154. https://doi.org/10.1016/j.jmp.2008.12.005

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, *117*(3), 864–901. https://doi.org/10.1037/a0019737

Polanía, R., Woodford, M., & Ruff, C. C. (2019). Efficient coding of subjective value. *Nature Neuroscience*, *22*(1), 134–142. https://doi.org/10.1038/s41593-018-0292-0

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*(1), 129–156. https://doi.org/10.1037/a0014420

Puterman, M. L. (2014). *Markov decision processes: Discrete stochastic dynamic programming*. Wiley.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*(2), 93–134. https://doi.org/10.1037/0033-295X.88.2.93

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356. https://doi.org/10.1111/1467-9280.00067

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481. https://doi.org/10.3758/BF03196302

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, *19*(1), 90–138. https://doi.org/10.1016/0010-0285(87)90005-3

Reder, L. M. (1988). Strategic control of retrieval strategies. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 22, pp. 227–259). Academic Press.

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 435–451. https://doi.org/10.1037/0278-7393.18.3.435

Russell, S., & Wefald, E. (1991). Principles of metareasoning. *Artificial Intelligence*, *49*(1–3), 361–395. https://doi.org/10.1016/0004-3702(91)90015-C

Savage, L. J. (1954). *The foundations of statistics*. Wiley.

Schacter, D. L., & Worling, J. R. (1985). Attribute information and the feeling-of-knowing. *Canadian Journal of Psychology / Revue canadienne de psychologie*, *39*(3), 467–475. https://doi.org/10.1037/h0080074

Schwartz, B. L. (2001). The relation of tip-of-the-tongue states and retrieval time. *Memory & Cognition*, *29*(1), 117–126. https://doi.org/10.3758/BF03195746

Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 1074–1083. https://doi.org/10.1037/0278-7393.18.5.1074

Schwartz, B. L., & Metcalfe, J. (2017). Metamemory: An update of critical findings. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (2nd ed., pp. 423–432). Academic Press. https://doi.org/10.1016/B978-0-12-809324-5.21056-0

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, *115*(4), 893–912. https://doi.org/10.1037/a0013396

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, *550*(7676), 354–359. https://doi.org/10.1038/nature24270

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Sobol', I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, *7*(4), 784–802. https://doi.org/10.1016/0041-5553(67)90144-9

Stewart, E. E. M., Ludwig, C. J. H., & Schütz, A. C. (2022). Humans represent the precision and utility of information acquired across fixations. *Scientific Reports*, *12*(1), Article 2411. https://doi.org/10.1038/s41598-022-06357-7

Suchow, J. W., & Griffiths, T. L. (2016). Deciding to remember: Memory maintenance as a markov decision process. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th annual conference of the cognitive science society* (pp. 2063–2068). Cognitive Science Society.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Tajima, S., Drugowitsch, J., Patel, N., & Pouget, A. (2019). Optimal policy for multi-alternative decisions. *Nature Neuroscience*, *22*(9), 1503–1511. https://doi.org/10.1038/s41593-019-0453-9

Tavares, G., Perona, P., & Rangel, A. (2017). The attentional Drift Diffusion Model of simple perceptual decision-making. *Frontiers in Neuroscience*, *11*, Article 468. https://doi.org/10.3389/fnins.2017.00468

Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity and Bayesian inference. *Behavioral and Brain Sciences*, *24*(4), 629–640. https://doi.org/10.1017/s0140525x01000061

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592. https://doi.org/10.1037/0033-295X.108.3.550

van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. *ELife*, *7*, Article e34963. https://doi.org/10.7554/eLife.34963

van den Berg, R., Zou, Q., Li, Y., & Ma, W. J. (2023). No effect of monetary reward in a visual working memory task. *PLOS ONE*, *18*(1), Article e0280257. https://doi.org/10.1371/journal.pone.0280257

Vesonder, G. T., & Voss, J. F. (1985). On the ability to predict one's own responses while learning. *Journal of Memory and Language*, *24*(3), 363–376. https://doi.org/10.1016/0749-596X(85)90034-8

Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, *13*(1), 37–58. https://doi.org/10.1080/00140137008931117

Williams, M., Hong, S. W., Kang, M.-S., Carlisle, N. B., & Woodman, G. F. (2013). The benefit of forgetting. *Psychonomic Bulletin & Review*, *20*(2), 348–355. https://doi.org/10.3758/s13423-012-0354-3

Williams, M. D., & Hollan, J. D. (1981). The process of retrieval from very long-term memory. *Cognitive Science*, *5*(2), 87–119. https://doi.org/10.1207/s15516709cog0502_1

Wixted, J. T., Vul, E., Mickes, L., & Wilson, B. M. (2018). Models of lineup memory. *Cognitive Psychology*, *105*, 81–114. https://doi.org/10.1016/j.cogpsych.2018.06.001

Woodford, M. (2014). Stochastic choice: An optimizing neuroeconomic model. *American Economic Review*, *104*(5), 495–500. https://doi.org/10.1257/aer.104.5.495

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). *SUN database: Large-scale scene recognition from abbey to zoo* [Conference session]. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, United States. https://doi.org/10.1109/CVPR.2010.5539970

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310–1321. https://doi.org/10.1098/rstb.2011.0416

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Aly, M., Wang, W.-C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. https://doi.org/10.1002/hipo.20864

Yoo, A. H., Klyszejko, Z., Curtis, C. E., & Ma, W. J. (2018). Strategic allocation of working memory resource. *Scientific Reports*, *8*(1), Article 16162. https://doi.org/10.1038/s41598-018-34282-1

Zhang, Q., Griffiths, T., & Norman, K. A. (2022). Optimal policies for free recall. *Psychological Review*, *130*(4), 1104–1124. https://doi.org/10.1037/rev0000375

Appendix A

Supplementary Results

Table A1

Full Response Type Proportions

Pretest accuracy	Correct	Empty	Intrusion	Other	Time-out
0%	.012	.806	.072	.093	.016
50%	.481	.380	.057	.068	.014
100%	.929	.056	.006	.007	.002

Note. Each cell shows the proportion of each possible response type conditional on pretest accuracy in Experiment 1.

Figure A1

Distributions of the Number of Fixations on Each Trial in Experiment 2

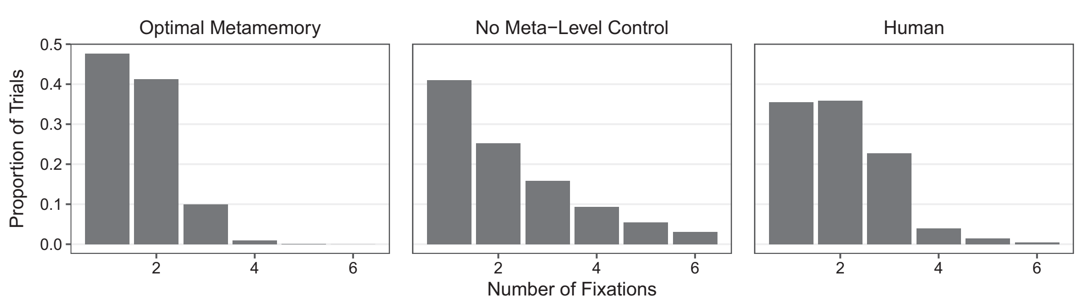
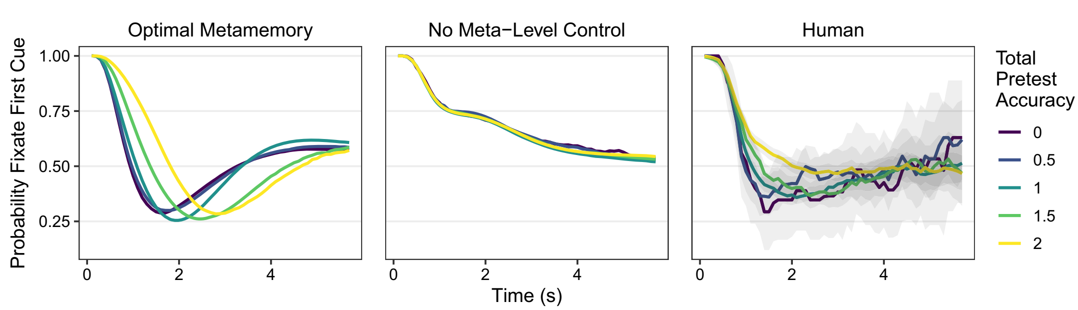


Figure A2

The Probability That the First-Seen Cue Is Currently Displayed Over the Course of the Trial, Split by the Total (Sum) Pretest Accuracy of Both Cues



Note. See the online article for the color version of this figure.

(*Appendices continue*)

Appendix B

Deviations From Preregistration

After preregistering and running the experiments, we discovered a flaw in the implementation of the lesioned models without meta-level control. In our first implementation, stopping times and fixation durations were sampled directly from the empirical distribution in the human data (discretized into 100 ms time steps). We reasoned that this would give the model the best chance of capturing the human behavior without a metacognitive component. However, we later realized that this approach effectively prevents the model from utilizing nondecision time, as any nondecision time would be added to a distribution that already perfectly matches the human data.

After realizing this, we implemented the more flexible lesioned model with stopping times and fixation durations drawn from arbitrary gamma distributions, as reported in the main text. Note that this model has more free parameters in the two-item case, thus we also decided to fit the lesioned model to data in Experiment 2 (contrary to our intention to use the fitted parameters from Experiment 1). The predictions of the original model for both experiments are shown in Figures D2–D5.

This analysis also revealed that two of the effects which we had believed to be evidence of meta-level control could in fact be produced by the model without meta-level control. The first such effect is shown in Figure D1B. The original logic of this plot was to look at how strength affects the decision to skip a trial over time, after factoring out the effect of strength on recall (which in turn prevents skipping). With a low or moderate nondecision time distribution, this effect cannot be captured by a model without meta-level control. However, when nondecision time is a sufficiently large portion of total response time, the effect can be produced through mechanistic means. Given this, we could not provide a useful interpretation of the effect and thus removed it from the main text.

The second effect is the prediction that final fixations are longer than nonfinal fixations (Figure 10A). We elected to keep this effect in the main text because its interpretation as an indication of rational metamemory was further supported by the additional analysis of the distribution of pretest accuracy for long fixations (Figure 10B).

Appendix C

Previously Preregistered Experiments

Before running the experiments presented in the main text, we ran another set of large experiments using the same experimental design. These were also preregistered. However, the results from these studies were not conclusive (for different reasons), so we reran them.

For Experiment 1, we initially planned to run 125 participants and to z score response times within participants before regressing them on pretest accuracy. This analysis yielded a marginally significant effect of pretest accuracy on response time for skip trials ($p = .062$). The original preregistration is available at https://aspredicted.org/ss8x3.pdf. While analyzing the data, we discovered that the z-scoring step actually reduced statistical power, as it selectively diminished the contribution of participants who showed the effect most robustly (and thus had more variable response times). We thus eliminated the z-scoring step, preregistered the slightly modified analysis, and reran the experiment with a target of 500 participants. Note that the mixed-effects analysis still accounts for individual variability in both average response times and sensitivity to pretest accuracy.

For Experiment 2 (which was actually run first), we had planned to use a definition of memory strength that combined response time and accuracy on the pretest trials. The logic was that a fast response indicated a stronger memory; thus, a higher response time in the pretest should predict less fixation time in the critical trials. However, we then discovered that response time on the pretest was actually *positively* correlated with fixation durations in the critical trials (contrary to our predictions). This may be due to factors other than memory strength, such as perceptual encoding, which would uniformly increase the amount of time spent looking at an item. Additionally, we had planned a different set of analyses, focusing on the duration of individual fixations as a function of the relative strength of the two cues, rather than breaking down the effect of the fixated and nonfixated cues, as we do now. The full preregistration is available at https://aspredicted.org/ss8x3.pdf. Given the extent of the changes we wished to make to the analysis, we preregistered the new analysis plan and reran the experiment.

As shown in the sections below, all of the results reported in the main text are also significant in this previous sample. We reproduce all the figures from the main text in Figures D6–D9.

Experiment 1

We recruited 124 participants and excluded 14 (11%) participants who did not provide a response on more than 90% of critical trials. This yielded 110 participants in our final analysis. Participants were faster to correctly recall targets that they reported greater confidence in ($B = -0.212$, 95% CI [-0.302, -0.122], $t(111.8) = -4.62$, $p < .001$), but slower to skip targets that they reported higher feeling-of-knowing for ($B = 0.474$, 95% CI [0.302, 0.646], $t(35.4) = 5.40$, $p < .001$). People were likewise faster to recall targets that they had previously recalled correctly ($B = -0.802$, 95% CI [-1.281, -0.322], $t(45.1) = -3.28$, $p = .002$). More importantly, they were also slower to skip such targets ($B = 0.511$, 95% CI [0.269, 0.753], $t(34.5) = 4.13$, $p < .001$).

Experiment 2

We recruited 459 participants and excluded 66 (14%) participants who failed to correctly recall a target on more than 50% of critical trials. This yielded 393 participants in our final analysis. Participants spent substantially more time looking at cues that were stronger than the other available cue ($B = 0.187$, 95% CI [0.178, 0.196], $t(326.9) = 41.46$, $p < .001$). Their nonfinal fixations increased with the pretest

(*Appendices continue*)

accuracy of the fixated cue (B = 0.101, 95% CI [0.069, 0.133], $t(219.9) = 6.22$, $p < .001$), and decreased with the pretest accuracy of the nonfixated cue (B = −0.432, 95% CI [−0.564, −0.300], $t(71.4) = -6.41$, $p < .001$; first fixations excluded). Final fixations were longer than their nonfinal fixations (B = 0.869, 95% CI [0.810, 0.929], $t(429.4) = 28.86$, $p < .001$). The probability of fixating the weaker memory significantly decreased with fixation duration (B = −1.376, 95% CI [−1.537, −1.215], $z = -16.78$, $p < .001$; logistic regression, excluding trials where the cues had equal pretest accuracy).

The final result is notable because it confirms the exploratory rational commitment analysis we developed using the new data set. Thus, although the result in the main text was not preregistered, the result presented above effectively is; we had finalized the analysis before examining the results it yielded with the old data set.

Appendix D

Optimizing the Lesioned Model to Predict Effects

To more conclusively determine which effects are inconsistent with the lesioned model, we conducted a thorough search of the parameter space to see if the model could produce each qualitative effect under any parameter setting. For each experiment, we sampled 100,000 parameter configurations and simulated 100,000 trials for each. From these, we excluded simulations with accuracy below 1% or above 99%, as these yield unreliable estimates for effects that condition on accuracy. For similar reasons, in Experiment 2, we excluded configurations for which fewer than 1% of trials had at least two fixations. For each unexcluded simulation, we then performed a standard linear regression corresponding to the regression we reported in the main text. Next, we selected the 100 configurations that produced the largest effect as estimated by the lower bound of a 95% confidence interval (this prevents selecting for models that simply produce highly variable estimates). If the lower confidence bound for any of these was larger than a "minimal interesting difference," then we concluded that the lesioned model could produce the effect. Note that this analysis occurred to us after running the experiments and was thus not preregistered.

Experiment 1

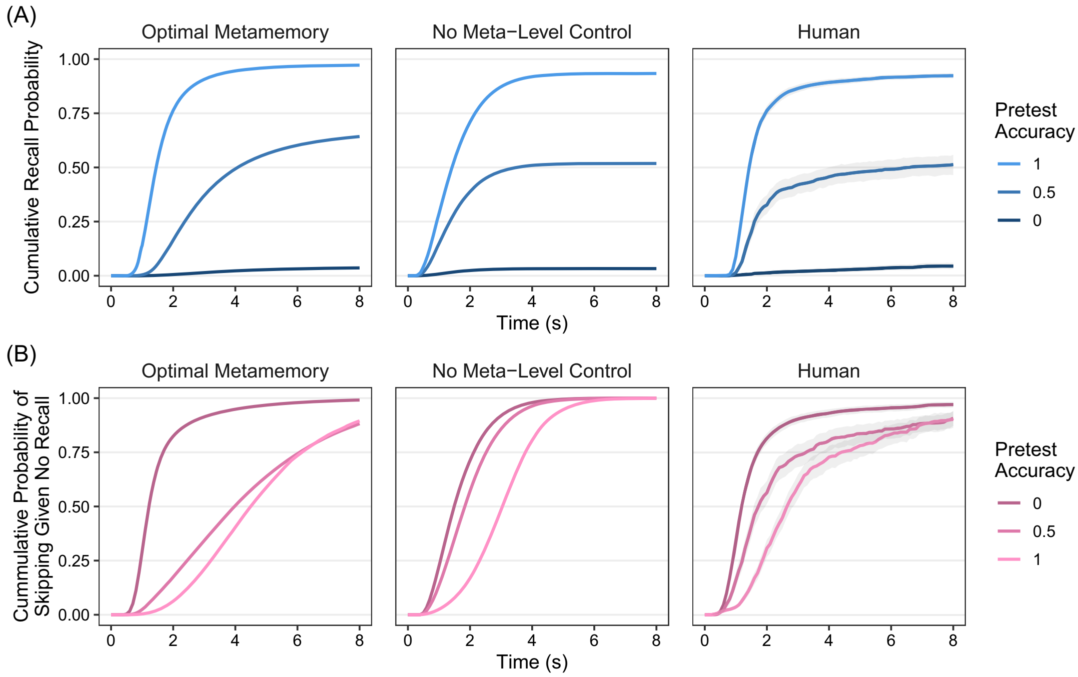
As expected, we found that the lesioned model could predict the observed negative effect of both judgments (B = −0.775; 95% CI [−0.778, −0.771]) and pretest accuracy (B = −6.031; 95% CI [−6.067, −5.996]) on response time for the correct trials. Note that these are the largest effect sizes we found; they are substantially larger than those seen in the data. Surprisingly, we also found that the lesioned model could predict a substantial positive relationship between judgment and response time on the skip trials (B = 0.601; 95% CI [0.590, 0.611]). However, this model failed to predict the negative relationship on correct trials (B = 0.053; 95% CI [0.051, 0.055]; note that B should be negative). No parameter configuration was able to predict the crossover pattern, with a negative relationship between judgment and response time on correct trials but a positive relationship on skip trials. Furthermore, no configuration was able to predict the positive relationship between pretest accuracy and response time on skip trials (even allowing the relationship for correct trials to be positive).

Experiment 2

Consistent with Figures 8A and 10A, the lesioned model was able to capture the overall proportion (B = 0.314; 95% CI [0.308, 0.320]) and the long final fixation (B = 2.372; 95% CI [2.370, 2.374]) effects. Perhaps surprisingly, we found that the lesioned model was also capable of capturing both fixation duration effects (fixated: B = 0.114; 95% CI [0.108, 0.121]; nonfixated: B = 0.099; 95% CI [0.091, 0.107]). However, these configurations achieved accuracies of only 7.4% and 7.4%, respectively. The model was able to capture the effect through a selection effect. It only provided a correct response on the small percentage of trials in which it happened to sample long fixations on the strongest cues. Running the analysis again with the requirement that the model achieves at least 65% accuracy (compared to 83% in the human data), we found that no configuration could capture the nonfinal fixation duration effects.

(*Appendices continue*)

Figure D1
Time Course of Recall and Stopping

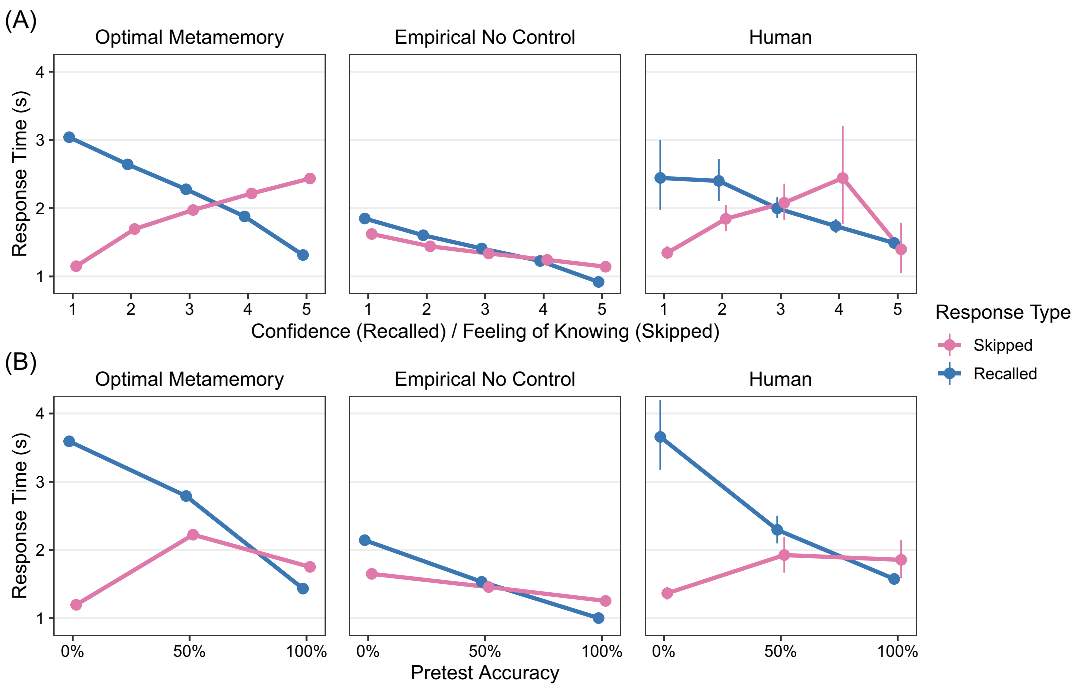


Note. (A) For each time point (in 100 ms steps), the proportion of trials on which the target was recalled before that point, grouping trials by the accuracy rate of the presented cue in the pretest phase. (B) For each time point, the proportion of trials that were skipped before that point, conditioning on the fact that the target had not already been recalled. Lines show means of participant means, and ribbons show 95% bootstrapped confidence intervals over participant means. See the online article for the color version of this figure.

(*Appendices continue*)

Figure D2

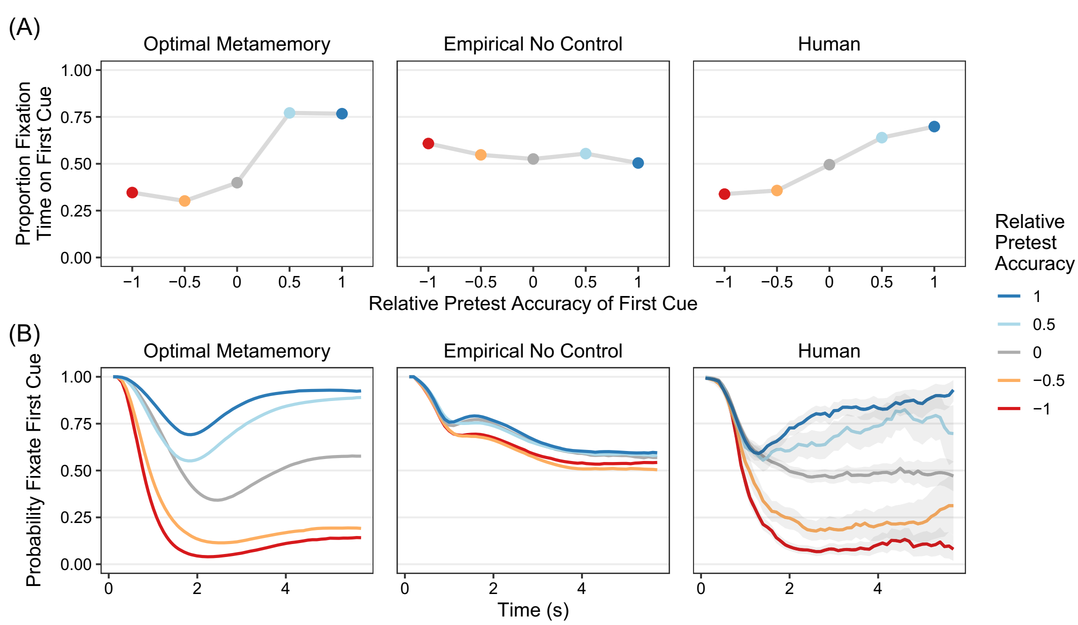
Alternative Version of Figure 5 Where the Lesioned Model Draws Stopping Times From the Empirical Distribution



Note. See the online article for the color version of this figure.

Figure D3

Alternative Version of Figure 8 Where the Lesioned Model Draws Stopping and Switching Times From the Empirical Distribution

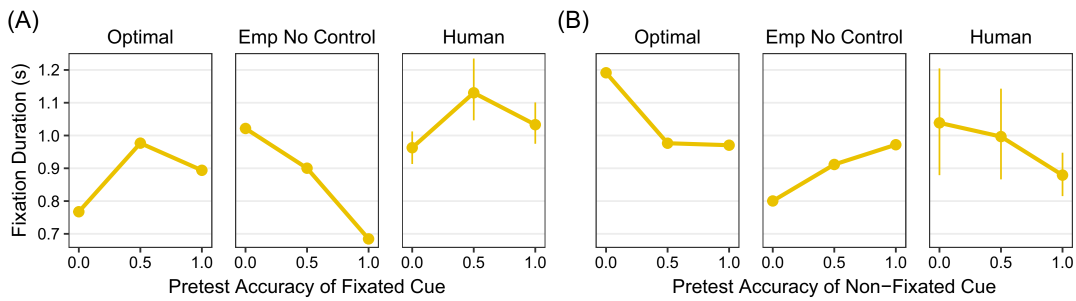


Note. See the online article for the color version of this figure.

(*Appendices continue*)

Figure D4

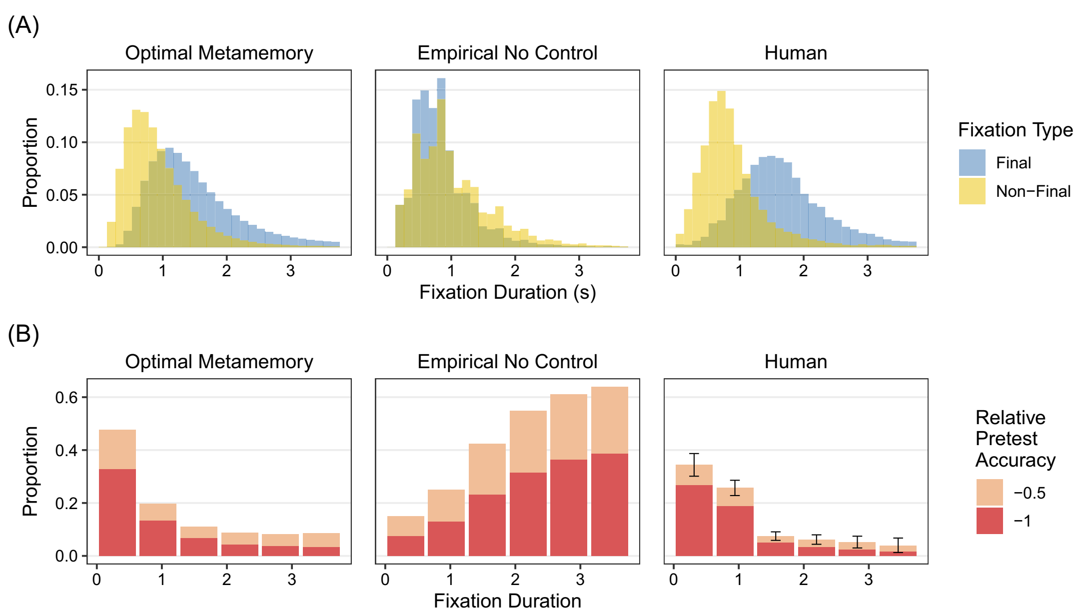
Alternative Version of Figure 9 Where Lesioned Model Draws Stopping and Switching Times From the Empirical Distribution



Note. See the online article for the color version of this figure.

Figure D5

Alternative Version of Figure 10 Where Lesioned Model Draws Stopping and Switching Times From the Empirical Distribution

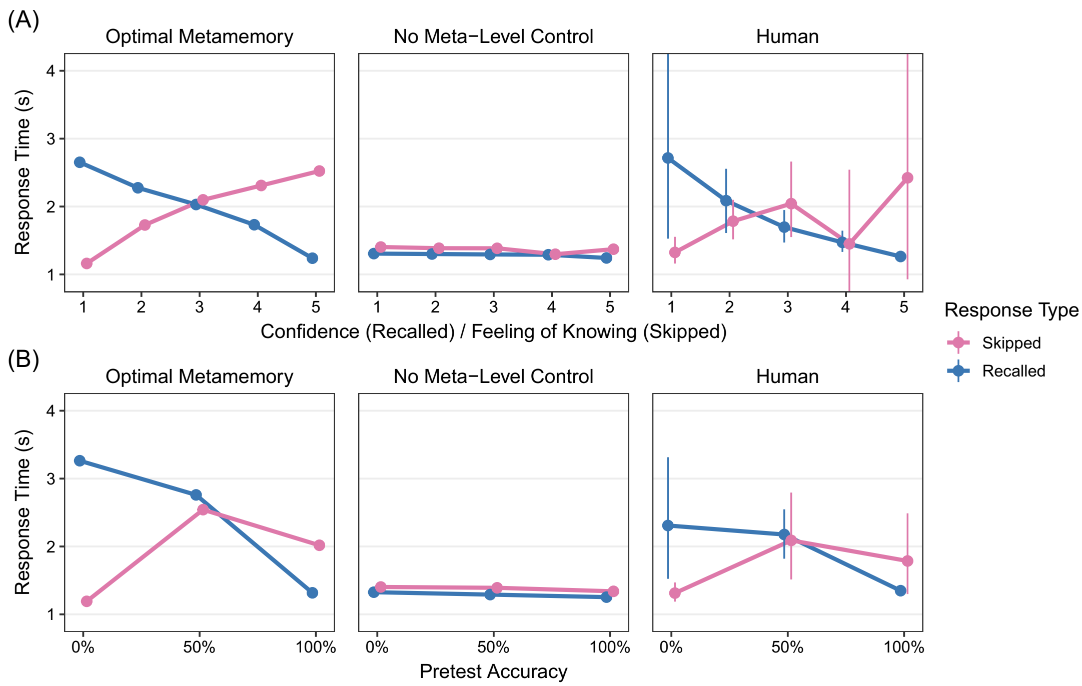


Note. See the online article for the color version of this figure.

(*Appendices continue*)

Figure D6

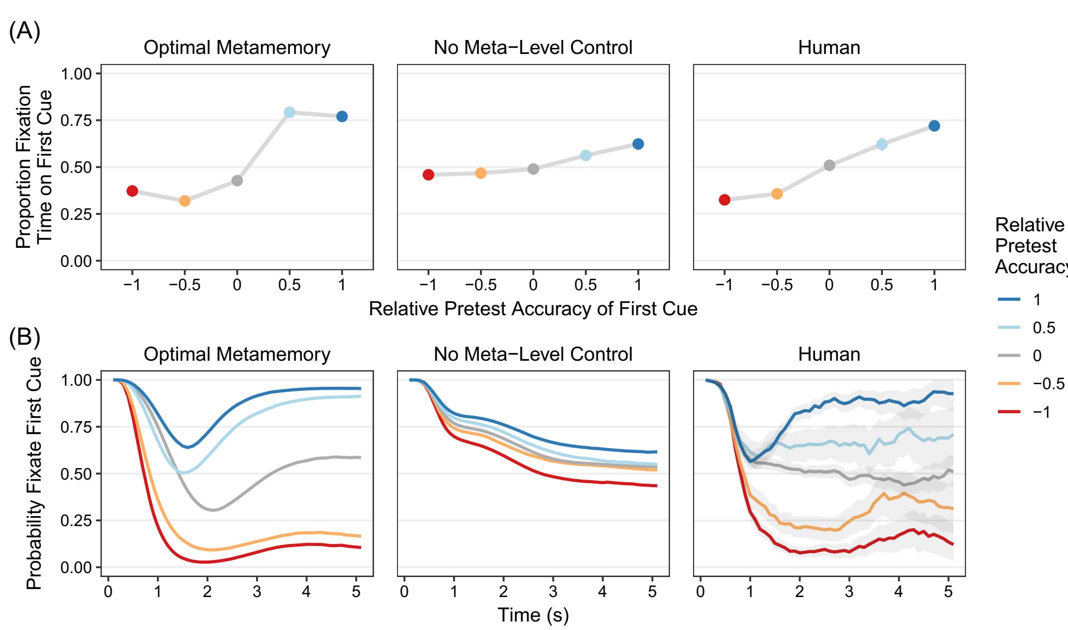
Figure 5 With Previous Experimental Data



Note. The models are fit to the data shown in the plot. We use the same axis limits as in the main text to facilitate comparison (the error bars extend beyond the plotted range). See the online article for the color version of this figure.

Figure D7

Figure 8 With Previous Experimental Data

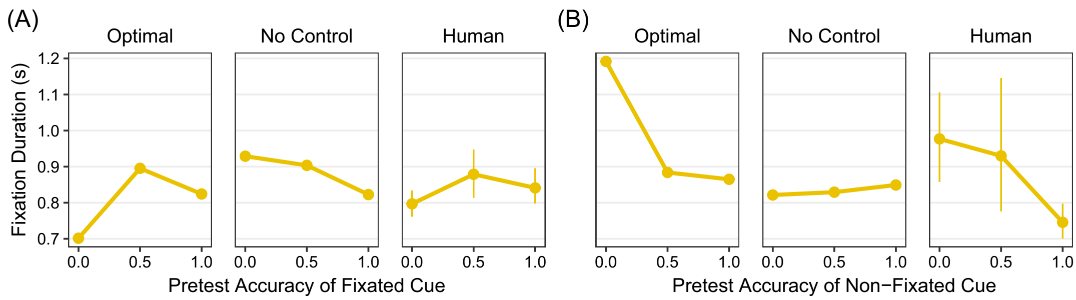


Note. See the online article for the color version of this figure.

(*Appendices continue*)

Figure D8

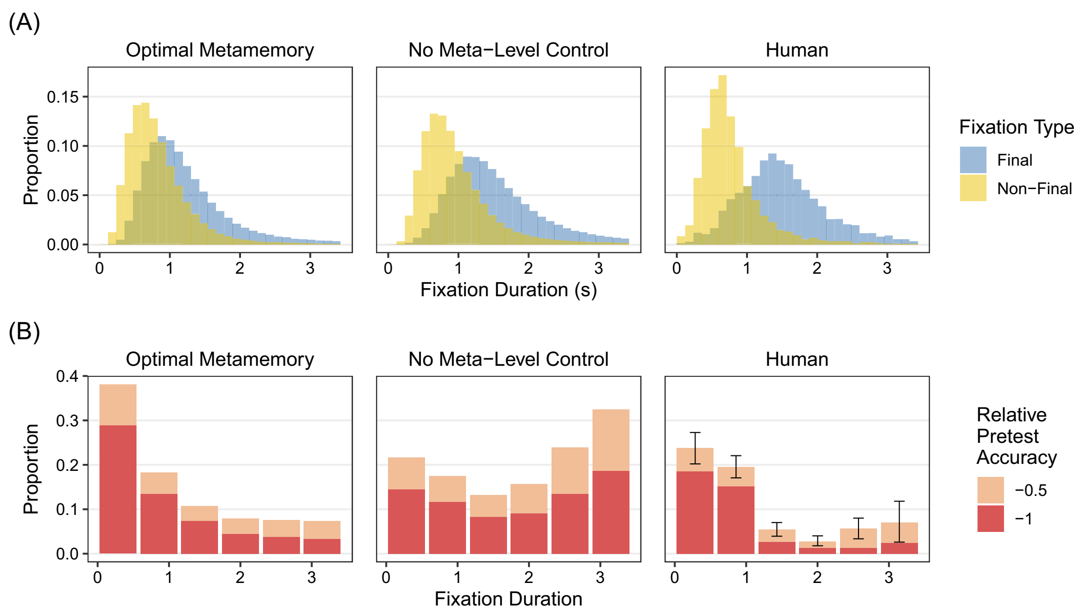
Figure 9 With the Old Previous Experimental Data



Note. See the online article for the color version of this figure.

Figure D9

Figure 10 With Previous Experimental Data



Note. See the online article for the color version of this figure.

Received September 16, 2022
Revision received April 11, 2023
Accepted June 10, 2023 ■