

# Memory Rehearsal and Belief Biases

John J. Conlon

First version: May 28, 2024

This version: June 24, 2026\*

## Abstract

I show experimentally that memory rehearsal generates systematic belief biases. Participants are randomized to have conversations about a subset of their experiences, which has large effects on what they recall a week later. Crucially, participants are naive about these effects: they take what they remember at face value when later forming beliefs about their experiences. Participants also make rehearsal choices without regard to their later distortionary effects. Intrinsic preferences for thinking about positive experiences instead drive rehearsal choices and therefore a positivity bias in recall and beliefs. This mechanism provides a new non-strategic channel through which overly optimistic beliefs arise.

---

\*Carnegie Mellon University: [jconlon@andrew.cmu.edu](mailto:jconlon@andrew.cmu.edu). I thank Chiara Aina, Kai Barron, Lukas Bolte, Lucas Coffman, Ben Enke, Thomas Graeber, Spencer Kwon, Muriel Niederle, Dev Patel, Pete Robertson, Chris Roth, Peter Schwardmann, Andrei Shleifer, Cassidy Shubatt, and Florian Zimmermann for helpful comments. Refine.ink was used to check the paper for consistency and clarity.

# 1 Introduction

Theories of motivated beliefs in economics posit that agents *want* to deceive themselves: they derive utility from optimism or self-confidence and so selectively seek out information, update their beliefs, or direct their attention to achieve it. This story is clearly right in some cases—for example, when people forgo a medical test because they would prefer to remain ignorant—but in other settings it requires a mixture of sophistication and naivete that can seem psychologically implausible: to succeed in biasing their future self’s beliefs, agents need to form and execute on a plan of self-deception only to later forget about it or fail to correct for it. Furthermore, they need *ex ante* to anticipate that their future self will be naive in this way. And yet, people do tend to end up with overly positive or self-serving beliefs. How?

In this paper, I study experimentally how selective *memory rehearsal* can generate overoptimism without agents attempting to bias their own beliefs. To motivate the design, I first describe a simple model of endogenous memory rehearsal. The agent derives in-the-moment utility from retrieving and thinking more about certain past experiences: for example, she might prefer to dwell on positive experiences and avoid ruminating on negative ones. But differential rehearsal of this type induces selection bias in her future recall by increasing the probability that more attended-to experiences will later come to mind. Whether this in turn distorts the agent’s beliefs (and the magnitude of any such bias) depends on if, and when, she is sophisticated or naive about this selection problem.

My experimental design allows me to measure the consequences for belief biases of naturalistic rehearsal of rich experiences, while maintaining tight control of the underlying environment. A first survey asks participants to list and then describe classes they took in college/high school that they felt they either did well or poorly in (henceforth, “successes” and “failures”). The number of successes vs failures is pre-determined, with the experiment eliciting a fixed number of each one-by-one in a random order.

Participants then wait a week before taking a second survey. One of the classes they listed in the first survey is chosen at random, without revealing its identity, and participants must estimate the percent chance that this random class was one of the successes. For an agent without memory constraints, this belief elicitation is trivial: participants directly provided the full set of classes and so could simply report as

their belief the true percent of the classes from the first survey that were successes. In practice, however, they must rely on what they can remember about the first survey in order to form their beliefs. Errors in participants' beliefs are tightly linked to their memory: participants are separately asked to free-recall (by listing the names of) these classes, and this measure of recall is highly correlated with their beliefs.

To study the effect of rehearsal, participants in the first survey have conversations with a large language model about three of the classes they named. These conversations are randomized to focus either on three successful classes (the *Successes* treatment) or three unsuccessful classes (the *Failures* treatment), and participants are aware that this is how the conversation topics are selected. These conversations have large effects on recall a week later. Participants in the second survey recall 45.9% of the classes they do not have conversations about. In contrast, they recall 72.1% of the classes they were randomized to have conversations about, a 26.2 percentage point increase ( $p < 0.01$ ). There are therefore significant corresponding differences across treatments in the number of successes and failures that participants recall: of the classes participants recall, 62.7% are successes in the *Successes* treatment but only 41.8% are successes in the *Failures* treatment (a 20.9 percentage point difference,  $p < 0.01$ ).

I then turn to my primary question: the implications of these rehearsal effects for belief biases. The model I describe makes the intuitive point that rehearsal can only bias beliefs if agents are naive *ex post*, when they are retrieving experiences to inform their beliefs, about the selection problem that earlier rehearsal induces in their memory. Such an agent takes what she recalls at face value, failing to account for the fact that her recall is distorted by earlier rehearsal. “Biases” in rehearsal (e.g., only having conversations about successes) therefore produce biases in her beliefs. In contrast, an *ex post* sophisticated agent, who realizes that some experiences were more likely to come to mind due to rehearsal, can “back out” accurate beliefs on average even from a biased memory.<sup>1</sup>

I find that participants appear almost wholly naive *ex post* about rehearsal's effect on their memory. Those in the *Successes* treatment, who had conversations about only successful classes, believe that 61.8% of the original classes were successes. This

---

<sup>1</sup>*A priori*, it is unclear whether we should expect this kind of sophistication about rehearsal. Agents appear naive *ex post* about other forces in memory such as cueing and associativity when forming beliefs (Bordalo et al. 2023), but rehearsal is perhaps the most well-known force that shapes human recall, as any student using flashcards to prepare for an exam can attest.

average belief is 16.5 percentage points lower ( $p < 0.01$ ) among participants in the *Failures* treatment whose conversations instead focused on unsuccessful classes. This treatment effect is not statistically distinguishable from the 20.9 percentage point effect on recall, so I cannot reject that participants are fully naive *ex post*. Participants in the *Successes* treatment are also 47.6 percentage points more likely ( $p < 0.01$ ) to bet that the randomly selected class is a success rather than a failure than those in the *Failures* treatment. In a sense, this *ex post* naivete represents an internal form of selection neglect (Enke 2020, Jin et al. 2021, Barron et al. 2024, Farina et al. 2024): participants take what they remember at face value, failing to account for the fact that factors beyond objective frequency contribute to recall.

My experiment exogenously varies rehearsal patterns to identify their effect on participants' memory and beliefs. But which experiences do participants choose to rehearse on their own? I elicit an incentivized measure of which classes participants would prefer to have conversations about, where I find that participants are dramatically more likely to choose to discuss successes than failures (44.6% vs 10.0%,  $p < 0.01$ ). Participants therefore endogenously choose rehearsal patterns that—given the large effects of rehearsal on beliefs—predictably lead them to have overly positive beliefs. I also analyze open-ended questions asking participants the reasons behind their rehearsal decisions. Over half of those who choose to discuss successes mention an in-the-moment intrinsic preference (i.e., not directed at manipulating their future beliefs) for thinking about times they were successful and for avoiding ruminating on failures.

Do participants fail to anticipate that their rehearsal choices will affect their later recall and beliefs? While I show that such *ex ante* naivete is not necessary for rehearsal preferences to produce belief biases (*ex post* naivete is sufficient), my experiment nonetheless provides suggestive evidence on this question. A randomly selected half of participants were warned immediately before making their rehearsal choice that they would later face monetary incentives for accurate beliefs about the fraction of classes that were successes vs failures. Of these, only 3.3% appear to mention anything about how their choice of conversation might affect these future beliefs or their future memory. Further, this manipulation does not have any effect on participants' rehearsal choices, nor does it attenuate the effect of rehearsal on later beliefs. These results suggest that, in addition to being naive *ex post* about the effect of rehearsal on their memory, many participants also appear naive *ex ante*.

These results show that intrinsic preferences for thinking about positive experiences drive rehearsal choices and therefore produce—through *ex post* naivete—overly positive beliefs as an unintended side effect. The literature on motivated cognition in economics has largely focused on models where sophisticated agents are intentionally optimizing their beliefs. Does distinguishing between these two mechanisms matter? I show that these two interpretations have diametrically opposed welfare implications: for example, a sophisticated agent with belief-based utility generally is *harmed* by an information intervention correcting her beliefs (by a simple revealed preference argument), while a naive agent with intrinsic rehearsal-based preferences is helped (because she is straightforwardly making an error while forming her beliefs *ex post*). Disentangling when each of these two mechanisms predominates is therefore relevant for policy makers considering whether de-biasing beliefs will tend to be beneficial. My results suggest that, at least in some cases, naive rehearsal-preferences may contribute to overoptimistic beliefs.

This paper contributes first to the literature in economics on motivated cognition. In addition to theoretical work (e.g., Bénabou & Tirole 2002, Brunnermeier & Parker 2005), many studies find that people hold overly optimistic or overconfident beliefs, and lab studies often find that belief updating is biased in a self-serving direction (e.g., Eil & Rao 2011, Schwardmann & van der Weele 2019, Saucet & Villeval 2019, Chew et al. 2020, Zimmermann 2020, Müller 2022, Möbius et al. 2022, Drobner 2022, Huffman et al. 2022, Amelio & Zimmermann 2023, Sial et al. 2023, Engelmann et al. 2024, Gödker et al. 2025). My results suggest that memory rehearsal may contribute to some of these patterns. People’s preference toward rehearsing positive memories also connects to work on selective attention, where agents prefer to focus on positive items even absent any effect on their beliefs (Loewenstein 1987, Caplin & Leahy 2001, Bolte & Raymond 2024).<sup>2</sup> This paper bridges these two literatures, as I show that rehearsing positive memories for the in-the-moment emotional benefits has unintended but predictable effects on belief biases in the future.

Next, this paper speaks to literatures on memory both in psychology and economics. Economists are increasingly studying the role of memory frictions in shaping beliefs. This work has largely focused on the role of associations and similarity in shaping what comes to mind (e.g., Bordalo et al. 2023, Enke et al. 2024, Graeber et

---

<sup>2</sup>The agents in these models are assumed to be fully rational (i.e., forward looking with correct expectations), precluding belief biases like those I study.

al. 2023, Augenblick et al. 2024), while rehearsal has received much less attention.<sup>3</sup> Work in psychology on rehearsal effects goes back at least to the nineteenth century. Most such studies focus on exploring how different variables (e.g., lag time, spacing, elaboration) increase or decrease the effect of rehearsal on later recall (Madigan 1969, Craik & Lockhart 1972, Roediger & Karpicke 2006, Cepeda et al. 2006, Bartsch et al. 2018) and how other patterns in memory are related to rehearsal (Rundus 1971, Brodie & Murdock 1977, Tan & Ward 2000).

## 2 Theoretical Framework

**Setup** I first describe a simple model to motivate the experimental design. There are three periods  $t \in \{0, 1, 2\}$ . In periods  $t = 1$  and  $t = 2$ , the agent has a memory database  $M$  of experiences that occurred in period  $t = 0$ , each with a characteristic  $y_m \in \mathbb{R}$ . For example,  $M$  might be the classes a student took in school, and  $y_m$  how successful she was in each. In period  $t = 1$  the agent decides how much to think about each experience; let  $a = (a_1, \dots, a_M)$ , with  $a_m \in \mathbb{R}$ , denote this vector of rehearsal choices. For simplicity, I assume that in period  $t = 1$ , the agent has perfect memory.

**Utility** In period  $t = 1$  the agent derives in-the-moment utility  $u_a = \sum_m \nu_m a_m - \frac{\phi}{2} a_m^2$  from her rehearsal choice: a constant marginal utility  $\nu_m$  from thinking about  $m$ —e.g., some experiences may simply be more pleasant to dwell on—minus a quadratic cost  $\phi$ . We can interpret  $a_m = 0$  as a baseline level of rehearsal, with  $a_m < 0$  or  $a_m > 0$  denoting less or more attention than that baseline.

In period  $t = 2$  the agent needs to form a belief about the average value  $y_m$ , perhaps to inform a decision. She derives utility from two sources. First, she derives instrumental utility  $u_c = -\frac{\beta}{2}(E_a[\bar{y}] - \bar{y})^2$  that increases in the accuracy of her beliefs, where  $\beta$  governs the cost of inaccurate beliefs. Second, I allow her to also potentially have belief-based utility  $u_b = \alpha E_a[\bar{y}]$  that increases in her subjective expectation of the average  $y_m$ : for example, she may prefer to hold high beliefs about her ability.

---

<sup>3</sup>Mullainathan (2002) and Bodoh-Creed (2020) incorporate rehearsal-like effects in models of exogenous memory retrieval, which is used in their models to generate autocorrelation in beliefs and thereby lasting distortions originally arising from other memory forces (i.e., associativity).

Equation 1 summarizes the agent’s utility:

$$u = \underbrace{\sum_m \nu_m a_m - \frac{\phi}{2} a_m^2}_{\text{Rehearsal-Based Utility } u_a} + \underbrace{\alpha E_a[\bar{y}]}_{\text{Belief-Based Utility } u_b} - \underbrace{\frac{\beta}{2} (E_a[\bar{y}] - \bar{y})^2}_{\text{Instrumental Utility } u_c} \quad (1)$$

**Beliefs by Sampling** To form her period-2 belief  $E_a[\bar{y}]$ , the agent *samples* experiences from her memory database  $M$  and takes a weighted average of their  $y_m$  values. She draws  $N$  iid samples (with replacement), and I study the limit as  $N \rightarrow \infty$ .<sup>4</sup> Let  $r(m, a) \in [0, 1]$  be the probability she recalls experience  $m$  given  $a$ , and  $\omega(m, a)$  the weight she gives  $m$  conditional on recalling it. Her beliefs are then given by equation 2:

$$E_a[\bar{y}] = \frac{1}{N} \sum_{n=1}^N y_{m(n)} \cdot \omega(m(n), a) \xrightarrow{P} \sum_{m \in M} y_m \cdot r(m, a) \cdot \omega(m, a) \quad (2)$$

**Rehearsal Effects and *Ex Post* Sophistication** The probability  $r(m, a)$  of recalling experience  $m$  in period 2 is higher the more attention  $a_m$  it received in period 1 (and the less other experiences received); for tractability, let  $r_m \propto \exp\{\gamma a_m\}$ , where  $\gamma$  governs the strength of rehearsal effects.

When forming beliefs, the agent tries to correct for any bias in her memory. She does so using inverse-probability weighting: she sets  $\omega(m, a) = \frac{1}{M \tilde{r}_2(m, a)}$ , where  $\tilde{r}_2(m, a)$  is her period-2 belief about the likelihood that  $m$  would come to mind. That is, she down-weights experiences she believes she was more likely to recall and up-weights those she believes were less likely to come to mind. A crucial question is whether the agent realizes how much rehearsal affects her recall. Let  $\tilde{r}_2(m, a) \propto \exp\{\lambda_2 \gamma a_m\}$ , where  $\lambda_2 \in [0, 1]$  is her (dogmatic) belief about the strength of rehearsal effects relative to the truth:  $\lambda_2 = 1$  is full *ex post* sophistication and  $\lambda_2 = 0$  full naivete. Proposition 1 then follows (all proofs in Appendix D):

**Proposition 1** *Rehearsal distorts beliefs if and only if it boosts recall of more attended-to experiences ( $\gamma > 0$ ) and agents are at least partially naive ex post ( $\lambda_2 < 1$ ):*

$$E_a[\bar{y}] \approx \bar{y} + (1 - \lambda_2) \cdot \gamma \cdot \text{Cov}(a_m, y_m) \quad (3)$$

---

<sup>4</sup>As is standard in this literature (e.g., Bordalo et al. 2023), the agent “double counts” experiences that come to mind more than once, so that beliefs do not converge to the truth as  $N \rightarrow \infty$ ; this is a tractable way to model experiences getting more weight in beliefs when they are more likely to be recalled.

**Rehearsal Choice** How does the agent choose which experiences to rehearse in period 1? Recall that she gains in-the-moment utility from thinking more about experiences with high  $\nu_m$ , but whether she anticipates that doing so distorts her future recall depends on her *ex ante* sophistication. Let her period-1 belief about her period-2 recall probability be  $\tilde{r}_1(m, a) \propto \exp\{\lambda_1 \gamma a_m\}$ , so she perceives rehearsal effects as a fraction  $\lambda_1 \in [0, 1]$  of their true magnitude:  $\lambda_1 = 1$  is full *ex ante* sophistication and  $\lambda_1 = 0$  represents full naivete.<sup>5</sup> Proposition 2 characterizes her choice.

**Proposition 2** *The agent “myopically” maximizes her in-the-moment rehearsal-based utility unless she is both sophisticated ex ante ( $\lambda_1 > 0$ ) and naive ex post ( $\lambda_2 < 1$ ):*

$$\underbrace{a_m^* - \frac{\nu_m}{\phi}}_{\text{Myopic rehearsal}} \propto \underbrace{\frac{1}{\phi} \lambda_1 (1 - \lambda_2) (y_m - \bar{y}) \left( \alpha - \lambda_1 \beta (E_{a^*}[\bar{y}] - \bar{y}) \right)}_{\text{Correction due to anticipated belief effect}} \quad (4)$$

The intuition is straightforward. An *ex ante* naive agent does not realize her rehearsal affects future recall, so she simply maximizes her in-the-moment rehearsal utility. A fully sophisticated agent does the same thing, but for a different reason: she knows her future self will optimally “back out” the correct beliefs from whatever she recalls, so her rehearsal choices today do not affect her beliefs in the future. Only an agent who is sophisticated *ex ante* but naive *ex post* foresees that today’s rehearsal will shift tomorrow’s beliefs, and so distorts her choice away from what would maximize in-the-moment rehearsal utility.

None of the above conclusions rely on the agent having or not having belief-based utility. First,  $\alpha$  alters rehearsal choices only for an *ex ante*-sophisticated and *ex post*-naive agent, as only such an agent perceives there to be an opportunity to use rehearsal to distort her future beliefs. Second, notice that even a fully *ex ante* sophisticated ( $\lambda_1 = 1$ ) agent without belief-based utility ( $\alpha = 0$ ) acts on her rehearsal-based preferences. She accepts some future unwanted belief bias as the price of indulging her in-the-moment rehearsal utility.

Proposition 3 below then characterizes the agent’s beliefs in period 2 given her rehearsal choices in period 1.

---

<sup>5</sup>For simplicity, I assume the agent knows her future *ex post* sophistication  $\lambda_2$ , treating her future self as believing recall probabilities are proportional to  $\exp\{\lambda_1 \lambda_2 \gamma a_m\}$ .

**Proposition 3** *A sufficient condition for rehearsal-based preferences ( $\text{Cov}(y, \nu) \neq 0$ ) to generate belief biases is that agents be ex post naive ( $\lambda_2 < 1$ ). Ex ante naivete ( $\lambda_1 < 1$ ) is unnecessary, but amplifies any such biases:*

$$\underbrace{E_{a^*}[\bar{y}] - \bar{y}}_{\text{Belief bias}} = \frac{(1 - \lambda_2)\gamma}{\phi + \beta M \xi^2 \text{Var}(y)} \left[ \underbrace{\text{Cov}(y, \nu)}_{\text{Effect of rehearsal-based utility}} + \underbrace{\alpha \xi \text{Var}(y)}_{\text{Effect of belief-based utility}} \right] \quad (5)$$

$$\text{where } \xi = \lambda_1(1 - \lambda_2) \frac{\gamma}{M}$$

**Welfare Implications** What are the welfare implications of such belief biases? Consider the effect of changing the agent’s period-2 beliefs, for example through a de-biasing information intervention. Let  $u_2 = \alpha E_{a^*}[\bar{y}] - \frac{\beta}{2}(E_{a^*}[\bar{y}] - \bar{y})^2$  be her actual utility from these corrected beliefs, holding fixed her earlier rehearsal choice  $a^*$ . Proposition 4 summarizes.

**Proposition 4** *Assume  $\text{Cov}(y, \nu) \geq 0$ ,  $\alpha \geq 0$ , and  $\lambda_2 < 1$  (ex post naivete). Then,*

(i) *Rehearsal-based preferences increase the welfare gains from correcting beliefs:*

$$\frac{d}{d\text{Cov}(y, \nu)} \left[ \underbrace{-\frac{du_2}{dE_{a^*}}}_{\text{Welfare effect of reducing beliefs}} \right] > 0 \quad (6)$$

(ii) *Suppose  $\lambda_1(1 - \lambda_1) < \frac{\phi}{\beta \text{Var}(y) \frac{\gamma^2}{M}}$  and that the agent has belief-based but not rehearsal-based utility:  $\alpha > 0 = \text{Cov}(y, \nu)$ . Then correcting her beliefs is harmful:  $-\frac{du_2}{dE_{a^*}} < 0$ .<sup>6</sup>*

(iii) *Suppose the agent’s rehearsal-based preferences are sufficiently strong:  $\text{Cov}(y, \nu) > \frac{\alpha \phi}{(1 - \lambda_2)\beta \gamma}$ . Then correcting her beliefs is beneficial:  $-\frac{du_2}{dE_{a^*}} > 0$ .*

The welfare effect of correcting biased beliefs thus depends critically on their source. An agent with belief-based utility ( $\alpha > 0$ ) *wants* biased beliefs despite their cost (from  $\beta$ ), so correcting them harms her by a simple revealed-preference logic. In contrast, a naive agent who biased her beliefs as an accidental byproduct of her

---

<sup>6</sup>The assumption of not-too-partial *ex ante* naivete prevents the agent from “overshooting” her utility-maximizing beliefs (given her belief-based utility).

rehearsal-based preferences is simply making a mistake when she forms inaccurate beliefs, so correcting them helps her.

### 3 Experimental Design

The experiment was designed to mimic the key features of the model described above. A total of 370 participants recruited from Prolific completed each of the two parts of the study during January of 2024. Attrition between the two parts was low (4.3%) and balanced across treatments. See Appendix B for more details on recruitment and attrition, as well as screenshots of the experiment. One of the incentivized questions across the two surveys was randomly selected to determine participants’ bonus (Azrieli et al. 2018).

**Memory database** The experiment first elicits from participants a “memory database” for the purposes of the study. In particular, in the first of two surveys participants completed, they are asked to name and describe nine classes they took during college or high school (depending on whether they have a college degree or not). In particular, they are asked one-by-one to name classes that they felt they did well or poorly in “compared to your normal academic performance.” Henceforth I call these “successes” and “failures,” though they were never framed to participants in these terms. In a second survey a week later, participants are asked to recall which nine classes they named in the first survey as well as the share of these classes that were successes vs failures. This binary success vs failure variable represents the  $y_m$  characteristic from the model in Section 2. The ground truth for this variable is pre-determined: the survey randomly asks either for four successes and five failures or for five successes and four failures. These elicitations occur in a random order. In addition to the name of the class, the first survey also asks a few followup questions about each class, including the gender of the instructor who taught it, the year participants took it, whether it involved a final project/paper, whether it was part of their major (for college classes), and whether it was an elective class (for high school classes).

**Rehearsal of Experiences** After listing the nine classes in the first survey, participants were informed that they would be having conversations about three of them. Their conversation partner was described as a “chatbot” and was in fact GPT-4,

a large language model. These conversations consisted of GPT-4 asking natural-language questions about the class in question, with participants answering in their own words and then followup questions being asked in response to their answers. Each conversation stopped after either 7.5 minutes or 12 question-and-answer pairs, whichever was shortest. The average participant wrote 209 words per conversation, while on average GPT-4 wrote 189 words per conversation.

After learning they would be having these conversations, participants were asked whether they would prefer to have conversations about three successes, three failures, or a randomly selected three classes. To make these elicitations incentive-compatible, 5% of participants had their choice actually implemented. I drop this 5% of the sample because for them I do not have exogenous variation in rehearsal. For the remaining 95% of participants, the computer simply flipped a coin to decide whether they would discuss three successes or three failures. Participants knew this procedure and were also told whether their preference was implemented and, if not, whether they were randomized to have conversations about three classes they did well or poorly in (the *Successes* and *Failures* treatments). These three classes were chosen randomly from the corresponding set.

Participants were randomly either warned that they would later face incentives for recall and accurate beliefs or were not warned. For the half of participants who were not warned, no wording in the experiment suggested that the study was about memory. The other half of participants, directly before being asked their conversation preferences, were told (truthfully) that “later, we’ll ask you to remember some of the answers you just gave (e.g., in how many classes you named did you have a male professor, do well vs poorly, etc.). You’ll have a better chance of winning the bonus if you answer those questions right.” They then had to correctly answer a comprehension question reiterating these facts. Thus, these participants were aware both that the study was about memory and also were informed of the exact question they would be asked (about the share of classes that were successes vs failures).<sup>7</sup>

**Beliefs and Recall** After having their three conversations, the first survey of the study ended. A week later, participants were invited back to take a short second survey that elicited their beliefs about and recall of the original set of nine classes.

---

<sup>7</sup>Participants were also asked beliefs about the other elicited characteristics of classes, such as the share of classes they said had a male professor. This part of the warning was thus also truthful.

The second survey began by asking participants to “Think back to all the classes we had you name during the first survey you took, including **both** the classes you had conversations about **and** the ones you did not have conversations about.” Note that this phrasing not only eliminates any ambiguity about which items the survey is asking about, it also equally cues participants with both types of classes.

A block of questions then asked participants to make incentivized binary bets about whether a randomly selected one of these classes had each of the binary characteristics that the first survey asked about, including the main variable indicating whether they felt they did well or poorly in it. Participants could earn a \$2.00 bonus if their bet was correct. A separate block of questions asked probabilistic versions of these same questions: e.g., “What do you think is the percent chance that the randomly selected class is one you said you did... well in? poorly in?” Participants could earn a \$2.00 bonus if their answer was within five percentage points of the true percentage. These two blocks occurred in a random order.

After these two belief-elicitation blocks, participants faced an incentivized free-recall task. In particular, they were asked to write down the names of as many classes as they could that they remembered listing in the first survey. If the free-recall question was chosen to determine participants’ bonus, each correct response increased their chances of winning by 10 percentage points, while each incorrect response reduced this chance by 10 percentage points. Their free-recall answers were graded by having GPT-4 assess whether each class they listed in the recall task matched a class they originally listed in the first survey (they need not have word-for-word matched).

## 4 Results

**Beliefs and Recall** I begin by describing some basic patterns in the beliefs and recall data. For all participants, the true share of classes from the first survey that were successes was either four or five out of nine: that is, 44% or 56%. The left panel of Figure 1 shows the distribution of participants’ beliefs about this share in the second survey. We see a wide range of beliefs, indicating that participants cannot perfectly think back on the first survey and report the actual distributions of experiences from it. Indeed, in the free-recall task the average participant only remembers 4.9 of the

nine classes.<sup>8</sup> The middle panel of Figure 1 calculates for each participant the share of the classes they recalled that were successes vs failures, and then plots the distribution of this variable. We see wide variation in this measure of what participants recall. The right panel of Figure 1 shows that these two variables—beliefs and recall—are tightly linked in the cross-section: participants who recall more successes than failures believe the original set of experiences contained a greater share of successes.

**Rehearsal Effects on Recall** Next, I ask to what extent the rehearsal manipulation in the first survey affected recall in the second survey. Note that all participants listed nine classes in the first survey. Of these, either four or five were successes, and therefore five or four were failures. The rehearsal manipulation involved conversations about three classes, all successes in the *Successes* treatment or all failures in the *Failures* treatment. Thus, every participant had at least one success and one failure that they did *not* have a conversation about.

Recall that there are two sources of randomization that the experimental design allows me to exploit. First, participants randomly discuss either three successes or three failures depending on their treatment status. Second, within treatment, which classes the conversations focus on are chosen randomly from among all the successes (in the *Successes* treatment) or all the failures (in the *Failures* treatment). I start by looking at this second source of variation, comparing recall rates within treatment across experiences.

Panel A of Figure 2 shows recall rates within each treatment for different types of classes. First, we see that within the *Successes* treatment, participants recall 47.5% of the successes they do *not* have conversations about but 72.3% of the successes they do have conversations about. This difference of 24.8 percentage points is significant at the  $p < 0.01$  level. We see very similar within-treatment differences in recall rates for the *Failures* treatment. Such participants recall 46.4% of the failures they did not have conversations about but 71.9% of those they did have conversations about. This 25.5 percentage point effect is significantly distinguishable from zero ( $p < 0.01$ ) and indistinguishable from the 24.8 percentage point effect on successes in the *Successes*

---

<sup>8</sup>Throughout the results, unless otherwise noted, I focus on the classes that participants recall correctly. For 4.6% of entries in the free-recall task, participants wrote a class name that does not (according to GPT-4) correspond to one of the classes they listed in the first survey. Further, among recalled classes that did appear in the first survey 93.2% are correctly identified as being a success/failure. None of the results presented are sensitive to including/excluding these “false memories.”

treatment ( $p = 0.89$ ). Rehearsal also makes experiences come to mind *faster*, not merely more often: Panel B of Figure 2 shows that among recalled classes, those participants had conversations about are listed about one position earlier in the free-recall task ( $p < 0.01$  for all comparisons). These results are consistent with the memory-sampling framework of Section 2.

The above results showed effects of conversations on later recall at the class-by-class level. Of course, the two treatments differed in whether these conversations focused on successes or failures. We should therefore expect treatment effects across individuals on the total number of successes and failures that participants recall. Indeed, the first pair of bars in Figure 3 shows that those in the *Successes* treatment recall on average 2.9 successes, 0.6 more than those in the *Failures* treatment ( $p < 0.01$ ). Similarly, the second pair of bars shows that those in *Failures* recall 2.8 failures, 0.9 more than those in *Successes* ( $p < 0.01$ ). Transforming these numbers, the third pair of bars shows that 62.7% of the classes participants in *Successes* recall are successes, 20.9 percentage points more than those in *Failures* ( $p < 0.01$ ). Thus, shifting rehearsal toward successes in turn systematically shifts the composition of participants’ recall toward successes.<sup>9</sup>

**Beliefs and *Ex Post* Naivete** All of the results presented thus far are consistent with a fully sophisticated agent reacting rationally to her memory constraints when forming beliefs. I now turn to discussing whether rehearsal biases beliefs in addition to recall. That is, do the distortions of recall from rehearsal in turn distort participants’ beliefs about the share of classes that were successes vs failures? Proposition 1 stated that belief biases arise from rehearsal effects if and only if agents are naive *ex post* about them. That is, a sophisticated agent who understands that what she can recall partly reflects what she has rehearsed more can nonetheless form accurate beliefs (in expectation). A naive agent will instead take what she recalls at face value: i.e., as representative of the true distribution of her experiences.

The fourth pair of bars in Figure 3 shows participants’ average beliefs about the share of the original set of classes that were successes rather than failures. Those in *Successes* on average believe 61.8% of classes were successes, while those in *Failures* believe only 45.3% were successes. This treatment effect of 16.5 percentage points

---

<sup>9</sup>These effects primarily reflect the extensive margin—whether a class is recalled at all—rather than the intensive margin: whether, conditional on recalling a class, it is (mis)remembered as being a success or a failure (see Appendix C.1).

( $p < 0.01$ ) is 79% of the 20.9 percentage point effect on recall that we saw above (third pair of bars in Figure 3) and is statistically indistinguishable from it ( $p = 0.16$ ). I therefore cannot reject that agents are fully naive *ex post* about rehearsal’s effect on their recall. Rehearsal therefore biases beliefs in addition to distorting recall.

Recall that in addition to eliciting probabilistic beliefs, the experiment also asked participants to make a binary bet on whether the randomly selected class was a success or failure. The final pair of bars in Figure 3 shows very large treatment effects of the rehearsal manipulation on these bets. Those in *Successes* are 47.6 percentage points ( $p < 0.01$ ) more likely to bet that the random class is a success than those in *Failures*.<sup>10</sup>

**Endogenous Rehearsal Choices** Thus far, I have focused on the effect of exogenously assigning rehearsal patterns to participants in order to measure their effects on recall and beliefs. I now consider participants’ rehearsal preferences to understand what memory-based belief biases would arise endogenously. Recall that participants could express a preference for having conversations about either three successes, three failures, or three randomly selected classes. Participants are over four times as likely to choose to rehearse successes as failures (44.6% vs 10.0%,  $p < 0.01$ ). In effect, therefore, participants are “choosing” (if unintentionally) to have overly positive future beliefs, in that the rehearsal choices they make predictably lead to this bias.

A reasonable worry one might have is rehearsal may affect beliefs differently when it is endogenously chosen rather than exogenously assigned by the experiment. For example, participants could be more sophisticated *ex post* about rehearsal they choose than about rehearsal imposed on them by a coin flip. In Appendix C.3, I provide evidence against this potential confound. Many participants happened to be randomized to rehearse the classes they themselves chose, and I find no larger discrepancy between recall and beliefs—as we would expect from greater sophistication—among those participants ( $p = 0.93$ ).

**Suggestive Evidence on *Ex Ante* Sophistication** Are participants also naive *ex ante*—that is, do they fail to take into account future belief distortions when deciding what to rehearse? Proposition 3 shows that *ex ante* naivete is not necessary

---

<sup>10</sup>In Appendix C.2, I show similar results for other features of classes besides success vs failure: e.g., being randomized to discuss classes with more female professors increases beliefs about the frequency of such classes.

for rehearsal-based preferences to bias beliefs (*ex post* naivete suffices) but that it amplifies any such bias. Proposition 2 suggests a test: an *ex ante* sophisticated agent who knows she will later be incentivized for accurate beliefs should shade her rehearsal away from the rehearsal options (all successes or all failures) that would bias her beliefs. Recall that half of participants were randomly assigned to be warned that they would face incentives for accurate beliefs, in effect varying the perceived  $\beta$  in equation 4. In practice, this warning has no effect: warned participants choose to rehearse successes at the same (indeed, slightly higher) rate as those not warned (46.4% and 42.9%,  $p = 0.49$ , see Table A.IV). Nor do we find evidence for indirect effects, for example through warned participants offsetting the within-experiment conversations by rehearsing the other type of class on their own after the experiment: Figure 4 shows no differential treatment effects on recall, beliefs, or binary bets across the two groups ( $p > 0.10$  for all five comparisons). I thus find no evidence of *ex ante* sophistication in my setting.

**What Drives Rehearsal Choices** Because participants appear naive *ex ante* about rehearsal effects, strategic motives related to managing future beliefs seem not to determine rehearsal choices within my experiment.<sup>11</sup> To provide qualitative evidence on what does drive them, participants are asked in an open-ended question to explain the reason for their rehearsal choice. I then code the responses with a large language model (see Appendix C.4 for more detail). The majority (56%) of participants who chose to rehearse successes are coded as reporting an in-the-moment preference for thinking about positive vs negative experiences (e.g., “It’s more pleasant to think about times I performed well than times I did not”). In contrast, almost none of the participants (3.3%) who were warned about the future incentives for accurate beliefs report anything about how their choice of what to rehearse might impact future recall or beliefs. These results corroborate my interpretation of the results: participants simply enjoy thinking about positive experiences, which incidentally biases their future beliefs by shifting what they will be able to recall.

---

<sup>11</sup>One might worry instead that choices fail to respond to the accuracy incentives simply because those incentives are too small. But almost no participants (3.2%) say anything suggesting awareness that their conversations would affect what they later remember or believe (and none states the logic clearly), and the mix of stated considerations does not vary with whether participants were warned about the incentives. See Appendix C.4 for details.

## 5 Discussion

Two directions appear especially promising for future work. First, while this study focuses on top-down preferences for thinking about experiences, bottom-up forces play a key role in determining what grabs our attention independently of our preferences (Bordalo et al. 2022). Exploring how such factors might produce later belief biases by altering rehearsal patterns remains an important question for future work. One natural hypothesis—that would push against the top-down preferences I study—is that some mental health conditions like depression are partly defined as an inability to avoid ruminating on negative experiences. Combined with the rehearsal effects and naivete that my experiment documents, such tendencies could lead to a vicious cycle, with depression exacerbating low self-confidence and vice versa.

Second, the participants in my experiment appear almost fully naive about rehearsal effects both *ex ante* and *ex post*, but this cannot be universally true. Students use flashcards to study for tests, leveraging rehearsal effects to boost their academic performance. When is the effect of rehearsal on recall clear enough that agents take it into account, either exploiting it to improve/change their memories or accounting for it *ex post*? Answering this question would aid both in predicting when (and which) belief biases arise as well as in understanding how to design policies to counteract rehearsal’s distortionary effects.

## References

- Amelio, A., & Zimmermann, F. (2023). Motivated memory in economics-a review. *Games*. Retrieved from [www.econtribute.de](http://www.econtribute.de)
- Augenblick, N., Jack, B. K., Kaur, S., Masiye, F., & Swanson, N. (2024). Retrieval failures and consumption smoothing: A field experiment on seasonal poverty. *Working paper*.
- Azrieli, Y., Chambers, C. P., & Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126.
- Barron, K., Huck, S., & Jehiel, P. (2024, 8). Everyday econometricians: Selection neglect and overoptimism when learning from others. *American Economic Journal: Microeconomics*, 16, 162-198. Retrieved from <https://pubs.aeaweb.org/doi/10.1257/mic.20200030> doi: 10.1257/mic.20200030
- Bartsch, L. M., Singmann, H., & Oberauer, K. (2018, 7). The effects of refreshing and elaboration on working memory performance, and their contributions to long-term memory formation. *Memory and Cognition*, 46, 796-808. doi: 10.3758/s13421-018-0805-9
- Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *The quarterly journal of economics*, 117(3), 871–915.
- Bodoh-Creed, A. L. (2020). Mood, memory, and the evaluation of asset prices. *Review of Finance*, 24, 227-262. doi: 10.1093/rof/rfz001
- Bolte, L., & Raymond, C. (2024). *Emotional inattention*.
- Bordalo, P., Conlon, J. J., Gennaioli, N., Kwon, S. Y., & Shleifer, A. (2023, 12). Memory and probability. *The Quarterly Journal of Economics*, 138, 265-311. doi: 10.1093/qje/qjac031
- Bordalo, P., Gennaioli, N., & Shleifer, A. (2022). Salience. *Annual Review of Economics*, 1-42.
- Brodie, D. A., & Murdock, B. B. (1977). Effect of presentation time on nominal and functional serial-position curves of free recall. *Journal of Verbal Learning and Verbal Behavior*, 16, 5-7.
- Brunnermeier, M. K., & Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4), 1092–1118.
- Caplin, A., & Leahy, J. (2001). Psychological expected utility theory and anticipatory feelings. *Quarterly Journal of Economics*. Retrieved from <https://academic.oup.com/qje/article/116/1/55/1939090>
- Cepeda, N. J., Pashler, H., Vul, E., Rohrer, D., & Wixted, J. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132.
- Chew, S. H., Huang, W., & Zhao, X. (2020). Motivated false memory. *Journal of Political Economy*, 128, 3913-3939. doi: 10.1086/709971
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research 1. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.

- Drobner, C. (2022, 3). Motivated beliefs and anticipation of uncertainty resolution. *American Economic Review: Insights*, 4, 89-105. doi: 10.1257/aeri.20200829
- Eil, D., & Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3, 114-138. doi: 10.1257/mic.3.2.114
- Engelmann, J. B., Lebreton, M., Salem-Garcia, N. A., Schwardmann, P., & Weele, J. J. V. D. (2024). Anticipatory anxiety and wishful thinking. *American Economic Review*.
- Enke, B. (2020, 8). What you see is all there is. *Quarterly Journal of Economics*, 135, 1363-1398. doi: 10.1093/qje/qjaa012
- Enke, B., Schwerter, F., & Zimmermann, F. (2024). *Associative memory and belief formation*.
- Farina, A., Fréchette, G. R., Lizzeri, A., & Perego, J. (2024). The selective disclosure of evidence: An experiment. *Working paper*.
- Graeber, T., Roth, C., & Zimmermann, F. (2023). *Stories, statistics, and memory*.
- Gödker, K., Jiao, P., & Smeets, P. (2025). Investor memory. *The Review of Financial Studies*.
- Huffman, D., Raymond, C., & Shvets, J. (2022, 10). Persistent overconfidence and biased memory: Evidence from managers. *American Economic Review*, 112, 3141-3175. doi: 10.1257/aer.20190668
- Jin, G. Z., Luca, M., & Martin, D. (2021). Is no news (perceived as) bad news? an experimental investigation of information disclosure. *American Economic Journal: Microeconomics*, 13, 141-173. doi: 10.1257/mic.20180217
- Loewenstein, G. (1987, 9). Anticipation and valuation of delayed consumption. *The Economic Journal*, 666-684.
- Madigan, S. A. (1969). *Intraserial repetition and coding processes in free recall* (Vol. 8).
- Mullainathan, S. (2002). A memory-based model of bounded rationality. *The Quarterly Journal of Economics*. Retrieved from <https://academic.oup.com/qje/article/117/3/735/1932979>
- Möbius, M. M., Niederle, M., & Niehaus, P. (2022). Managing self-confidence. *Management Science*.
- Müller, M. W. (2022). *Selective memory around big life decisions*.
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning taking memory tests improves long-term retention. *Psychological Science*, 17.
- Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, 89, 63-77.
- Saucet, C., & Villeval, M. C. (2019, 9). Motivated memory in dictator games. *Games and Economic Behavior*, 117, 250-275. doi: 10.1016/j.geb.2019.05.011
- Schwardmann, P., & van der Weele, J. (2019, 10). *Deception and self-deception* (Vol. 3). Nature Research. doi: 10.1038/s41562-019-0666-7

- Sial, A. Y., Sydnor, J. R., & Taubinsky, D. (2023). Biased memory and perceptions of self-control. *NBER Working Paper 30825*. Retrieved from <http://www.nber.org/papers/w30825>
- Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1589-1625.
- Zimmermann, F. (2020). *The dynamics of motivated beliefs* (Vol. 110). American Economic Association. doi: 10.1257/aer.20180728

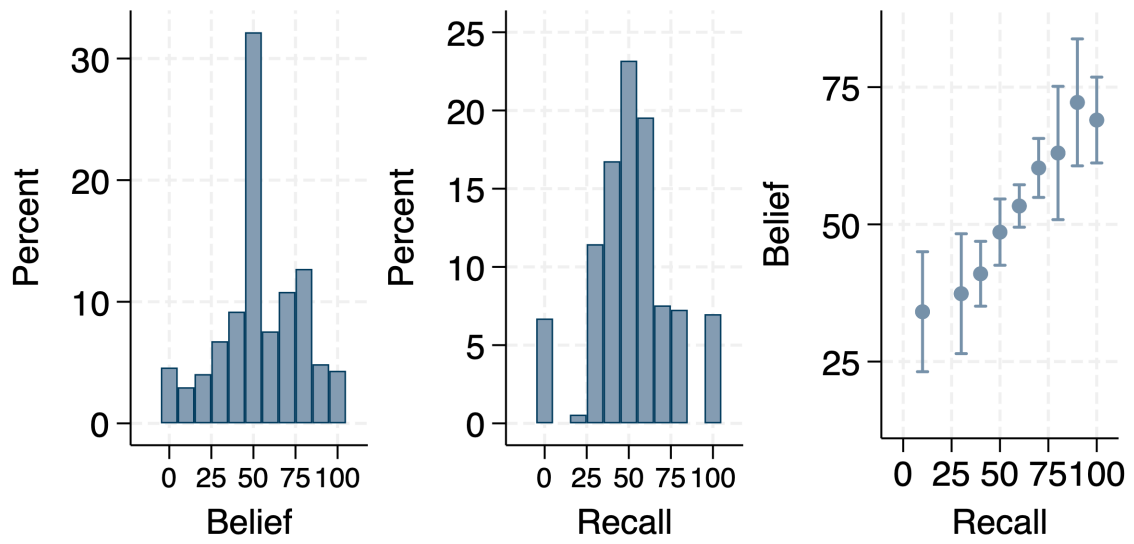


Figure 1: Beliefs and Recall

*Notes:* The left panel shows the distribution of participants' beliefs in the second survey about the percent of classes they listed in the first survey that were successes vs failures. The middle panel shows the distribution of the share of recalled classes that are successes. The right panel shows a binscatter correlating beliefs and recall.

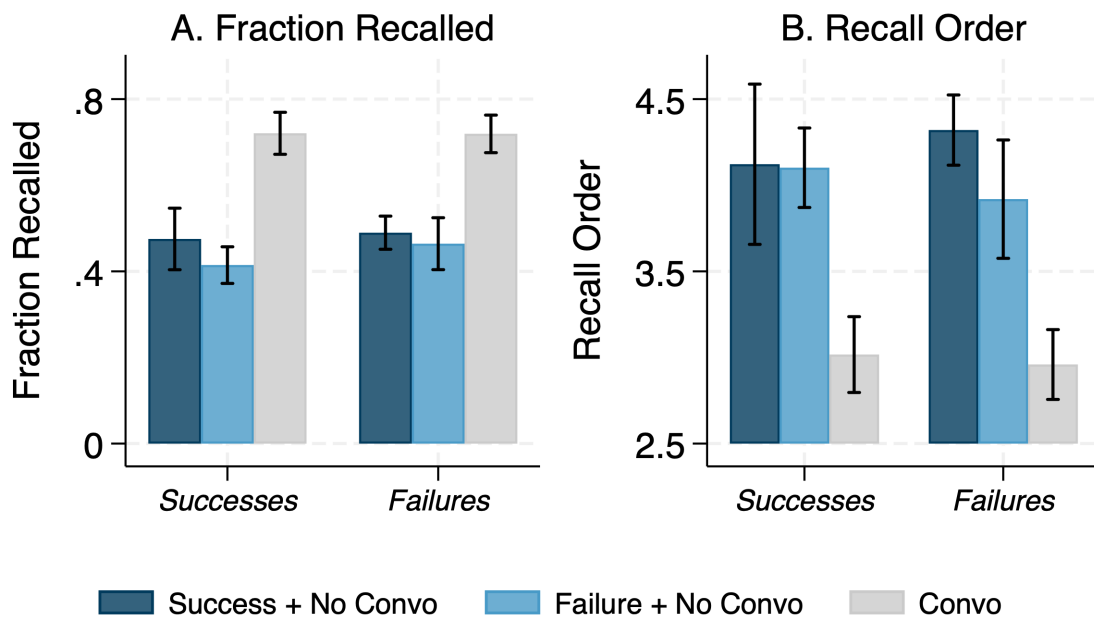


Figure 2: Effects of Rehearsal on Recall

*Notes:* Panel A shows recall rates by treatment status; Panel B shows the average position at which recalled classes appear in the free-recall task (lower means recalled earlier), among classes that were recalled. “Success + No Convo” refers to successful classes that participants did not have conversations about; “Failure + No Convo” to unsuccessful classes they did not discuss; and “Convo” to classes they did discuss (all successes in *Successes* and all failures in *Failures*). Whiskers show 95% confidence intervals. Table A.I shows the underlying numbers.

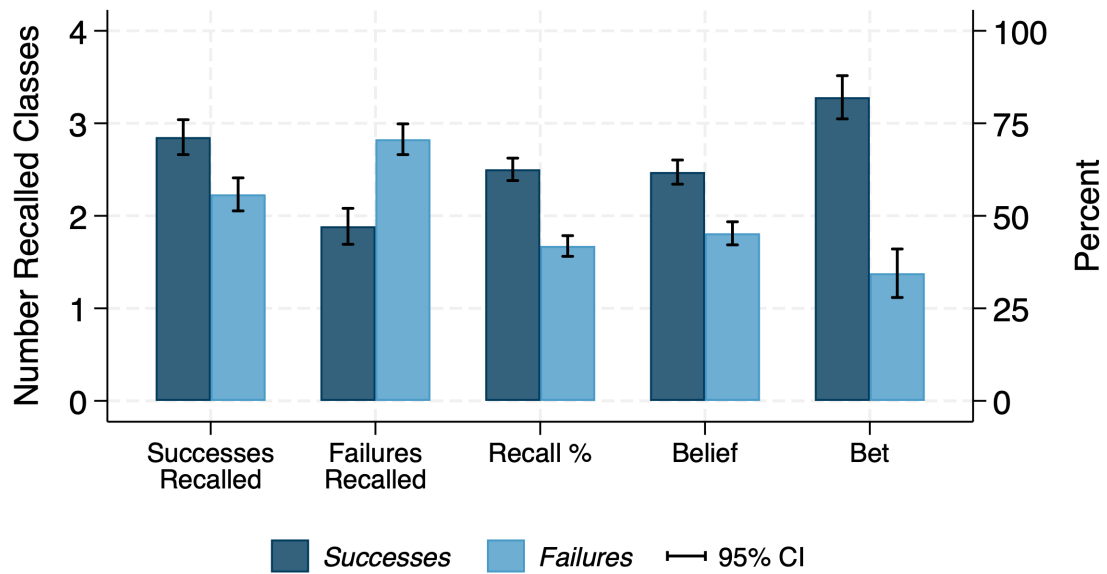


Figure 3: Effects of Rehearsal on Recall and Beliefs

*Notes:* The first two pairs of bars from the left show the number of successes and failures that participants correctly recalled in the free-recall task. The third pair of bars shows the average percent of recalled classes that were successes. The fourth pair shows the average belief about the percent of classes that were successes. The final pair shows the percent of participants who bet the randomly selected class was a success. Whiskers show 95% confidence intervals. Table A.II shows the underlying numbers. Table A.III shows analogous regressions controlling for a variety of covariates, where the results look identical.

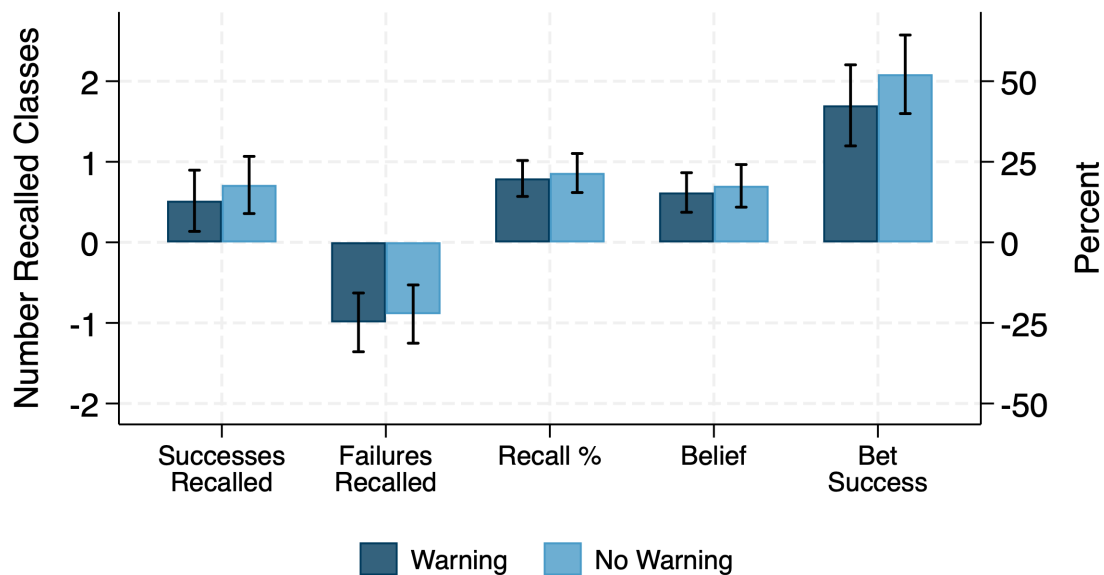


Figure 4: (Lack of) Heterogeneous Effects of Rehearsal by Warning

*Notes:* This figure shows treatment effects of being assigned to the *Successes* treatment, splitting the sample by those who were or were not warned about later incentives for accurate beliefs. The first two pairs of bars from the left show effects on the number of successes and failures that participants correctly recalled in the free-recall task. The third pair of bars shows effects on the percent of recalled classes that were successes. The fourth pair shows effects on beliefs about the percent of classes that were successes. The final pair shows effects on whether participants bet that the randomly selected class was a success. Table A.V shows the underlying regressions. Whiskers show 95% confidence intervals.

# A Supplemental Appendix: Additional Tables and Figures

Table A.I: Effects of Rehearsal on Recall

	Recalled Class			Order of Recall		
	Pooled	<i>Successes</i>	<i>Failures</i>	Pooled	<i>Successes</i>	<i>Failures</i>
	(1)	(2)	(3)	(4)	(5)	(6)
Conversation	0.720 (0.017)	0.721 (0.025)	0.719 (0.022)	2.985 (0.076)	3.017 (0.113)	2.959 (0.104)
Success + No Conversation	0.487 (0.017)	0.475 (0.037)	0.490 (0.020)	4.280 (0.096)	4.122 (0.238)	4.320 (0.104)
Failure + No Conversation	0.428 (0.018)	0.414 (0.022)	0.464 (0.031)	4.047 (0.098)	4.102 (0.118)	3.919 (0.175)
Observations	3,330	1,503	1,827	1,818	791	1,027
Individuals	370	167	203	358	161	197
<i>p</i> -value: Convo = Success + No Convo	0.00	0.00	0.00	0.00	0.00	0.00
<i>p</i> -value: Convo = Failure + No Convo	0.00	0.00	0.00	0.00	0.00	0.00
<i>p</i> -value: No Convo, Success = Failure	0.01	0.08	0.43	0.09	0.94	0.06

*Notes:* Table shows average recall rates, with standard errors (clustered at the individual level) in parentheses. Columns 1 and 4 show rates for the sample as a whole, columns 2 and 5 restrict to the *Successes* treatment, and columns 3 and 6 restrict to the *Failures* treatment. “Conversation” shows recall rates for the classes participants had conversations about, while “Success + No Conversation” and “Failure + No Conversation” show rates for the successes and failures that participants did not have conversations about. \*, \*\*, and \*\*\* show significance at the  $p < 0.10$ , 0.05, and 0.01 levels respectively.

Table A.II: Effect of Rehearsal on Recall and Beliefs

	Successes Recalled (1)	Failures Recalled (2)	Recall % (3)	Belief (4)	Bet (5)
<i>Successes</i>	0.619*** (0.133)	-0.941*** (0.131)	0.207*** (0.021)	0.165*** (0.023)	0.476*** (0.045)
Constant	2.232*** (0.091)	2.828*** (0.085)	0.418*** (0.014)	0.453*** (0.016)	0.345*** (0.033)
Observations	370	370	358	370	370

*Notes:* Table shows OLS regressions. The dependent variables are the number of successes participants correctly recalled (column 1), the number of failures they correctly recalled (column 2), the fraction of recalled classes that were successes (column 3), participants' beliefs about the share of classes that were successes (column 4), and whether participants bet that the randomly selected classes was a success (column 5). *Successes* is an indicator variable for having conversations about three successes (rather than three failures). Robust standard errors in parentheses. \*, \*\*, and \*\*\* show significance at the  $p < 0.10$ , 0.05, and 0.01 levels respectively.

Table A.III: Effect of Rehearsal on Recall and Beliefs with Controls

	Successes	Failures			
	Recalled	Recalled	Recall %	Belief	Bet
	(1)	(2)	(3)	(4)	(5)
<i>Successes</i>	0.645*** (0.131)	-0.932*** (0.133)	0.204*** (0.020)	0.158*** (0.023)	0.477*** (0.045)
Constant	1.175 (0.748)	4.144*** (0.690)	0.121 (0.115)	0.284** (0.136)	-0.044 (0.263)
Observations	370	370	358	370	370

*Notes:* Table shows OLS regressions. The dependent variables are the number of successes participants correctly recalled (column 1), the number of failures they correctly recall (column 2), the fraction of recalled classes that were successes (column 3), participants' beliefs about the share of classes that were successes (column 4), and whether participants bet that the randomly selected classes was a success (column 5). *Successes* is an indicator variable for having conversations about three successes (rather than three failures). Robust standard errors in parentheses. \*, \*\*, and \*\*\* show significance at the  $p < 0.10$ , 0.05, and 0.01 levels respectively. All regressions include linear controls for participants' self-reported memory quality (1-5 scale), their age, the share of experiences in the experiment that are successes (either 4/9 or 5/9), their belief about the share of actual classes at college/high school that they felt they did well in, whether they expressed a preference to discuss successes, whether they expressed a preference to discuss failures, and whether they are female.

Table A.IV: Effect of Knowing about Belief Incentives on Rehearsal Choices

	Chose	Chose	Chose
	Successes	Failures	Random
	(1)	(2)	(3)
Warned about Incentives	0.036 (0.052)	-0.012 (0.031)	-0.024 (0.052)
Constant	0.429*** (0.036)	0.106*** (0.022)	0.466*** (0.036)
Observations	370	370	370

*Notes:* Table shows OLS regressions. The dependent variables are indicators for whether participants chose to discuss three successes, a random three classes, or three failures (columns 1-3, respectively). Robust standard errors in parentheses. \*, \*\*, and \*\*\* show significance at the  $p < 0.10$ , 0.05, and 0.01 levels respectively.

Table A.V: Effects of Rehearsal on Recall and Beliefs, Depending on Warning about Belief Incentives

	Successes Recalled		Failures Recalled		Recall %		Belief		Bet	
	NW	W	NW	W	NW	W	NW	W	NW	W
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Successes</i>	0.712*** (0.181)	0.516*** (0.194)	-0.890*** (0.184)	-0.993*** (0.186)	0.215*** (0.031)	0.198*** (0.028)	0.175*** (0.034)	0.155*** (0.031)	0.521*** (0.062)	0.425*** (0.064)
Constant	2.086*** (0.123)	2.388*** (0.134)	2.819*** (0.112)	2.837*** (0.129)	0.404*** (0.019)	0.434*** (0.021)	0.437*** (0.023)	0.469*** (0.021)	0.276*** (0.044)	0.418*** (0.050)
Observations	189	181	189	181	186	172	189	181	189	181
<i>p</i> -value: NW = W		0.46		0.69		0.69		0.65		0.28

*Notes:* Table shows OLS regressions. The dependent variables are the number of successes participants correctly recalled (column 1-2), the number of failures they correctly recall (column 3-4), the fraction of recalled classes that were successes (column 5-6), participants' beliefs about the share of classes that were successes (column 7-8), and whether participants bet that the randomly selected classes was a success (column 9-10). *Successes* is an indicator variable for having conversations about three successes (rather than three failures). "NW" and "W" indicate whether the sample is restricted to those who were not warned or who were warned, respectively, about future incentives for accurate beliefs. Robust standard errors in parentheses. \*, \*\*, and \*\*\* show significance at the  $p < 0.10$ , 0.05, and 0.01 levels respectively.

## B Supplemental Appendix: Experimental Details

This appendix gives more details on the experimental procedures, recruitment, and incentives. The survey was implemented in Qualtrics; screenshots of the interface can be viewed in [this online document](#).

**Recruitment** Participants were recruited on Prolific, restricting the subject pool to US residents who had completed at least 50 prior studies with an approval rate of at least 95% and who were between the ages of 18 and 40 (to make remembering classes not too difficult). A total of 439 participants completed the first survey.<sup>12</sup> From these, I first drop 29 participants whose rehearsal preferences were randomly chosen to be implemented, because for such respondents I do not have random variation in rehearsal. I then drop 24 participants who gave any duplicate class names in the first survey. Finally, I drop 16 participants who did not complete the second survey, reflecting a 95.7% rate of taking the second survey (balanced across treatments). This set of 370 remaining participants constitutes the final sample. The median respondent spent 38 minutes on the first survey and 6 minutes on the second survey.

**Incentives** Participants were told that one question across the two parts of the study would be randomly selected to determine their bonus payment. In addition to the incentives for recall and belief accuracy described in the main text, the conversations participants had with GPT-4 were incentivized so as to encourage participants to be engaged and pay attention to the classes they were about. In particular, participants were told (truthfully) that their responses in each of the three conversations would be graded “on a scale from 0 to 100 according to how thoughtful they were”. If a conversation was the randomly chosen response that determined their bonus, they would earn a bonus with a percent chance equal to the grade they received for it. In practice, this grade was provided by GPT-4, though participants were not told this.

---

<sup>12</sup>The original planned sample size was 600 participants. Fewer than that number successfully completed the first survey due to a server crash that made the chatbot unresponsive for some participants. I include only respondents who did not encounter any technical issues stemming from this because those who did encounter this issue were not invited to take the second survey.

## C Supplemental Appendix: Additional Results

### C.1 Misremembering Successes vs Failures

My primary measure of recall counts only those classes that participants both name in the second survey and correctly remember as having been a success or a failure in the first survey. A natural question is whether the rehearsal manipulation led participants to *misremember* successes as failures or vice versa. Table C.I shows only limited such effects. Only 6.8% of recalled classes are misclassified, with participants being somewhat more likely to misidentify failures as successes than vice versa. Conversations do improve misclassification rates, but these effects are much smaller than the treatment effects on recall itself: compared to classes they did not have conversations about, participants are 3.4 percentage points less likely to mislabel discussed successes as failures and 4.7 percentage points less likely to mislabel discussed failures as successes. Correspondingly, Table C.II shows that treatment variation across participants has only a small effect on the number of misclassified successes and failures: the effect on the number of “recalled” false successes and false failures is only 13.1% and 7.5% as large as the effect on true successes and failures, respectively. The subjective classification of classes into successes versus failures therefore does not drive the main results, in part because participants are extremely internally consistent in how they label classes across the two surveys.

Table C.I: Misremembering is Rare

	Misclassified		
	Pooled (1)	<i>Successes</i> (2)	<i>Failures</i> (3)
Conversation	0.041 (0.008)	0.019 (0.007)	0.058 (0.013)
Success + No Conversation	0.056 (0.009)	0.042 (0.018)	0.060 (0.011)
Failure + No Conversation	0.126 (0.017)	0.135 (0.021)	0.105 (0.026)
Observations	1,951	852	1,099
Individuals	363	165	198
<i>p</i> -value: Convo = Success + No Convo	0.18	0.26	0.89
<i>p</i> -value: Convo = Failure + No Convo	0.00	0.00	0.10
<i>p</i> -value: No Convo, Success = Failure	0.00	0.00	0.12

*Notes:* Table shows average rates at which classes were recalled but incorrectly (i.e., as a success if initially a failure or *vice versa*), with standard errors (clustered at the individual level) in parentheses. Column 1 shows rates for the sample as a whole, column 2 restricts the sample to the *Successes* treatment, and column 3 restricts the sample to the *Failures* treatment. “Conversation” shows recall rates for the classes participants had conversations about, while “Success + No Conversation” and “Failure + No Conversation” show rates for the successes and failures that participants did not have conversations about.

Table C.II: Small Treatment Effects on Misremembering

	False Successes	False Failures
	(1)	(2)
<i>Successes</i>	0.082 (0.061)	-0.071** (0.033)
Constant	0.212*** (0.038)	0.143*** (0.026)
Observations	370	370

*Notes:* Table shows OLS regressions. The dependent variables are the number of failures participants incorrectly recalled as being successes (column 1) and the number of successes they incorrectly recalled as being failures (column 2). *Successes* is an indicator variable for having conversations about three successes (rather than three failures). Robust standard errors in parentheses. \*, \*\*, and \*\*\* show significance at the  $p < 0.10$ , 0.05, and 0.01 levels respectively.

## C.2 Effects on Other Class Attributes

Besides whether each class was a success or a failure, the first survey also recorded four other attributes of each class: the instructor's gender, whether the class was taken in the first two years of high school or college, whether it involved a final project, and whether it was part of the participant's major (for college classes) or an elective (for high-school classes). Because the treatment only directly manipulated whether participants rehearsed three successes or three failures, I do not have direct independent variation in how many classes with these other attributes participants discussed. The treatment does, however, induce partially random variation in that number: a student whose successes tended to be classes with female instructors, for example, will rehearse more such classes in the *Successes* treatment than in the *Failures* treatment, and a female-instructed success may have been randomly chosen for discussion instead of a male-instructed one. Exploiting this variation, and controlling for the underlying number of classes with each attribute, Table C.III shows parallel effects on recall, beliefs, and bets. Being randomly assigned to discuss more classes with a female instructor, say, makes participants more likely to recall such classes in the second survey and raises both their beliefs about and their willingness to bet on whether the randomly selected class had a female instructor.

Table C.III: Effects of Rehearsal on Recall and Beliefs about Other Attributes

	Recalled w/ Attribute (1)	Recalled w/o Attribute (2)	Recall % (3)	Belief (4)	Bet (5)
# Conversations	0.215*** (0.051)	-0.298*** (0.056)	0.060*** (0.008)	2.271** (0.954)	0.048*** (0.018)
True #	0.473*** (0.028)	-0.535*** (0.025)	0.093*** (0.003)	5.069*** (0.442)	0.070*** (0.008)
Constant	0.137** (0.060)	5.486*** (0.127)	-0.003 (0.007)	27.000*** (1.470)	0.165*** (0.025)
Observations	1,480	1,480	1,452	1,480	1,480
Individuals	370	370	363	370	370

*Notes:* Table shows OLS regressions pooled across the four attributes (instructor gender, whether the class was taken in the first two years of school, whether it involved a final project, and whether it was part of the participant’s major or an elective), so that each of the 370 participants contributes four observations. The regressor of interest, “# Conversations,” is the number of the three classes a participant discussed that had the attribute in question; “True #” is the participant’s actual number of classes (out of nine) with that attribute, included as a control. The dependent variables are the number of recalled classes that had the attribute (column 1), the number of recalled classes that did not have the attribute (column 2), the share of recalled classes that had the attribute (column 3), the participant’s belief about the percent of classes that had the attribute (column 4), and whether the participant bet that the randomly selected class had the attribute (column 5). Robust standard errors, clustered at the individual level, in parentheses. \*, \*\*, and \*\*\* show significance at the  $p < 0.10$ , 0.05, and 0.01 levels respectively.

### C.3 Chosen versus Imposed Rehearsal

In the main text, I implicitly assume that the rehearsal patterns participants would choose for themselves affect recall and beliefs in the same way as the patterns my experiment exogenously imposes by a coin flip. One might reasonably wonder, for example, whether we would see more sophistication *ex post* about the kinds of rehearsal participants actually tend to choose than about unusual patterns imposed on them. My experiment was not designed to test this directly, but some patterns in the data appear inconsistent with it. Table C.IV reports average recall, beliefs, and bets for each of the six groups defined by treatment assignment interacted with participants' elicited rehearsal preference. Column 6 reports, for each group, the average deviation between participants' *belief* about the share of successes and the share of the classes they *recalled* that were successes. For a sophisticated agent this deviation should be negative in the *Successes* treatment (revising beliefs downward to correct for successes being overrepresented in memory through rehearsal) and positive in the *Failures* treatment (revising upward to correct for failures being overrepresented).

Two facts stand out. First, these deviations are small, centered on zero ( $p = 0.331$  for the hypothesis that all six are zero), and statistically indistinguishable from one another ( $p = 0.421$ ), consistent with the overall *ex post* naivete documented in the main text. Second, I can compare participants who happened to be randomized into the treatment they would have chosen anyway—those who preferred *Successes* and were assigned to *Successes*, and those who preferred *Failures* and were assigned to *Failures*. If participants were more sophisticated about self-chosen rehearsal, the column-6 deviations should differ sharply between these two groups (more negative for the former and more positive for the latter). In fact they do not differ ( $p = 0.934$ ), and both are small and indistinguishable from zero. My finding of *ex post* naivete therefore does not appear to be driven by the experiment sometimes inducing participants to rehearse experiences they would not themselves have chosen.

Table C.IV: Average Recall and Beliefs by Rehearsal Preference and Treatment Assignment

	Successes Recalled (1)	Failures Recalled (2)	Recall % (3)	Belief (4)	Bet (5)	Belief minus Recall % (6)
<i>Successes</i> × Preferred <i>Successes</i>	2.901*** (0.147)	1.642*** (0.147)	0.671*** (0.024)	0.645*** (0.025)	0.852*** (0.040)	-0.022 (0.032)
<i>Successes</i> × Preferred <i>Failures</i>	2.429*** (0.301)	2.214*** (0.291)	0.497*** (0.048)	0.566*** (0.048)	0.857*** (0.094)	0.069 (0.057)
<i>Successes</i> × Preferred <i>Random</i>	2.875*** (0.139)	2.097*** (0.146)	0.601*** (0.020)	0.598*** (0.025)	0.778*** (0.049)	0.003 (0.027)
<i>Failures</i> × Preferred <i>Successes</i>	2.405*** (0.131)	2.702*** (0.123)	0.469*** (0.020)	0.499*** (0.025)	0.429*** (0.054)	0.023 (0.027)
<i>Failures</i> × Preferred <i>Failures</i>	1.522*** (0.247)	2.696*** (0.259)	0.312*** (0.048)	0.296*** (0.051)	0.130* (0.071)	-0.028 (0.063)
<i>Failures</i> × Preferred <i>Random</i>	2.250*** (0.139)	2.969*** (0.129)	0.399*** (0.021)	0.450*** (0.021)	0.323*** (0.048)	0.052** (0.026)
Observations	370	370	358	370	370	358
<i>p</i> -value: $S \times PS = F \times PF$	0.000	0.000	0.000	0.000	0.000	0.934
<i>p</i> -value: All means equal	0.000	0.000	0.000	0.000	0.000	0.421
<i>p</i> -value: All means zero	0.000	0.000	0.000	0.000	0.000	0.331

*Notes:* Table shows mean outcomes for each of the six mutually exclusive groups defined by treatment assignment (*Successes* vs. *Failures*) interacted with the participant’s elicited rehearsal preference (whether she preferred to discuss successes, failures, or a randomly selected set of classes). The dependent variables are the number of successes the participant correctly recalled (column 1), the number of failures correctly recalled (column 2), the share of recalled classes that were successes (column 3), the belief about the share of classes that were successes (column 4), whether the participant bet that the randomly selected class was a success (column 5), and the difference between the belief in column 4 and the recall share in column 3 (column 6). The final three rows report, for each column, *p*-values for the hypotheses that the group mean is equal across the *Successes*×Preferred *Successes* and *Failures*×Preferred *Failures* groups, that all six group means are equal, and that all six are zero. Robust standard errors, clustered at the individual level, in parentheses. \*, \*\*, and \*\*\* show significance at the  $p < 0.10$ , 0.05, and 0.01 levels respectively.

## C.4 Stated Reasons for Rehearsal Choices

After making their rehearsal choice, participants were asked to explain in their own words why they selected the option they did. I then have GPT-4o code whether each response mentions any of five considerations. These five were chosen in an *ex post* fashion, but I also allow GPT-4o to record whether a participant raises any other consideration. In total, 92.7% of participants are coded as mentioning at least one of the five, and only 14.1% mention anything else (6.8% mention both one of the five and something else), so the five considerations comprise the large majority of what participants say. Figure C.I reports the share mentioning each consideration, split by which option the participant chose; Table C.V reports the same shares for the whole sample and separately for participants who were and were not warned about future incentives for accurate beliefs.

The two considerations most often tied to choosing successes are enjoyment and ease of conversation. Participants coded as finding it “more enjoyable to discuss positive memories than negative ones” (25.7% of the sample) write things like “It’s more pleasant to think about times I performed well than times I did not,” “I feel good about how well I did in the classes and prefer to focus on the positive,” and “They bring back great memories.” Participants coded as mentioning “how easy it would be to engage thoughtfully in conversations” (37.0%)—the category meant to capture the monetary incentives in the conversations, which were graded on how thoughtful participants’ responses were—write things like “I feel that these classes are the easiest to remember to have an informed discussion about” and “I believe it will be easier for me to respond thoughtfully.”

One might worry that rehearsal choices fail to respond to the experiment’s accuracy incentives simply because those incentives are too small: participants might be aware that rehearsal distorts their future beliefs but choose to discuss successes anyway because the in-the-moment stakes dominate. I cannot conclusively rule this out, but I can check whether participants articulate any such tradeoff, and almost none do. Only 3.2% are coded as mentioning anything about how their conversations might affect what they later remember or believe, and even among these twelve responses none states the logic of rehearsal clearly.<sup>13</sup> Moreover, as Table C.V shows, the share of participants mentioning each consideration does not meaningfully vary with whether they were warned about the accuracy incentives—despite the warning coming immediately before the rehearsal choice, when it should be most salient. At the same time, participants clearly do respond to the monetary incentives within the conversations, often choosing topics they felt would let them engage more thoughtfully and so earn a higher score.

---

<sup>13</sup>The closest such participant writes “Less chance of me hyper focusing on a specific, better chance at memory recall in my opinion.”

Table C.V: Reasons for Rehearsal Choices

	Chose Successes	Chose Failures	Chose Random
<b>Full Sample</b>			
Enjoy Discuss Success	0.249	0.003	0.005
Prefer Variety	0.008	0.000	0.111
Easier to Discuss	0.265	0.059	0.046
Affect Later Memory	0.005	0.014	0.014
Indifferent	0.003	0.005	0.289
Other	0.035	0.030	0.076
<b>No Warning about Belief Incentives</b>			
Enjoy Discuss Success	0.217	0.005	0.011
Prefer Variety	0.011	0.000	0.116
Easier to Discuss	0.280	0.063	0.058
Affect Later Memory	0.005	0.011	0.016
Indifferent	0.000	0.011	0.275
Other	0.048	0.026	0.095
<b>Warning about Belief Incentives</b>			
Enjoy Discuss Success	0.282	0.000	0.000
Prefer Variety	0.006	0.000	0.105
Easier to Discuss	0.249	0.055	0.033
Affect Later Memory	0.006	0.017	0.011
Indifferent	0.006	0.000	0.304
Other	0.022	0.033	0.055

*Notes:* Table shows the share of participants who select each rehearsal option and mention each consideration when justifying their choice. The top panel shows data for the whole sample, while the lower two panels restrict to those who were not warned or who were warned about future incentives for accurate beliefs.

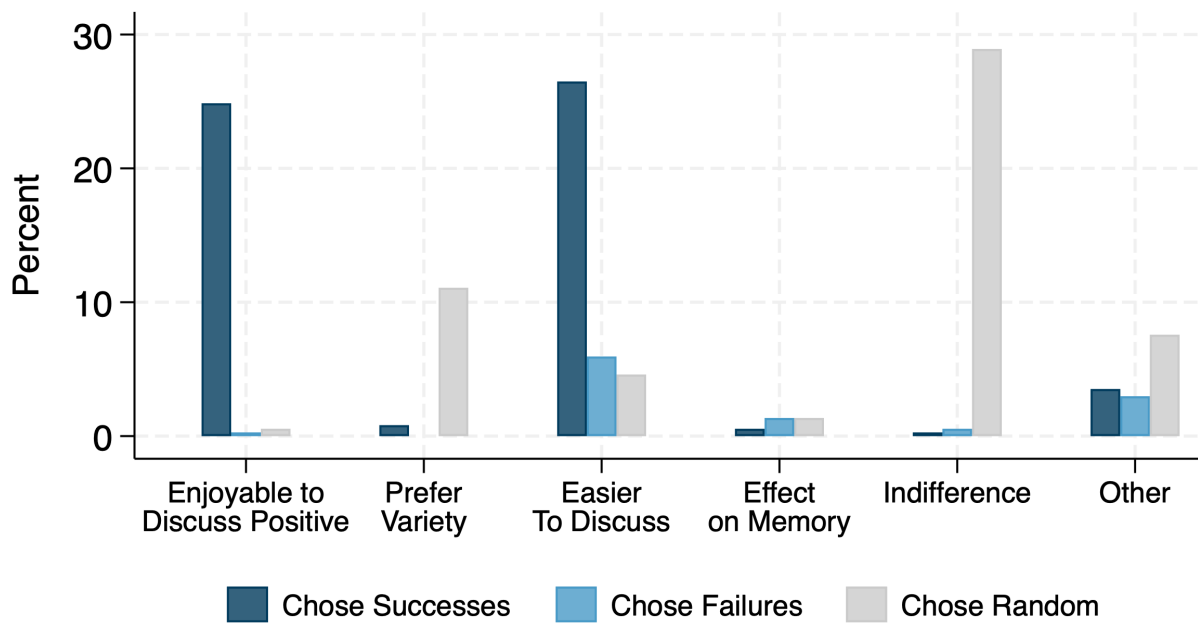


Figure C.I: Reasons behind Rehearsal Choices

*Notes:* The figure shows the share of participants who mention each of the five considerations when explaining their rehearsal choice, split by which option they selected. Table C.V reports the underlying numbers (its top panel), along with breakdowns by whether participants were warned about the accuracy incentives.

## D Supplemental Appendix: Proofs

I start by solving for the period-1 agent's beliefs about what her period-2 self's beliefs will be, which I denote by  $\tilde{E}[\bar{y}|a]$ . Recall that  $\tilde{r}_t(m, a)$  is the agent's belief in period  $t$  about her probability of recalling  $m$  given attention  $a$ . Let  $\tilde{r}_{1,2}(m, a)$  be her belief at time  $t = 1$  about what her future self at  $t = 2$  will believe this probability to be, which as in the main text I assume takes the form  $\tilde{r}_{1,2}(m, a) \propto \exp\{\lambda_1 \lambda_2 \gamma a_m\}$ . Then,

$$\begin{aligned}
\tilde{E}[\bar{y}|a] &= \sum_m y_m \cdot \tilde{r}_1(m, a) \cdot \frac{1}{M\tilde{r}_{1,2}(m, a)} \\
\Rightarrow \frac{d\tilde{E}[\bar{y}|a]}{da_l} &= \sum_m y_m \left[ \frac{d\tilde{r}_1(m, a)}{da_l} \frac{1}{M\tilde{r}_{1,2}(m, a)} + \tilde{r}_1(m, a) \frac{d\frac{1}{M\tilde{r}_{1,2}(m, a)}}{da_l} \right] \\
&= \sum_m y_m \left[ \frac{1}{M\tilde{r}_{1,2}(m, a)} \frac{d}{da_l} \frac{\exp\{\lambda_1 \gamma a_m\}}{\sum_k \exp\{\lambda_1 \gamma a_k\}} + \frac{1}{M} \tilde{r}_1(m, a) \frac{d}{da_l} \frac{\sum_k \exp\{\lambda_1 \lambda_2 \gamma a_k\}}{\exp\{\lambda_1 \lambda_2 \gamma a_m\}} \right] \\
&= y_l \left[ \frac{1}{M\tilde{r}_{1,2}(l, a)} \frac{d}{da_l} \frac{\exp\{\gamma \lambda_1 a_l\}}{\sum_k \exp\{\lambda_1 \gamma a_k\}} + \frac{1}{M} \tilde{r}_1(l, a) \frac{d}{da_l} \frac{\sum_k \exp\{\lambda_1 \lambda_2 \gamma a_k\}}{\exp\{\lambda_1 \lambda_2 \gamma a_l\}} \right] \\
&+ \sum_{m \neq l} y_m \left[ \frac{1}{M\tilde{r}_{1,2}(m, a)} \frac{d}{da_l} \frac{\exp\{\lambda_1 \gamma a_m\}}{\sum_k \exp\{\lambda_1 \gamma a_k\}} + \frac{1}{M} \tilde{r}_1(m, a) \frac{d}{da_l} \frac{\sum_k \exp\{\lambda_1 \lambda_2 \gamma a_k\}}{\exp\{\lambda_1 \lambda_2 \gamma a_m\}} \right] \\
&= y_l \left[ \frac{1}{M\tilde{r}_{1,2}(l, a)} \frac{\sum_k \exp\{\lambda_1 \gamma a_k\} \lambda_1 \gamma \exp\{\lambda_1 \gamma a_l\} - \exp\{\lambda_1 \gamma a_l\} \lambda_1 \gamma \exp\{\lambda_1 \gamma a_l\}}{(\sum_k \exp\{\gamma \lambda_1 a_k\})^2} \right. \\
&+ \left. \frac{1}{M} \tilde{r}_1(l, a) \frac{\exp\{\lambda_1 \lambda_2 \gamma a_l\} \lambda_1 \lambda_2 \gamma \exp\{\lambda_1 \lambda_2 \gamma a_l\} - \sum_k \exp\{\lambda_1 \lambda_2 \gamma a_k\} \lambda_1 \lambda_2 \gamma \exp\{\lambda_1 \lambda_2 \gamma a_l\}}{\exp\{\lambda_1 \lambda_2 \gamma a_l\}^2} \right] \\
&+ \sum_{m \neq l} y_m \left[ \frac{1}{M\tilde{r}_{1,2}(m, a)} \frac{-\exp\{\lambda_1 \gamma a_m\} \lambda_1 \gamma \exp\{\lambda_1 \gamma a_l\}}{(\sum_k \exp\{\lambda_1 \gamma a_k\})^2} \right. \\
&+ \left. \frac{1}{M} \tilde{r}_1(m, a) \frac{\exp\{\lambda_1 \lambda_2 \gamma a_m\} \lambda_1 \lambda_2 \gamma \exp\{\lambda_1 \lambda_2 \gamma a_l\}}{\exp\{\lambda_1 \lambda_2 \gamma a_m\}^2} \right]
\end{aligned}$$

Note that when  $a = 0$ ,  $\tilde{r}_1(m, a) = \tilde{r}_{1,2}(m, a) = \frac{1}{M}$  for all  $m$ . Thus,

$$\begin{aligned}
\frac{d\tilde{E}[\bar{y}|a]}{da_l} \Big|_{a=0} &= y_l \left[ \frac{M\lambda_1\gamma - \lambda_1\gamma}{M^2} + \frac{1}{M^2} \frac{\lambda_1\lambda_2\gamma - M\lambda_1\lambda_2\gamma}{1} \right] + \sum_{m \neq l} y_m \left[ -\frac{\lambda_1\gamma}{M^2} + \frac{1}{M^2} \frac{\lambda_1\lambda_2\gamma}{1} \right] \\
&= \frac{\lambda_1\gamma}{M} (1 - \lambda_2) \left[ y_l - \frac{1}{M} \sum_m y_m \right] \\
&= \lambda_1 (1 - \lambda_2) \frac{\gamma}{M} (y_l - \bar{y})
\end{aligned}$$

Then, taking a first-order approximation around  $a = 0$ ,

$$\begin{aligned}
\tilde{E}[\bar{y}|a] &\approx \bar{y} + \lambda_1(1 - \lambda_2) \frac{\gamma}{M} \sum_m (y_m - \bar{y}) a_m \\
&= \bar{y} + \gamma \lambda_1(1 - \lambda_2) E[(y_m - \bar{y}) a_m] \\
&= \bar{y} + \gamma \lambda_1(1 - \lambda_2) \left( E[y_m a_m] - E[y_m] E[a_m] \right) \\
&= \bar{y} + \gamma \lambda_1(1 - \lambda_2) \text{Cov}(y, a)
\end{aligned}$$

We can then quickly prove Proposition 1 by supposing that the period-1 self is sophisticated ( $\lambda_1 = 1$ ) since such an agent will be correct about what her future self will believe. That is, recall that

$$E[\bar{y}|a] = \sum_m y_m \cdot r(m, a) \cdot \frac{1}{M \tilde{r}_2(m, a)}$$

Note that, when  $\lambda_1 = 1$ ,  $\tilde{r}_1(m, a) = r(m, a)$  and  $\tilde{r}_{1,2}(m, a) = \tilde{r}_2(m, a)$ . Therefore, we can simply redo the above derivation but letting  $\lambda_1 = 1$  to get that

$$E[\bar{y}|a] \approx \bar{y} + \gamma(1 - \lambda_2) \text{Cov}(y, a) \tag{7}$$

which is Proposition 1. The first-order expansion linearizes two forces: nonlinearity in how attention maps to recall (recall probabilities cannot be negative), and the fact that rehearsing one experience should crowd out already-rehearsed experiences more than others.

Next, I prove Proposition 2. The first-order conditions for  $a_m$  are

$$\begin{aligned}
\nu_m &= \phi a_m^* - \left( \alpha - \beta(\tilde{E}[\bar{y}|a^*] - \bar{y}) \right) \frac{d\tilde{E}[\bar{y}|a^*]}{da_m} \\
\nu_m &= \phi a_m^* - \left( \alpha - \beta \lambda_1(1 - \lambda_2) \frac{\gamma}{M} \sum_k (y_k - \bar{y}) a_k^* \right) \frac{\gamma}{M} \lambda_1(1 - \lambda_2)(y_m - \bar{y}) \\
\nu_m &= \phi a_m^* - \alpha \frac{\gamma}{M} \lambda_1(1 - \lambda_2)(y_m - \bar{y}) + \beta \left[ \lambda_1(1 - \lambda_2) \frac{\gamma}{M} \right]^2 (y_m - \bar{y}) \sum_k (y_k - \bar{y}) a_k^* \\
\phi a_m^* &= \nu_m + (y_m - \bar{y}) \left[ \alpha \frac{\gamma}{M} \lambda_1(1 - \lambda_2) - \beta \left( \lambda_1(1 - \lambda_2) \frac{\gamma}{M} \right)^2 \sum_k (y_k - \bar{y}) a_k^* \right] \quad (8) \\
a_m^* - \frac{\nu_m}{\phi} &= \frac{1}{\phi} (y_m - \bar{y}) \frac{\gamma}{M} \lambda_1(1 - \lambda_2) \left[ \alpha - \beta \gamma \lambda_1(1 - \lambda_2) \text{Cov}(y, a^*) \right] \\
a_m^* - \frac{\nu_m}{\phi} &= \frac{1}{\phi} (y_m - \bar{y}) \frac{\gamma}{M} \lambda_1(1 - \lambda_2) \left[ \alpha - \lambda_1 \beta (E[\bar{y}|a^*] - \bar{y}) \right] \quad (9)
\end{aligned}$$

which is Proposition 2. Next, to prove Proposition 3, we can add the first-order conditions given by equation 8 across all experiences  $m$ :

$$\begin{aligned}
\phi \sum_m a_m^* (y_m - \bar{y}) &= \sum_m \nu_m (y_m - \bar{y}) + \sum_m (y_m - \bar{y})^2 \left[ \alpha \frac{\gamma}{M} \lambda_1(1 - \lambda_2) - \beta \left( \lambda_1(1 - \lambda_2) \frac{\gamma}{M} \right)^2 \sum_k (y_k - \bar{y}) a_k^* \right] \\
\sum_m a_m^* (y_m - \bar{y}) &= \frac{\sum_m \nu_m (y_m - \bar{y}) + \sum_m (y_m - \bar{y})^2 \left[ \alpha \frac{\gamma}{M} \lambda_1(1 - \lambda_2) \right]}{\phi + \beta \sum_m (y_m - \bar{y})^2 \left( \lambda_1(1 - \lambda_2) \frac{\gamma}{M} \right)^2} \\
\text{Cov}(a^*, y) &= \frac{\text{Cov}(y, \nu) + \text{Var}(y) \left[ \alpha \frac{\gamma}{M} \lambda_1(1 - \lambda_2) \right]}{\phi + \beta M \text{Var}(y) \left( \lambda_1(1 - \lambda_2) \frac{\gamma}{M} \right)^2}
\end{aligned}$$

Plugging this expression into equation 7 then yields Proposition 3.

To derive Proposition 4, note that the utility effect of increasing beliefs above the agent's

actual beliefs (given her endogenous rehearsal choices  $a^*$ ) is given by:

$$\begin{aligned}
\frac{du_2}{dE_a[\bar{y}]}\Big|_{a^*} &= \alpha - \beta(E_{a^*}[\bar{y}] - \bar{y}) \\
&= \alpha - \beta\gamma(1 - \lambda_2)\text{Cov}(y, a^*) \\
&= \alpha - \frac{\beta\gamma(1 - \lambda_2)}{\phi + \beta M\text{Var}(y)\left(\lambda_1(1 - \lambda_2)\frac{\gamma}{M}\right)^2} \left[ \text{Cov}(y, \nu) + \text{Var}(y) \left[ \alpha \frac{\gamma}{M} \lambda_1(1 - \lambda_2) \right] \right] \\
&= \frac{1}{\phi + \beta M\text{Var}(y)\left(\lambda_1(1 - \lambda_2)\frac{\gamma}{M}\right)^2} \left[ -\beta\gamma(1 - \lambda_2)\text{Cov}(y, \nu) + \right. \\
&\quad \left. \alpha \left( \phi + \beta\text{Var}(y)\left(\lambda_1(1 - \lambda_2)\right)^2 \frac{\gamma^2}{M} - \beta(1 - \lambda_2)^2\text{Var}(y)\frac{\gamma^2}{M}\lambda_1 \right) \right] \\
&= \frac{1}{\phi + \beta M\text{Var}(y)\left(\lambda_1(1 - \lambda_2)\frac{\gamma}{M}\right)^2} \left[ -\beta\gamma(1 - \lambda_2)\text{Cov}(y, \nu) + \right. \\
&\quad \left. \alpha \left( \phi - \beta\text{Var}(y)\lambda_1(1 - \lambda_2)^2 \frac{\gamma^2}{M}(1 - \lambda_1) \right) \right]
\end{aligned}$$

Note that Proposition 4(i) follows directly from the above expression, since  $\text{Cov}(y, \nu)$  only appears in the negative first term. Note that this comparative static is implicitly holding the distribution of experiences  $y$  constant (since it appears  $\text{Var}(y)$  appears elsewhere in the expression). So this result should be thought of (and is phrased as) being about how changes in the agent's rehearsal-based preferences (i.e.,  $\nu$ ) affects beliefs and the welfare impact of correcting beliefs.

Next, suppose  $\text{Cov}(y, \nu) = 0$ . Then,

$$\begin{aligned}
&-\frac{du_2}{dE_a[\bar{y}]}\Big|_{a^*} < 0 \\
\iff &-\frac{1}{\phi + \beta M\text{Var}(y)\left(\lambda_1(1 - \lambda_2)\frac{\gamma}{M}\right)^2} \left[ \alpha \left( \phi - \beta\text{Var}(y)\lambda_1(1 - \lambda_2)^2 \frac{\gamma^2}{M}(1 - \lambda_1) \right) \right] < 0 \\
\iff &\phi - \beta\text{Var}(y)\lambda_1(1 - \lambda_2)^2 \frac{\gamma^2}{M}(1 - \lambda_1) > 0 \\
\iff &\lambda_1(1 - \lambda_1) < \frac{\phi}{\beta\text{Var}(y)(1 - \lambda_2)^2 \frac{\gamma^2}{M}}
\end{aligned}$$

A sufficient condition for the above to hold is that  $\lambda_1(1 - \lambda_1) < \frac{\phi}{\beta\text{Var}(y)\frac{\gamma^2}{M}}$ , which is Proposition 4(ii).

Next,

$$\begin{aligned}
& -\frac{du_2}{dE_a[\bar{y}]}\Big|_{\alpha^*} > 0 \\
& \iff \alpha \left( \phi - \beta \text{Var}(y) \lambda_1 (1 - \lambda_2)^2 \frac{\gamma^2}{M} (1 - \lambda_1) \right) < \beta \gamma (1 - \lambda_2) \text{Cov}(y, \nu) \\
& \iff \frac{\alpha \phi - \alpha \beta \text{Var}(y) \lambda_1 (1 - \lambda_2)^2 \frac{\gamma^2}{M} (1 - \lambda_1)}{\beta \gamma (1 - \lambda_2)} < \text{Cov}(y, \nu)
\end{aligned}$$

Notice that the left-hand side is bounded above by  $\frac{\alpha \phi}{\beta \gamma (1 - \lambda_2)}$ , hence the sufficient condition described in Proposition 4(iii).