**The effects of false feedback on state memory distrust toward commission and omission, and recognition memory errors**

Yikang Zhang1, 2, Henry Otgaar2, 3, Robert A. Nash4, Chunlin Li2, 3

1 Max Planck Institute for the Study of Crime, Security and Law, Germany

2 Faculty of Psychology and Neuroscience, Maastricht University, the Netherlands

3 Faculty of Law and Criminology, KU Leuven, Belgium

4 School of Psychology, Aston University, the United Kingdom

**Author Information**

Yikang Zhang ([kang.y.zhang@outlook.com](mailto:kang.y.zhang@outlook.com) )  0000-0001-5173-562X

Henry Otgaar ([henry.otgaar@maastrichtuniversity.nl](mailto:henry.otgaar@maastrichtuniversity.nl))  0000-0002-2782-2181

Robert A. Nash ([R.Nash1@aston.ac.uk](mailto:R.Nash1@aston.ac.uk) )  [0000-0002-2284-2001](https://orcid.org/0000-0002-2284-2001)

Chunlin Li ([ElsaLee214@hotmail.com](mailto:chunlin.li@kuleuven.be))  0000-0002-6814-8103

Correspondence concerning this article should be addressed to Yikang Zhang, Günterstalstr. 73, Freiburg i. Br. · Germany. Postcode: 79100. E-mail: [kang.y.zhang@outlook.com](mailto:kang.y.zhang@outlook.com).

## Contributor Roles Taxonomy

In the table below, we employed CRediT ([Contributor Roles Taxonomy](https://credit.niso.org/)) to identify the contributions and roles played by the contributors in the current research effort.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Yikang Zhang | Henry Otgaar | Robert A. Nash | Chunlin Li |
| Conceptualization | X | X | X | X |
| Data curation | X |  |  |  |
| Formal analysis | X |  |  |  |
| Funding acquisition |  | X |  |  |
| Investigation | X | X | X | X |
| Methodology | X | X | X | X |
| Project administration | X |  |  |  |
| Resources |  |  |  |  |
| Software |  |  |  |  |
| Supervision |  | X | X |  |
| Validation |  |  |  |  |
| Visualization | X |  |  |  |
| Writing-original draft | X | X | X | X |
| Writing-review and editing | X | X | X | X |

# Abstract

Memory distrust refers to the subjective appraisal of one’s memory functioning and it has two aspects: distrust over making omission errors (e.g., forgetting) and distrust over making commission errors (e.g., falsely remembering). Although these two aspects of memory distrust have been studied, how they are associated with memory validation (e.g., the formation of autobiographical belief) and reporting remains unclear. In the present study, we experimentally examined the effect of metacognitive appraisals on the memory validation process as well as the commission and omission errors in memory reporting. Specifically, participants (*N* = 622, *M*age = 38.67, *SD*age = 12.23) completed a memory task in which they either received inaccurate feedback regarding making commission errors or making omission errors, or they received no feedback. Then, they completed a second recognition task. Compared to the control group with no feedback, people who received feedback that indicated a tendency to make commission errors showed a shift toward a more conservative response criterion; in contrast, participants who received feedback indicating a tendency to make omission errors showed a shift toward a more liberal response criterion. However, manipulation checks failed to show that our manipulations had the expected effect on state memory distrust; we therefore did not find sufficient evidence that the effect of feedback was through its effect on state memory distrust toward commission or omission. Possible explanations for the results and future directions are discussed.

*Keywords*: Memory distrust, misinformation, response criterion, commission errors, omission errors

# Introduction

In most instances of remembering, the person trusts that the recollected event truly happened in the past. Such event representations, which encompass both vivid recollection and a firm belief in the event’s occurrence, are referred to as believed memories (Scoboria et al., 2014). However, there are other types of event representations in addition to believed memories. For some events, people can hold strong beliefs about their occurrence without any recollections about them, such as the celebration of your first birthday (believed-but-not-remembered events). More interestingly, in some instances, people can recollect certain events vividly but do not believe the events have happened in the past. The latter phenomenon is known as nonbelieved memories (NBMs, Mazzoni et al., 2010; Otgaar et al., 2014). Cases where the recollection and belief of a memory diverge reveal that autobiographical recollection and belief of an event are two related but distinct constructs (Clark et al., 2012; Mazzoni et al., 2010; Otgaar et al., 2014, Scoboria et al., 2014).

The dissociation between recollection and belief can also be observed in other ways. Scoboria et al. (2014) reported that memory characteristics that predict recollection ratings well (e.g., perceptual, re-experiencing, emotion intensity, event specificity) did not predict belief ratings and vice versa (e.g., plausibility judgment). Furthermore, studies on nonbelieved memory showed that autobiographical beliefs can be altered relatively easily —more so than can recollections—in response to social information contradicting one’s memories (Li et al., 2020; Otgaar et al., 2013, 2017; Scoboria et al., 2018; Wang et al., 2017). In particular, the credibility of social information influences the formation of false autobiographical beliefs but not the recollective features (Scoboria et al., 2014). On the other hand, in the course of forming false memories, false beliefs can be created more easily relative to false recollections (Mazzoni et al., 2001; Pezdek et al., 1997). People who start to falsely believe a suggested event may then utilize general scripts for similar events, or memory details of other episodes, to develop their recollections (Pezdek et al., 1997).

Blank (2017) posited that these recollection-belief divergences, such as nonbelieved memories or believed-not-remembered events, arise from normal, healthy metacognitive monitoring and control processes that balance recollections and reality constraints. Specifically, Blank (2009) proposed a theoretical framework of remembering that distinguishes memory, belief, and the communication of memory (e.g., memory reporting) and elaborates on the role of social influence in remembering. First, memory traces are activated through either internal processes (e.g., intentional searches) or external cues (e.g., verbal prompts). The second stage involves a validation process in which the validity of the memory is inferred from both the retrieved internal information and from external information (e.g., physical evidence such as photos or social feedback) to form a (memory) belief. Finally, the rememberer decides whether or not to communicate and how to communicate the belief to others. In this view, an autobiographical belief is the summative evaluation of the truth status of the remembered events at the time of retrieval (see also Otgaar et al., 2014; Scoboria et al., 2014, Scoboria & Henkel, 2020). Notably, since the information available could change in each instance of remembering, the truth status assigned to the events could therefore also be revisited and substantially altered, possibly resulting in the change of memory reports such as the retraction of allegations of sexual abuse (Blank, 2017; Ost, 2017).

Expanding this theoretical framework, recent research brings to the discussion the roles of beliefs and appraisals in the process of memory validation (e.g., Nash et al., 2023; Zhang et al., 2022a, 2022b). More specifically, researchers have argued that memory distrust, defined as the skeptical perception or appraisal of one’s own memory functioning, contributes to the formation of false memory and nonbelieved memory (Nash et al., 2023; Zhang et al., 2022b). The term “memory distrust” was first coined by Gudjonsson and MacKeith (1982) to describe a phenomenon in which interviewees in a forensic context develop profound distrust towards their memories due to inappropriate investigation practices. Subsequently, researchers expanded the concept of memory distrust to denote a more general and stable appraisal of one’s memory functioning (Nash et al., 2023; van Bergen et al., 2010). As of now, there exist two validated scales measuring trait memory distrust: the adapted Squire Subjective Memory Questionnaire (SSMQ, Squire et al., 1979; van Bergen et al., 2010) focuses on people’s beliefs about their susceptibility to memory omission errors (i.e., forgetting) whereas the Memory Distrust Scale (MDS, Nash et al., 2022), which measures beliefs about susceptibility to commission errors (i.e., falsely remembering events which never happened).

Zhang et al. (2022a, 2022b) argued that individuals vary in how much they trust their recollections when comparing them to external information. Specifically, people with high memory distrust (vs. low memory distrust) would give less credence to their recollections and would in turn lower their memory beliefs. Building on this premise, Zhang et al. (2022a, 2022b) examined the relationship between memory distrust (as measured by the SSMQ) and nonbelieved memories, finding that people with high memory distrust were more likely to report nonbelieved memories than those with low memory distrust. Similarly, using an indirect recall method, in which participants were given cues to recall past events and then rated their recollections and beliefs about those events, Nash et al. (2023) found that people who reported nonbelieved memories, compared with people who did not, scored higher on average on memory distrust (as measured by the MDS). Furthermore, in one study, individuals with high or low memory distrust differed in their chosen strategies for verifying personal memories (Zhang et al., 2023a). In Zhang et al.’s (2023a) research, people with high (vs. low) memory distrust, when confronted with the possibility of erroneous memory, were more likely to use low-cost but less reliable verification strategies (e.g., seeking information from other people), thus exhibiting a greater ‘cheap-and-easy strategy’ bias (Nash et al., 2017; Wade et al., 2014), and they seemed less likely to rely on their recollections for making belief judgments, even in the absence of external information (Zhang et al. 2024).

A more nuanced picture of the relationship between memory distrust and memory appraisal was revealed through analyses using the Signal Detection Theory (SDT, Green & Swets, 1966). SDT is a framework designed to analyze people’s ability to differentiate between signal and noise (i.e., sensitivity or discriminability) and their thresholds for deciding that a given stimulus is a signal (i.e., response criterion), in decision-making processes such as recognition memory. For example, some people may have a very conservative response criterion and therefore only judge a stimulus as signal when the evidence is strong, resulting in them making few false alarms but also having relatively poor correct recognition of the signal (i.e., hits). Others, in contrast, might respond more liberally, requiring more moderate evidence for judging a stimulus as signal, and therefore making more hits and more false alarms. Zhang et al. (2023c) reported a positive association between participants’ scores on the MDS (but not SSMQ) and the response criterion index *β*, with higher values indicating a more conservative response bias (i.e., a bias toward judging stimuli as new in a recognition task). However, in another study (Zhang et al., 2024), higher memory distrust as indexed by the MDS was negatively associated with *β,* whereas higher memory distrust as indexed by the SSMQ was positively associated with *β*. Put differently, in one study, people who believed that they were susceptible to certain memory errors exhibited response biases that aligned with those subjective beliefs (Zhang et al., 2024); whereas in another study (Zhang et al., 2023c), people who believe themselves being susceptible to certain memory errors exhibited response biases that seemed to compensate for that susceptibility.

Two major differences in the design may have caused these inconsistent results, as argued by Zhang et al. (2024). First, in Study 2 of Zhang et al. (2023c), participants were given (correct or incorrect) feedback allegedly from another participant after each judgment. Second, there was a cash incentive for high performers in the memory task of that study. Both features were absent in Zhang et al. (2024). Peer feedback and performance incentives might lead participants to adjust their responses. They may do so by applying subjective memory appraisals to their judgments. For example, if a participant thinks that they often make commission errors, they should then have a higher threshold to make an “old” judgment in this task, and in the absence of incentive, this participant may rely heavily on recollection to make this judgment (Blank, 2017), as reflected in the positive associations of memory distrust with recollection and with *β* in Zhang et al. (2024). But in the presence of an incentive, a participant who thinks they often make commission errors might aim to improve their accuracy by adopting a more conservative response criterion. Taken together, accumulating evidence leads us to propose that under specific conditions, people apply their beliefs about their subjective memory functioning to calibrate their evaluations of their recollections.

So far, the limited work on the role of memory distrust in the memory validation process considered only trait memory distrust (e.g., Zhang et al., 2024, 2023c), which complicates the picture as it is associated with the objective functioning of memory. That is, trait memory distrust may influence memory performance through various mechanisms. First, given that trait memory distrust is associated with objective memory functioning, we can expect people who are high (vs. low) on memory distrust to perform worse in memory tasks (i.e., lower sensitivity, Zhang et al., 2024; but see Kuczek et al., 2018 for a different result). Second, in situations where individuals reflect upon their subjective memory functioning and calibrate their memory validation based on it, we would expect memory distrust to have a different impact on memory performance, e.g., response bias. Further, the moderation effect of memory distrust on recollection-belief correspondence is complicated by the fact that memory distrust is associated with both belief and recollection ratings (Zhang et al., 2024).

To single out the potential calibration effects of memory distrust on memory validation, the current study aims to experimentally manipulate participants’ concern over making omission errors or commission errors, and examine the effect of this manipulation on subsequent memory reporting, including recognition, recollection, and belief judgments. We hypothesized that with an incentive to be as accurate as possible, participants who are told that they often make commission errors will adjust their response criterion, and show a shift toward a more conservative response bias. On the other hand, participants who are told that they often make omission errors will show a shift towards a more liberal response bias. The idea that feedback can induce criterion shifts is by no means novel. Research has shown that biased positive feedback on incorrect old classification of new items (i.e., false alarm), or incorrect new classification of old items (i.e., miss), caused participants to shift toward a more lax or conservative response criterion, respectively, without them being aware of the biased nature of the feedback (Han & Dobbins, 2008; 2009). Despite the similarity, the focus of the current study is not to examine the incremental criterion shift based on trial-by-trial positive feedback, but rather, a more general appraisal of one’s memory functioning that is influenced by negative feedback.

# Method

All study materials are available at the Open Science Framework (OSF) [https://osf.io/8qbkn/?view\_only=a663f2e3619545edafe86b0aee885603]. The present study received ethical approval from the Ethics Review Committee Psychology and Neuroscience (ERCPN) of Maastricht University (reference: ERCPN-OZL\_246\_167\_12\_2021\_S5). The accepted Stage 1 manuscript can be accessed at <https://osf.io/x69qt>.

## Design

The study used a between-participant design with three conditions (Feedback: omission vs. commission vs. no feedback).

## Participants

### Sample Size Justification

**Expected Effect Size of Criterion Shift.** We expected the effect of the memory distrust manipulation on criterion shift to be smaller than the effect of explicit instructions to respond liberally or conservatively (ca conservative = 0.34, ca liberal = -0.50, ca diff = 0.84, Azimian-Faridani & Wilding, 2006) but larger than the effect of implicit biased feedback (Experiment 2: ca conservative = 0.39, ca liberal = 0.02, ca diff = 0.37; Experiment 3: ca conservative = 0.19, ca liberal = 0.05, ca diff = 0.14, Han & Dobbins, 2008). We therefore set a conservative expected effect size as a difference of c = 0.15 between the omission condition and the control condition and between the control and commission condition (in between Experiments 2 and 3 in Han & Dobbins, 2008).

**Smallest Effect Size of Interest of Criterion Shift.** We consider the following data pattern to be the smallest effect size of interest (SESOI) in the current experimental setup: the top 25% of participants who are most receptive to the memory distrust (Omission) manipulation will make one more hit response and one more false alarm response compared to their control counterparts in the current experiment. Similarly, the top 25% of participants who are most receptive to the memory distrust (Commission) manipulation will make one less hit response and one less false alarm response compared to their control counterparts[[1]](#footnote-1). The justification behind one hit and one false alarm comes from the idea that the remembrance of just one (false) detail can have practical value (e.g., misremembrance of the face of a culprit). Using the data from Study 2 of Zhang et al. (2023c) as the control condition and the expected differences between conditions, we created a synthetic dataset and calculated the SDT indices (See Table A1). In the synthetic dataset (See Table A1), the average difference of response criterion c is 0.06 between the Feedback-Omission and control or between control and the feedback commission with a standard deviation of 0.30 (i.e., Cohen’s d = 0.06/0.30 = 0.2). With the same standard deviation assumed, our expected effect size would translate to a difference of Cohen’s d = 0.50.

**Expected Effect Size of State Memory Distrust.** Given that in Dudek and Polczyk (2023), the difference of memory distrust between the experimental group and control group was close to that of Cohen’s *d* = 1.0, we consider the expected effect size of our manipulation on state memory distrust to be Cohen’s *d* = 1.0.

**Smallest Effect Size of Interest of State Memory Distrust**. Assuming that memory distrust is the underlying mechanism of response criterion change, and that the correlation between state memory distrust and response criterion c is at least r = 0.25[[2]](#footnote-2), this requires that the strength of the manipulation to be no smaller than Cohen’s d = 0.2/ 0.25 = 0.8 for comparisons between memory distrust conditions and the control condition (calibration, Dienes, 2021). That is, state memory distrust toward omission should be 0.8 SD higher in the Feedback-Omission condition than in the control condition. Similarly, the state memory distrust toward commission should be 0.8 SD higher in the Feedback-Commission condition than in control condition.

**Power Analyses.** We performed simulation-based power analysis for minimal effect testing and equivalence testing (see Riesthuis, 2024 for the tutorial) for the pairwise comparisons of response criterion and state memory distrust between the (a) control and omission conditions and (b) control and commission conditions. Specifically, the minimal effect testing calculated the percentage of simulations wherein the lower bound of an 80% Confidence Interval (CI) of the effect was greater than our SESOI (i.e., cdiff = 0.06, raw score difference of state memory distrust = 1.6 scale-points). The equivalence testing calculated the percentage of simulations wherein the 80% CI fell within the two equivalence bounds (i.e., c: [-0.06, 0.06], state memory distrust: [-1.6, 1.6]). This analysis on criterion c showed that when the true effect size is cdiff = 0.15 (Cohen’s *d* = 0.5), a group size of 100 participants would have 80% power to detect the minimal effect. Looking instead at memory distrust, the analysis showed that when the true effect size on memory distrust is a raw score difference of 2 points (Cohen’s *d* = 1.0), a group size of 210 participants would have 80% power to detect the minimal effect of 1.6. We therefore decide to set the minimum number of participants as 210 per group based on the requirement of state memory distrust (<https://osf.io/5x43j>).

For the analyses on criterion c, a group of 210 participants would afford 96% power to detect the minimal effect of cdiff = 0.06. With this group size, there is a 54% probability that the 80% CI would fall between the equivalence bounds [-0.06, 0.06] when the true effect is 0. For the analyses on state memory distrust, with the same group size and when the true effect is 0, the 80% CI will fall between the equivalence bounds [-1.6, 1.6] almost 100% of the time (<https://osf.io/hxb6c>). For both criterion c and state memory distrust, when the true effects are zero, with a sample of 210, in zero out of 1000 simulations, the 80% CIs are above the minimal effects of interest (i.e., a false positive).

Sensitivity analysis with G\*Power 3.1 (Faul et al, 2009) showed that a sample of 630 (210\*3) would allow us to detect a slope of 0.015 criterion (c) units/Likert unit of state memory distrust (commission or omission) in a linear regression examining the association between state memory distrust and response criterion c (α = .05 and 1-β = .80; See appendices- Sensitivity Analysis protocol).

We recruited 787 participants using Connect (<https://connect.cloudresearch.com/>), among which, 653 participants completed both sessions. After excluding three participants who had participated in similar studies and 28 participants who questioned the validity of the performance feedback, per our preregistration, the final sample consisted of 622 participants (*n* women = 306, *n* men = 299, *n* nonbinary = 2, *n* other = 1, *n* missing = 14; *M* age = 38.67, *SD* age = 12.23). The only inclusion criteria were that participants should be aged 16 or above, have normal (corrected) vision, and be fluent English speakers. Participants could only sign up for and complete the experiment using laptops or PCs (i.e., no tablet or mobile answering) and they received $1.20 for the first session (approximately 10 minutes) and $1.80 for the second session (approximately 18 minutes) one day later. Participants whose memory task performance was among the top 10% also received an additional bonus of $3 (see the Procedure below).

## Materials and Procedure

### Stimuli for the Memory Task

To increase the generalizability of our findings across a wider range of stimuli but also to minimize the potentially unnecessary emotional impact on participants, a total of 80 valenced (valence ≥ 3 and arousal ≤ 5 on 7-point scales) colour images were selected from the open affective standardized image set (OASIS, Kurdi et al., 2017), an open-access stimulus set depicting a broad spectrum of natural or social situations (e.g., pets, people, buildings, or car accidents). We then randomly divided these 80 scenes into two 40-scene blocks that were used either as old (i.e., appearing in encoding) or new stimuli (i.e., only appearing in tests as fillers). Each block was then subdivided into 2 sub-blocks of 20 scenes that were tested either during Test 1 or Test 2. Using Kurdi et al.’s (2017) norming data we performed two-way between-subject ANOVAs to ensure that there was no statistically significant difference in the mean level of arousal and valence ratings between different blocks and old vs. new scenes (see Table A2 in the Appendix).

Forty scenes were presented during encoding (hereafter referred to as old scenes) and another 40 scenes were used as fillers for the first and second recognition tests. Old scenes and new scenes were counterbalanced between participants. That is, each scene was an old scene for half of the participants while being a new scene for the other half of the participants. Designated old scenes and new scenes were further randomly divided into two blocks (*n* = 20). For the first recognition task, one block of old scenes and one block of new scenes as fillers were randomly selected and presented to participants, leading to a recognition test containing 40 scenes. In the second recognition task, participants were presented with the remaining block of old scenes and the remaining block of 20 new scenes. The order of the blocks is thus also counterbalanced.

### Manipulation Check

Two statements measuring state memory distrust were completed using a 10-point Likert scale (from 1 = Extremely disagree to 10 = Extremely agree). The first statement is “At this moment, if I am asked to retrieve something from memory, I think I likely would remember things that did not happen” (state memory distrust toward commission). The second statement is “At this moment, if I am asked to retrieve something from memory, I think I likely would forget things that happened” (state memory distrust toward omission). Half of the participants answered the manipulation checks before the second test and half of the participants answered the manipulation checks after the second test. This random assignment was intended to examine whether the expected effect is influenced by the timing of the manipulation check, which could in principle make the feedback on commission /omission errors more salient. For example, it could be that there is a stronger effect of feedback in the manipulation-check-first conditions than in the memory-test-first conditions.

### Trait Memory Distrust

As stated earlier, previous research showed that aspects of trait memory distrust are differently associated with response bias and memory performance. We therefore also include measures of trait memory distrust in the current study, in an attempt to further examine the associations. The SSMQ (Squire et al., 1979) adapted by van Bergen et al. (2010) has 18 items (e.g., “my ability to pay attention to what goes on around me is” from -4 = *Disastrous* to 4 = *Excellent)* measuring distrust toward making omission errors. The Memory Distrust Scale (MDS; Nash et al., 2023) consists of 20 items (e.g., “I am sometimes uncertain whether an event that I recall really happened to me, or whether I saw it on TV or in a movie” from 1 = *Strongly disagree* to 7 = *Strongly agree*) measuring distrust toward commission errors. To ease the comparison of results, after establishing internal consistency of the SSMQ and the MDS in the current sample (Cronbach’s *α* =.95 and .97 respectively), we reverse-coded the SSMQ and then calculated the mean of all items for the two scales, so that a higher mean score in both scales reflects a higher level of memory distrust. Following this transformation, the two scales were positively correlated in the current sample, *r* = .46, *p* <.001, 95% CI [.40, .51].

### Procedure

**Session 1.** After reading the information letter and giving informed consent, participants first answered demographic questions about their age, gender, and education level, followed by the SSMQ and the MDS in counterbalanced order. Embedded in the two scales, we included three attention checks asking participants to choose a specific answer for that item (e.g., for this item, please choose “strongly agree”). Participants then viewed 40 scene images, one at a time and in randomized order. Each scene was presented for 3 seconds with an inter-stimulus interval of 1 second. At the end of Session 1, participants were reminded to sign up for and complete Session 2 at the same time the next day. In the current study, we expected to complete Session 1 data collection within 3 hours. If, however, there were not enough participants signing up for the study within a 3-hour window, we closed the signup for Session 1 after 3 hours and ran another cycle of data collection until the planned sample size was met.

**Session 2.** Twenty-four hours later, Session 2 was made available online. Participants then had a 6-hour window to sign up for and complete Session 2, after which the session was closed. They were first reminded of the tasks involved in the study and informed that they would receive feedback on their incorrect judgments after each of their decisions. The session began with a first recognition task, presenting 20 previously seen (old) scenes alongside 20 new scenes, displayed individually in a random sequence without a time constraint. In this task, participants first indicated whether a scene was old or new (“Did you see this exact scene in yesterday’s session?” Options: *yes* or *no*). Regardless of their recognition judgment, participants then assessed their levels of recollection (“Do you actually remember seeing this exact scene in yesterday’s session?”) and belief in occurrence (“Regardless of whether or not you remember this scene, do you believe that this scene appeared in yesterday’s session?") on an 8-point Likert-like scale (Memory: 1 = *No memory at all*, 8 = *Clear and complete memory;* Belief: 1 = *Definitely did not appear*, 8 = *Definitely did appear*). These questions were adapted from Scoboria et al. (2004) and Li et al., (2020; 2023).

After making recognition, recollection, and belief judgments for each scene, participants were shown feedback on supposedly incorrect recognitions. In reality, the feedback contained both true and false feedback, depending on their assigned condition, to increase the credibility of the feedback manipulation. For participants who were in the feedback-commission condition, they received false feedback on some of the old items that they had correctly identified as old, and true feedback on new items they had incorrectly identified as old. For each correctly recognized old item, there was a 20% probability that participants received false feedback that this was actually a new scene. For each incorrectly recognized new item, participants received true feedback that this is a new scene. No feedback on non-identified old items (i.e., misses) was given. For participants who were in the feedback-omission condition, they received false feedback on some of the new items that they had correctly identified as new, and true feedback on old items that they had incorrectly identified as new. For each correctly recognized new item, there was a 20% probability that participants received false feedback that this was actually an old scene. For each incorrectly recognized old item, participants received correct feedback that this was an old scene. No feedback on incorrectly identified new items (i.e., false alarms) was given. For participants in the control group, no feedback was given.

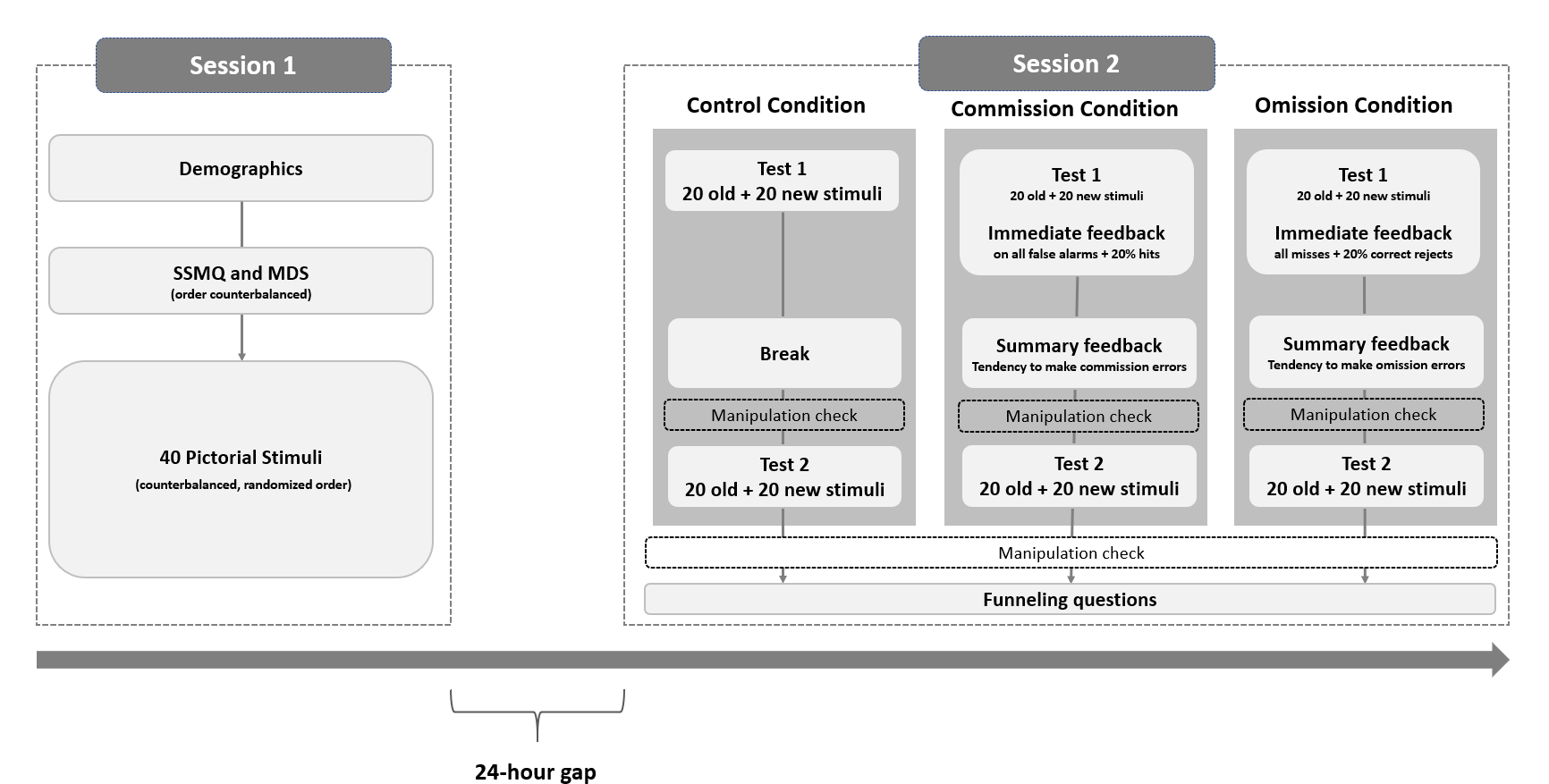
Immediately after the first recognition test, participants in the feedback-commission condition received inaccurate summary feedback claiming that they made several mistakes in the test, misidentifying as ‘old’ several scenes that were not seen in Session 1. Participants in the omission error condition received inaccurate summary feedback claiming that they have made several mistakes in the test, misidentifying as ‘new’ several scenes that were actually seen in Session 1. Although the majority of participants made both commission and omission errors, consistent with the manipulation during the first test, they only received feedback focusing on one type of error (or no feedback, for participants in the control group). Moreover, irrespective of their actual performance, they were led to believe that they have made more commission/omission errors than the average participant. Accompanying the feedback, we added general descriptions of memory distrust (similar to Dudek & Polczyk, 2023) to increase the credibility of the feedback (see Appendices - Feedback Script). They were further reminded that participants who ranked in the top 10% of scorers in terms of accuracy in the second memory task would receive an additional monetary bonus ($3), and that it was important to take the memory tasks seriously. The feedback page automatically turned over after 1 minute. For participants in the no-feedback condition, they were given an instruction stating that this is a 1-minute break between tests and were reminded of the monetary bonus for top performers, after which the second test will commence. Half of the participants in each condition completed the manipulation checks before the second test.

Then participants completed the second recognition task, in which their memory for the remaining block was tested. All the measures were the same as the first recognition task with no feedback. After the memory tasks, the other half of the participants in each condition completed the manipulation checks. Finally, all participants answered six funneling questions: 1) “To what extent did you take into account your tendency to have memory errors while making recognition judgments in the second memory task?” (1 = *not at all*, to 7 = *very much*); 2) “To what extent did you find the experiment procedures difficult to understand?” (1 = *not difficult at all*, to 7 = *very difficult*); 3) “When completing this study, how seriously did you take answering the questions with care? (Your answer will not affect your payment for the experiment)” (1 = *not serious at all*, to 5 = *very serious*); 4) “Have you seen the materials of this experiment before in other studies” (Yes or No); 5) “What do you think is the purpose of the experiment?” (Open-ended); 6) “Did you notice any errors in the experiment or do you have any suggestions to improve the experiment?” (Open-ended).[[3]](#footnote-3)

We then paid participants, debriefed them, and paid them the bonus where applicable after the data collection was done. We calculated and ranked the accuracy for each participant within their experimental conditions based on the second recognition task only.

**Figure 1**

***Flowchart of the Study Procedures***



**Note.** Dashed lines surrounding manipulation check indicate that the order of manipulation check (before and after Test 2) was randomized.

## Data Analyses

All data analyses were carried out in R (version 4.2.2, R Core Team, 2021). All anonymized datasets and coding scripts are available via the OSF (<https://osf.io/8qbkn/>). Response criterion indices *β* and c were calculated using the *psycho* package (Makowski, 2018), β was calculated based on the likelihood ratio of the two distributions (noise and signal) while c represents the distance between the response criterion and the unbiased point, expressed in units of standard deviations. A higher value of either *β* or c would indicate a greater tendency to recognize stimuli as new instead of old (i.e., a more conservative response criterion). Since *β* is based on a ratio and more likely to violate distribution assumptions than c (Zhang et al., 2024), when the results of the two indices diverged, we gave more weight to c when reaching conclusions. Please see Table 1 for the plans on hypothesis testing.

### Exclusions

We excluded from analyses the data from participants who (1) failed one or more of the attention check questions, i.e., did not respond to the attention checks by selecting the correct or the required answer; (2) reported being not serious when completing the experiment (score <3); (3) reported having seen the photo materials before in other studies; or (4) indicated, in the open question regarding the goal of the experiment, that they distrusted the feedback they received. Two independent coders coded all participants’ open-text responses regarding the latter exclusion criterion based on the gist of their responses. For example, if a participant mentioned gaslighting, manipulation, or making people doubting their memories by means of feedback, they were classified as distrusting the feedback (Cohen’s Kappa = .77). Any disagreements were resolved during discussion.

**Table 1**

***Study Design Table***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis Plan** | **Rationale** | **Interpretation** | **Theory** |
| Is the manipulation of memory distrust toward commission effective? | The manipulation of memory distrust toward the commission is effective. | For the single-item measuring state memory distrust toward commission, the expected effect size of the manipulation is Cohen’s *d* = 1.00 given previous research (Dudek & Polczyk 2023).  We will recruit 630 participants (210 in each condition) using Connect.  We performed simulation-based power analysis for minimal effect testing and equivalence testing (see Riesthuis, 2024 for the tutorial) for the pairwise comparisons of state memory distrust. In this analysis, the minimal effect testing calculated the percentage of results wherein the lower bound of 80% CI of the effect is greater than SESOI (raw score difference of state memory distrust = 1.6). The equivalence testing calculated the percentage of results wherein the 80% CI falls within the two equivalence bounds [-1.6, 1.6]. When the true effect size on memory distrust is a raw score difference of 2 (Cohen’s *d* = 1), a group size of 210 participants will have 80% power to detect the minimal effect of 1.6. | One-way ANOVA (Feedback: commission vs. omission vs. no feedback) was conducted for state distrust toward commission.  Only if the lower bound of the 80% CI on the effect size is not lower than Cohen’s *d* = 0.80 (raw score difference of 1.6 assuming a SD = 2) for the pairwise comparisons between control and distrust-commission, will we consider the manipulation(s) adequate. | Assuming that the correlation between state memory distrust and response criterion c is at least *r* = 0.25, given the minimal effect of memory distrust manipulation on criterion is set to be Cohen’s *d* = 0.20 (see row 3), this requires that the strength of the manipulation to be no smaller than Cohen’s *d* = 0.80. | If the manipulation to increase distrust toward commission errors is successful, we would expect that compared to feedback on omission errors and no feedback conditions, participants in the feedback on commission errors condition will report higher distrust toward falsely remembering.  Only if the lower bound of the 80% CI on the effect size is above the minimal effect of interest (raw score difference of 1.6 with a SD of 2), will we consider the manipulation adequate. If the 80% CI is within the equivalence bounds [-1.6, 1.6], we will conclude that the manipulation did not reach an adequate strength. If the 80% CI partially overlaps the equivalence region, we will suspend judgment. | NA |
| Is the manipulation of memory distrust toward omission effective? | The manipulation of memory distrust toward omission is effective. | For the item measuring state memory distrust toward omission, the expected effect size of the manipulation is Cohen’s *d* = 1.00 given previous research (Dudek & Polczyk 2023).  We will recruit 630 participants (210 in each condition) using Connect.  We performed simulation-based power analysis for minimal effect testing and equivalence testing (see Riesthuis, 2024 for the tutorial) for the pairwise comparisons of state memory distrust. In this analysis, the minimal effect testing calculated the percentage of results wherein the lower bound of 80% CI of the effect is greater than SESOI (raw score difference of state memory distrust = 1.6). The equivalence testing calculated the percentage of results wherein the 80% CI falls within the two equivalence bounds [-1.6, 1.6]. When the true effect size on memory distrust is a raw score difference of 2 (Cohen’s *d* = 1), a group size of 210 participants will have 80% power to detect the minimal effect of 1.6. | One-way ANOVA (Feedback: commission vs. omission vs. no feedback) was conducted for state distrust toward omission.  Only if the lower bound of the 80% CI on the effect size is not lower than Cohen’s *d* = 0.80 (raw score difference of 1.6 assuming an SD = 2) for the pairwise comparisons between control and distrust-omission, will we consider the manipulation(s) adequate. | Assuming that the correlation between state memory distrust and response criterion c is at least *r* = 0.25, given the minimal effect of memory distrust manipulation on criterion is set to be Cohen’s *d* = 0.20 (see row 3), this requires that the strength of the manipulation to be no smaller than Cohen’s *d* = 0.80. | If the manipulation to increase distrust toward omission errors is successful, we would expect that compared to feedback on commission errors and no feedback conditions, participants in the feedback on omission errors condition will report higher distrust toward forgetting.  Only if the lower bound of the 80% CI on the effect size is above the minimal effect of interest (raw score difference of 1.6 with a SD of 2), will we consider the manipulation adequate. If the 80% CI is within the equivalence bounds [-1.6, 1.6], we will conclude that the manipulation did not reach an adequate strength. If the 80% CI partially overlaps the equivalence region, we will suspend judgment. | NA |
| Will the increase in memory distrust toward commission lead to a more conservative response criterion while the increase of memory distrust toward omission lead to a more liberal response criterion? | The increase in memory distrust toward commission leads to a more conservative response criterion while the increase in memory distrust toward omission leads to a more liberal response criterion. | We performed simulation-based power analysis for minimal effect testing and equivalence testing (see Riesthuis, 2024 for the tutorial) for the pairwise comparisons of response criterion. In this analysis, the minimal effect testing calculated the percentage of results wherein the lower bound of 80% CI of the effect is greater than SESOI (c diff = 0.06). The equivalence testing calculated the percentage of results wherein the 80% CI falls within the two equivalence bounds, c: [-0.06, 0.06]. This analysis on criterion c showed that when the true effect size is c diff = 0.15 (Cohen’s *d* = 0.5), a group size of 100 participants will have 80% power to detect the minimal effect. We therefore decide to set the minimum number of participants as 210 per group, which is the larger required sample size. | One-way ANOVA (Feedback: commission vs. omission vs. no feedback) was conducted for response criterion *β* and c.  Simple effects would reveal that the response criterion is the most conservative in the feedback-commission condition, followed by the control condition. The feedback-omission condition will have the least conservative (most liberal) response criterion.  Only if the lower bound of the 80% CI on the effect size is equal to or greater than Cohen’s d = 0.20 (difference in c is 0.06, assuming an SD = 0.30) for the pairwise comparisons (distrust-commission vs. control; control vs. distrust-omission), will we consider the hypothesis supported | **Expected Effect Size.** We expected the effect of memory distrust manipulation on criterion shift to be smaller than explicit instructions to respond liberally or conservatively (ca conservative = 0.34, ca liberal = -0.50, ca diff = 0.84, Azimian-Faridani & Wilding, 2006) but larger than implicit biased feedback (Experiment 2: ca conservative = 0.39, ca liberal = 0.02, ca diff = 0.37; Experiment 3: ca conservative = 0.19, ca liberal = 0.05, ca diff = 0.14, Han & Dobbins, 2008). We therefore set a conservative expected effect size (close to Experiment 3 in Han & Dobbins, 2008) as a difference of c = 0.15 between the omission condition and control condition and between control and commission condition.  **Smallest Effect Size of Interest.** Using the data from Study 2 of Zhang et al. (2023c) as the control condition and the expected differences between conditions, we created a synthetic dataset and calculated the SDT indices In the synthetic dataset (See Table A1), the average difference of response criterion c is 0.06 between Feedback-Omission and control or between control and feedback commission with a standard deviation of 0.30 (i.e., Cohen’s *d* = 0.06/0.30 = 0.2). With the same standard deviation, the expected effect size translates a difference of Cohen’s *d* = 0.50. | If our hypothesis is supported, we would expect a main effect of the manipulation. Simple effects would reveal that the response criterion is the most conservative in the commission feedback condition, followed by the control condition. The omission feedback condition will have the least conservative and most liberal response criterion.  Only if the lower bound of the 80% CI on the effect size of pairwise comparisons (distrust-commission vs. control, control vs. distrust omission) is above the minimal effect of interest (criterion c difference of 0.06 with a *SD* of 0.30), will we consider the hypothesis supported. If the 80% CI is within the equivalence bounds [-0.06, 0.06], we will conclude that the hypothesis is rejected and accepted the null hypothesis. If the 80% CI partially overlaps the equivalence region, we will suspend judgment. | If our manipulation is successful yet we do not observe the shift in response criterion, the notion that people calibrate their response criterion based on their memory appraisal will be disconfirmed. The results will have an impact on the frameworks of memory reporting regulation by eyewitnesses. |
| Is state memory distrust toward commission errors associated with a more conservative response criterion while state memory distrust toward omission errors associated with a more liberal response criterion? | State memory distrust toward commission errors is associated with a more conservative response criterion while state memory distrust toward omission errors is associated with a more liberal response criterion. | Sensitivity analysis showed that a sample of 630 would allow us to detect a slope of 0.015 c units/Likert unit of state memory distrust (*α* = .05 and 1-*β* = .80) in linear regression (See appendices- Sensitivity Analysis protocol) | Linear regression with response criterion c (SD = 0.30) as the DV and either state memory distrust toward commission or omission (SD = 2.00) as IV. 90% CI was calculated for the regression coefficients to compare against the SESOI. | We expect that the correlation between state memory distrust and response criterion c is at least *r* = 0.25 (raw slope = 0.0375 c units/Likert unit of state memory distrust). | If our hypothesis is supported, we would see that memory distrust toward commission positively predicts β and c while memory distrust toward omission negatively predicts β and c. | Even if we receive support that the manipulation affects response criterion, if state memory distrust is not associated with response criterion as we hypothesized, our hypothesis that state memory distrust influences response criterion will be disconfirmed and the effect of manipulation might be explained by other mechanisms. |

# Results

## Primary Analysis

### Manipulation Check: State Distrust Measures

Two one-way ANOVAs, each with the experimental condition (Feedback: commission vs. omission vs. no feedback) as a between-participant factor, were run for the two manipulation checks (state distrust toward commission and omission) separately. Regarding state distrust toward commission, we failed to detect a statistically significant main effect of the experimental condition, *F* (2, 619) = 2.75, *p* = .065, η2 = .009, 95% [.00, .03][[4]](#footnote-4). The mean difference between the commission condition (*M* = 4.49, *SD* = 2.36) and control condition (*M* = 3.96, *SD* = 2.18) was 0.55 (Cohen’s *d* = 0.23). The equivalence test was significant, *t* (375.21) = -5.70, *p* <.001, given equivalence bounds of -1.82 and 1.82 and an alpha of 0.1. The 80% CI [0.24, 0.82] was completely within the equivalence bounds. For the rating distributions, please see supplementary materials-Plots for State Memory Distrust.

Similar results were found for the state distrust toward omission analysis. We did not detect a statistically significant main effect of the experimental condition, *F* (2, 617) = 2.96, *p* = .053, *η*2 = .010, 95% [.00, .03][[5]](#footnote-5). The mean difference between the omission condition (*M* = 5.07, *SD* = 2.53) and control condition (*M* = 4.52, *SD* = 2.17) was 0.55 (Cohen’s *d* = 0.23). The equivalence test was significant, *t* (410.62) = -5.92, *p* <.001, given equivalence bounds of -1.89 and 1.89 (on a raw scale) and an alpha of 0.1. The 80% CI [0.26, 0.84] was completely within the equivalence bounds.

Looking at the data from another point of view, for all pairwise comparisons of the state distrust toward commission ratings between the commission condition and the control conditions, 50.83% of the time, the ratings from the commission condition were higher while 37.13% of the time, the reverse was true. For all pairwise comparisons of the state distrust toward omission ratings between the omission condition and the control conditions, 50.58% of the time, the ratings from the commission condition were higher while 38.09% of the time, the reverse was true, reaffirming that the effects of manipulation were smaller than we have hoped for. We therefore concluded that the manipulation did not reach an adequate strength in manipulating state memory distrust. Despite the manipulation not producing the intended effects, we proceeded with the main analyses to examine potential effects that may still inform our research question.

### The Effect of Feedback on Response Bias

To examine whether the experimental feedback manipulation would cause a shift in response bias, we first calculated the SDT response criterion indices *c* and *β* for the second recognition test (see Table 2 for descriptive statistics). Recall that a higher value of c or *β* would indicate a more conservative response criterion (i.e., more likely to make ‘new’ judgments). Next, we performed one-way ANOVAs with the experimental condition (Feedback: commission vs. omission vs. no feedback) as a between-participant factor, and with c or *β* in the second test as the dependent variable, respectively. The effects of manipulations are reported averaging over manipulation check order conditions[[6]](#footnote-6).

**Table 2**

***Descriptive Statistics of Response Criterion***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | c | | *β* | |
|  | *Mean* | *SD* | *Mean* | *SD* |
| Control | 0.16 | 0.46 | 1.70 | 1.44 |
| Feedback Commission | 0.40 | 0.48 | 2.32 | 1.84 |
| Feedback Omission | -0.15 | 0.56 | 1.32 | 1.36 |

Results indicated that there was a large and statistically significant main effect of experimental conditions on criterion c, *F* (2, 619) = 60.06, *p* < .001, *η*2 = .16, 95% [.11, .21]. Criterion c was the highest (i.e., strictest) in the feedback-commission condition, followed by the control condition, and the feedback omission condition. All pairwise comparisons were statistically significant (*p*s < .001). The equivalence test for the commission-control comparison was non-significant, *t* (381.47) = 2.98, *p* = .998, given equivalence bounds of -0.094 and 0.094, equal to a Cohen’s *d* = 0.2) and an alpha of 0.1. The 80% confidence interval for the mean difference between commission condition and control condition was [0.17, 0.29], which lies entirely above the equivalence bounds. The equivalence test for the omission-control comparison was non-significant as well, *t* (401.98) = -4.40, *p* > .999, given equivalence bounds of -0.102 and 0.102 (equal to a Cohen’s *d* = 0.2) and an alpha of 0.1. The 80% confidence interval for the mean difference between commission condition and control condition was [-0.38, -0.26], which lies entirely below the equivalence bounds.

As for the analysis on *β,* results showed that there was a statistically significant main effect of experimental conditions, *F* (2, 619) = 20.70, *p* < .001, *η*2 = .06, 95% [.03, .10]. Similarly, *β* was the highest in the feedback-commission condition, followed by the control condition, and the feedback omission condition. All pairwise comparisons were statistically significant (*p*s < .026). The equivalence test for the commission-control comparison was non-significant, *t* (337.98) = 1.73, *p* = 0.958, given equivalence bounds of [-0.33, 0.33] and an alpha of 0.1. The 80% confidence interval for the mean difference between commission condition and control condition was [0.40, 0.83], which lies entirely above the equivalence bounds. The equivalence test for the omission-control comparison was non-significant as well, *t* (435.95) = -0.78, *p* = .781, given equivalence bounds of [-0.28, 0.28] and an alpha of 0.1. The 80% confidence interval for the mean difference between commission condition and control condition was [-0.60, -0.16], which partly overlaps with the equivalence region. Given that *β* is a ratio statistic and more prone to bias, we draw conclusions primarily based on the results of criterion c. The results therefore suggest that both forms of our experimental manipulation led to a significant criterion shift. As hypothesized, the response criterion was the most conservative in the feedback-commission condition, followed by the control condition. The feedback-omission condition had the least conservative (most liberal) response criterion.

Further, we ran regression analyses on response criterion in the second test, with state memory distrust toward commission and omission errors being entered in the model separately. We did not find a significant association of criterion c with either state memory distrust toward commission (*B* = -0.002, *SE* = 0.01, *p* = .821) or state memory distrust toward omission (*B* = -0.003, *SE* = 0.01, *p* = .739). Likewise, neither state memory distrust toward commission (*B* = -0.02, *SE* = 0.03, *p* = .556) nor state memory distrust toward omission (*B* = -0.04, *SE* = 0.03, *p* = .114) was a significant predictor of *β*.

Taken together, these results suggest that the experimental manipulation mainly influenced criterion. However, this effect on criterion was likely not through the hypothesized effect on state memory distrust, assuming that our manipulation check measures were indeed measuring state memory distrust as intended. In the following section, we performed a few exploratory analyses to gain more insight into the current results.

## Exploratory Analyses

### Effect of Feedback on Recognition, Recollection, and Belief

To examine whether the experimental manipulation influenced recognition judgments, recollection, and belief ratings in the second test, we ran linear mixed models (LMM) each with one of these variables as the dependent variable, and with the experimental conditions as fixed effects. We also included random intercepts for stimuli and participants in the models. These exploratory analyses showed that the experimental manipulations had an effect on recognition, recollection and belief judgments (See Table 3). Compared to participants in the control condition, participants in the feedback-commission condition made fewer recognition judgments (i.e., ‘old’), and reported lower recollection and belief ratings, whereas participants in the feedback-omission condition made more recognition judgments, and reported higher recollection and belief ratings, thus suggesting a calibration of metamemory judgments after receiving feedback.

We performed another analysis on the recognition judgments including belief rating as a covariate. Results showed that after controlling for belief ratings, the effects of feedback manipulation became non-significant (Commission: *B* = 0.06, *SE* = 0.22, *p* = .796; Omission: *B* = 0.23, *SE* = 0.21, *p* = .258).

**Table 3**

***The Effect of Experimental Manipulation on Recollection, Belief, and Recognition***

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Recollection | | | Belief | | | Recognition | | |
| **Fixed effects** |  |  |  |  |  |  |  |  |  |
|  | *B* | *SE* | *p* | *B* | *SE* | *p* | *B* | *SE* | *p* |
| Intercept | 4.01 | 0.09 | <.001 | 4.21 | 0.08 | <.001 | -0.20 | 0.06 | .001 |
| Commission vs. control | -0.28 | 0.12 | .017 | -0.46 | 0.10 | <.001 | -0.31 | 0.07 | <.001 |
| Omission vs. control | 0.48 | 0.11 | <.001 | 0.46 | 0.10 | <.001 | 0.45 | 0.07 | <.001 |
| **Random effects** |  |  |  |  |  |  |  |  |  |
|  |  | SD | ICC |  | SD | ICC |  | SD | ICC |
| Participants |  | 1.13 | 0.16 |  | 0.93 | 0.12 |  | 0.61 | 0.10 |
| Stimuli |  | 0.44 | 0.02 |  | 0.44 | 0.03 |  | 0.38 | 0.04 |
| Pseudo-R² (fixed/total) | 0.01/ 0.19 | | | 0.02/ 0.16 | | | 0.02/ 0.16 | | |

### Correlation between State and Trait Memory Distrust

Given that state memory distrust was not correlated with response criterion c, we investigated the validity of the state memory distrust items by examining their partial association with the SSMQ and MDS. Recall that the SSMQ primarily concerns errors of omission, whereas the MDS primarily concerns errors of commission; recall also that we reverse-scored the SSMQ such that high scores on both measures would signify higher levels of memory distrust. Our results showed that MDS scores positively predicted state memory distrust toward commission (*ß* = 0.35, *SE* = 0.04, *p* <.001) whereas SSMQ did not (*ß* = -0.07, *SE* = 0.04, *p* = .124). For state memory distrust toward omission, though, both MDS and SSMQ positively predicted the ratings (MDS: *ß* = 0.29, *SE* = 0.04, *p* <.001; SSMQ: *ß* = 0.17, *SE* = 0.04, *p* <.001). These results therefore suggest that the two manipulation checks did capture elements of memory distrust to an extent, but that they may not have been sufficiently selective in capturing the distinctive aspects of distrust toward commission or omission.

### Correlation of State Memory Distrust with Recollection and Belief

To further explore the validity of the state memory distrust measures, we examined whether they predicted belief and recollection ratings. We ran linear mixed models (LMM) with either recollection or belief rating in the second test as the dependent variable, and state memory distrust toward commission errors and omission errors as fixed effects. We also included random intercepts for stimuli and participants in the models. State memory distrust toward commission errors positively predicted (*B* = 0.06, *SE* = 0.03, *p* = .018) while state memory distrust toward omission negatively predicted recollection ratings (*B* = -0.05, *SE* = 0.02, *p* = .037). No significant associations were found for belief ratings (Distrust Commission: *B* = 0.03, *SE* = 0.02, *p* = .226; Distrust Omission: *B* = -0.02, *SE* = 0.02, *p* = .428) nor recognition judgments (Distrust Commission: *B* = 0.004, *SE* = 0.02, *p* = .796; Distrust Omission: *B* = -0.002, *SE* = 0.02, *p* = .873). Therefore, in our data, people who considered themselves more likely to make commission errors at the present moment reported slightly higher recollection ratings, while people who were consider themselves more likely to make omission errors at the present moment reported slightly lower recollection ratings only. The non-significant correlation between the state memory distrust measures and belief and recognition judgments could suggest that our hypothesized calibration was at work. For example, although people with higher state memory distrust toward commission had on average more recollective experiences compared with their low distrust counterparts, they could have taken into their tendency of making memory errors and adjusted belief ratings downwardly, resulting the non-significant relationship between state memory distrust and belief ratings.

Further analyses on belief and recognition judgements with recollection ratings included as a covariate showed that once conditioned on recollection, state memory distrust toward commission was negatively associated, and state memory distrust toward omission positively associated, with belief ratings (Distrust Commission: *B* = -0.03, *SE* = 0.01, *p* = .016; Distrust Omission: *B* = -0.03, *SE* = 0.01, *p* = .008) and recognition judgments (Distrust Commission: *B* = -0.16, *SE* = 0.05, *p* = .004; Distrust Omission: *B* = -0.03, *SE* = 0.05, *p* = .008), consistent with the calibration hypothesis. Note that the correlations were small and the previous analysis showed that our state memory distrust measure may not be up to the task; we therefore caution against overinterpreting this result. Below we further test the calibration with the established trait measures.

### Correlation of Trait Memory Distrust with Recollection and Belief

Because of the unexpected results regarding state memory distrust, we run exploratory analyses with trait memory distrust measures to examine potential calibration in reporting. First, we examined the correlation between trait memory distrust and recollection ratings. Results showed that the MDS positively predicted recollection while the SSMQ negatively predicted recollection, consistent with Zhang et al. (2024). Then we ran a second model for belief rating, with recollection rating added as a covariate. This model reveals that when conditioned on recollection rating, the MDS negatively predicted while the SSMQ positively predicted belief ratings. Put differently, when recollection judgments are same, people who considered themselves as prone to making commission errors tended to give lower belief ratings while people who considered themselves as prone to making omission errors tended to give higher belief ratings. Similar results were found regarding recognition judgments. This again seems to corroborate our calibration hypothesis.

**Table 4**

***The Effect of Trait Memory Distrust on Recollection, Belief, and Recognition***

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Recollection | | | Belief | | | Recognition | | |
| **Fixed effects** |  |  |  |  |  |  |  |  |  |
|  | *B* | *SE* | *p* | *B* | *SE* | *p* | *B* | *SE* | *p* |
| Intercept | 4.03 | 0.06 | <.001 | 0.44 | 0.08 | <.001 | -7.58 | 0.12 | <.001 |
| MDS | **0.27** | 0.05 | <.001 | **-0.06** | 0.02 | .009 | **-0.62** | 0.09 | <.001 |
| SSMQ | **-0.36** | 0.05 | <.001 | **0.11** | 0.02 | <.001 | **0.59** | 0.09 | <.001 |
| Recollection |  |  |  | 0.93 | 0.00 | <.001 | 1.87 | 0.02 | <.001 |
| **Random effects** |  |  |  |  |  |  |  |  |  |
|  |  | SD | ICC |  | SD | ICC |  | SD | ICC |
| Participants |  | 1.03 | 0.13 |  | 0.53 | 0.29 |  | 1.88 | 0.51 |
| Stimuli |  | 0.43 | 0.02 |  | 0.06 | 0.00 |  | 0.38 | 0.02 |
| Pseudo-R² (fixed/total) | 0.01/ 0.17 | | | 0.88/ 0.91 | | | 0.80/ 0.91 | | |

Recall that we explicitly asked participants whether they calibrated their responses in the experiment (on a 7-point scale). Participants on average gave a rating of 4.48 (*SD* = 1.76) and 60.45% of all participants give a rating of 5 and above, indicating that the majority of them did take into account their (perceived) tendency to make memory errors. To further test this possibility of calibration, we included this calibration rating in the models predicting belief ratings and recognition judgments and had it interact with each of the two memory distrust measures (See Table 5). Results showed that the interaction term between calibration rating and MDS was statistically significant as a predictor of belief ratings. Put differently, when having the same recollection, people who considered themselves more prone to commission errors and who took into consideration this tendency, reported lower belief ratings. No other interaction terms were statistically significant. Additional analysis was performed to examine whether this self-reported calibration moderated response criterion c in the second test. We did not find a significant interaction effect on criterion c between calibration effort and experimental manipulation (Calibration \* Commission vs. control: *B* = 0.07, *SE* = 0.05, *p* = .139; Calibration \* Omission vs. control: *B* = 0.01, *SE* = 0.05, *p* = .771).

**Table 5**

***Self-reported Calibration Effort and Belief as well as Recognition***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Belief | | | Recognition | | |
| **Fixed effects** |  |  |  |  |  |  |
|  | *B* | *SE* | *p* | *B* | *SE* | *p* |
| Intercept | 0.46 | 0.02 | <.001 | -7.56 | 0.13 | <.001 |
| Calibration | -0.04 | 0.02 | .083 | 0.03 | 0.09 | .751 |
| MDS | -0.06 | 0.02 | .028 | -0.63 | 0.09 | <.001 |
| SSMQ | 0.11 | 0.02 | <.001 | 0.59 | 0.09 | <.001 |
| Recollection | 0.93 | 0.00 | <.001 | 1.87 | 0.02 | <.001 |
| MDS \* Calibration | -0.11 | 0.02 | <.001 | -0.11 | 0.09 | .234 |
| SSMQ \*Calibration | 0.02 | 0.02 | .345 | 0.08 | 0.09 | .377 |
| **Random effects** |  |  |  |  |  |  |
|  |  | SD | ICC |  | SD | ICC |
| Participants |  | 0.52 | 0.29 |  | 1.88 | 0.51 |
| Stimuli |  | 0.06 | 0.00 |  | 0.38 | 0.02 |
| Pseudo-R² (fixed/total) | 0.88/ 0.91 | | | 0.80/ 0.91 | | |

# Discussion

In this experiment, we aimed to experimentally manipulate state memory distrust toward commission and omission separately, and examine their effect on criterion shift. Although we observed hypothesized effects of the manipulation on criterion shift, we did not our find our hypothesized effect of the manipulation on state memory distrust itself. As a matter of fact, our analysis showed that the experimental manipulation had a negligible effect on the state memory distrust measures. This may indicate that the feedback manipulation was not strong enough to alter participants’ memory distrust levels, or that the state memory distrust measure did not adequately capture the change in state memory distrust levels. In the following section, we discuss the results in relation to the literature.

### Biased Feedback Effects on Criterion Shift

Expanding on previous research which showed that biased positive feedback can shift participants’ response criterion in an either more lax or strict way (Han & Dobbins, 2008), we showed that biased negative feedback can have a similar effect. Consistent with our expectations, the observed differences in criterion *c* between the feedback-commission and feedback-omission conditions in the current study were *c*diff = 0.40 in the first test and *c*diff = 0.55 in the second test, in between the effect sizes observed in prior studies using explicit instruction (Azimian-Faridani & Wilding, 2006) and studies using biased positive feedback (Han & Dobbins, 2008). Equivalence tests showed that the observed effects were significantly larger than our smallest effect size of interest, which corresponds to a Cohen’s *d* of 0.2. In fact, the point estimate of pairwise comparison between feedback-commission condition and feedback-omission condition, cdiff = 0.55, translates to Cohen’s *d* = 1.0. To put this effect size in more intuitive terms, there is a 76.0% chance that a person randomly selected from the feedback-commission group would have a stricter response criterion than a person randomly selected from the feedback-omission group (probability of superiority).

The large observed effect of the feedback manipulation on criterion shift is in stark contrast with the fact that we did not find the expected effect of the feedback manipulation on state memory distrust, even though the differences in state memory distrust between feedback conditions and control conditions were in the expected directions. Equivalence tests, however, showed that the effects were too small to be considered relevant. Further, we did not find evidence that state memory distrust was associated with response criterion, further casting doubt on the hypothesized mechanism.

There could be several explanations for this surprising finding. The first explanation is that the experimental manipulation affected response criterion not through state memory distrust but through another mechanism, such as a conscious shift of response strategy given the feedback and the incentive to achieve high accuracy. That is, participants could have strategically responded to feedback on commission (omission) errors by choosing ‘new’ (‘old’) whenever their memory of the stimulus was ambiguous and uncertain. This adjustment of strategy does not necessarily require participants to doubt their memory. A second possibility is that including both aspects of memory distrust in the summary feedback might have introduced ambiguity, which unintentionally influenced participants’ responses to the manipulation checks. Third, it is possible the feedback on performance *did* influence levels of memory distrust, but not on participants’ *global* memory appraisals. Since our manipulation check questions were framed quite generally (e.g., “At this moment, if I am asked to retrieve something from memory, I think I likely would remember things that did not happen”), it could be that participants’ answers were based on expectations about their memory performance not restricted to the test format. In this case, if we were to measure participants’ state memory distrust with a more task-specific prompt such as ‘if I continue with this memory test, I think I likely would recognize photos that in fact were not in yesterday’s session’, then we might capture more precise assessments of state memory distrust that are more strongly influenced by the feedback. Third, although we deliberated on the item construction for the manipulation check questions and they appear to have good face validity, it is possible that the participants interpreted them in a different way. The fact that state memory distrust toward omission was correlated more strongly with the MDS than with the SSMQ, might indicate cause to speculate that some participants did not interpret the item as intended. Finally, the use of a single-item measure may have also suffered from low-reliability, allowing too much noise in responses.

Despite these possible issues with the measure, exploratory analyses revealed some meaningful patterns. When recollection was controlled for, participants with higher (vs. lower) state memory distrust toward commission errors tended to give lower belief ratings and were less likely to recognize an item as old. On the other hand, a reversed pattern is found regarding state memory distrust toward omission. Given these possibilities, we conclude that the current evidence cannot offer a clear picture whether or not state memory distrust indeed played a role in the response criterion shift.

### Evidence for Calibration based on Trait Memory Distrust Measures

Trait memory distrust could be associated with metacognitive judgments through different mechanisms. First, trait memory distrust could be partly reflecting objective memory functioning. In this case, we would expect that for example, people with high distrust toward commission errors indeed are more likely to have more vivid recollective experiences and therefore higher recollection and belief ratings. Second, when motivated, people can also calibrate their belief judgments based on their tendency of committing memory errors, leading people with high distrust toward commission to down-regulate their belief ratings and people with high distrust toward omission to up-regulate their belief ratings. Our data suggest that people who considered themselves as more prone in general to making commission errors reported, on average, the highest recollection and belief ratings, while the reverse was true for those with higher distrust toward omission errors. Further analysis also showed that, when conditioned on recollection ratings, people with higher distrust toward commission errors reported on average lower belief ratings, whereas those with higher distrust toward omission errors reported on average higher belief ratings. Taken together, the results suggest that people indeed can calibrate their response criterion when being incentivized; however, the overall effect is much smaller than anticipated.

### Limitations and Future Directions

As discussed earlier, a major limitation of the current study is the lack of clarity over the extent to which we successfully manipulated state memory distrust. Our first attempt at manipulating the two aspects of memory distrust revealed less than optimal results. While we observed some differences in state memory distrust between conditions, the manipulation did not produce the expected changes in state memory distrust that would align with our hypothesized mechanism of criterion shift. In the recent work of Dudek and Polczyk (2023), the authors also found some unexpected results regarding their memory distrust manipulation. Although they found significant differences in state memory distrust between conditions, this effect did not lead to an increase in susceptibility to misinformation as hypothesized. This suggests that memory distrust, particularly in an experimental context, may be more difficult to manipulate than previously thought. The aim to manipulate the two aspects of distrust separately further increased its difficulty. The disconnect between memory distrust and susceptibility to external influences highlights the need for more nuanced and effective manipulations. Based on these findings, we recommend future research to build on our work as well as those of others to further refine the manipulation procedures (e.g., measuring state memory distrust both pre and post manipulation) and establish better state memory distrust measures, such as multi-item scale with established reliability and validity. Research on the effects of memory distrust on various memory phenomena needs a stronger foundation.

# Conclusion

This study aimed to manipulate state memory distrust toward commission and omission separately, and examine their effect on criterion shift. Our findings indicate that feedback manipulation did influence response criteria: participants who were told they frequently made commission errors adopted a stricter response criterion, while those who received feedback about omission errors shifted toward a more liberal criterion. However, we did not find sufficient evidence that our manipulations had the expected effect on state memory distrust. Therefore, we could not offer evidence that the effect of feedback on criterion shift was through memory distrust. Overall, these results highlight the potential for feedback to influence metacognitive control of memory, but also point to the limitations in effectively manipulating aspects of state memory distrust separately. More research is needed to refine and improve state memory distrust manipulation and measurement.

# References

Azimian-Faridani, N., & Wilding, E. L. (2006). The influence of criterion shifts on electrophysiological correlates of recognition memory. *Journal of Cognitive Neuroscience*, *18*(7), 1075-1086.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Blank, H. (2009). Remembering: A theoretical interface between memory and social psychology. *Social Psychology*, *40*(3), 164-175.

Blank, H. (2017). Recollection, belief and metacognition: A reality check. *Memory*, *25*(7), 869-875.

Clark, A., Nash, R. A., Fincham, G., & Mazzoni, G. (2012). Creating non-believed memories for recent autobiographical events. *PLoS One*, *7*(3), e32998.

Dienes, Z. (2021). Obtaining evidence for no effect. *Collabra: Psychology*, *7*(1), 28202.

Dudek, I., & Polczyk, R. (2023). Memory distrust and suggestibility: A registered report. *Legal and Criminological Psychology*. DOI: 10.1111/lcr p.12 249

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1, pp. 1969-2012). Wiley.

Gudjonsson, G. H., & MacKeith, J. A. C. (1982). False confessions: Psychological effects of interrogation. *Reconstructing the past: The role of psychologists in criminal trials*, 253-269.

Han, S., & Dobbins, I. G. (2008). Examining recognition criterion rigidity during testing using a biased-feedback technique: Evidence for adaptive criterion learning. *Memory & Cognition*, *36*(4), 703-715.

Han, S., & Dobbins, I. G. (2009). Regulating recognition decisions through incremental reinforcement learning. *Psychonomic Bulletin & Review*, *16*, 469-474.

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior Research Methods*, *49*, 457-470.

Kuczek, M., Szpitalak, M., & Polczyk, R. (2018). Psychometric properties and correlates of the Polish version of the Squire Subjective Memory Questionnaire (SSMQ). *Personality and Individual Differences*, *120*, 271-275.

Li, C., Otgaar, H., Muris, P., Zhang, Y., & Wang, J. (2023). Inducing emotionally negative nonbelieved memories using negative pictures. *Memory & Cognition*, 1-16. https://doi.org/10.3758/s13421-023-01441-3

Li, C., Wang, J., & Otgaar, H. (2020). Creating nonbelieved memories for bizarre actions using an imagination inflation procedure. *Applied Cognitive Psychology*, *34*(6), 1277-1286.

Makowski, (2018). The psycho Package: an Efficient and Publishing-Oriented Workflow for Psychological Science. *Journal of Open Source Software, 3*(22), 470. <https://doi.org/10.21105/joss.00470>

Mazzoni, G. A. L., Loftus, E. F., & Kirsch, I. (2001). Changing beliefs about implausible autobiographical events: A little plausibility goes a long way. *Journal of Experimental Psychology: Applied, 7,* 51–59.

Mazzoni, G., Scoboria, A., & Harvey, L. (2010). Nonbelieved memories. *Psychological Science*, *21*(9), 1334-1340.

Nash, R. A., Saraiva, R. B., & Hope, L. (2023). Who doesn't believe their memories? Development and validation of a new Memory Distrust Scale. Journal of Applied Research in Memory and Cognition, 12(3), 401–411. <https://doi.org/10.1037/mac0000061>

Nash, R. A., Wade, K. A., Garry, M., & Adelman, J. S. (2017). A robust preference for cheap-and-easy strategies over reliable strategies when verifying personal memories. *Memory*, *25*(7), 890-899.

Ost, J. (2017). Adults’ retractions of childhood sexual abuse allegations: High-stakes and the (in) validation of recollection. *Memory*, *25*(7), 900-909.

Otgaar, H., Howe, M. L., Smeets, T., & Wang, J. (2016). Denial-induced forgetting: False denials undermine memory, but external denials undermine belief. *Journal of Applied Research in Memory and Cognition*, *5*(2), 168-175.

Otgaar, H., Moldoveanu, G., Wang, J., & Howe, M. L. (2017). Exploring the consequences of nonbelieved memories in the DRM paradigm. *Memory*, *25*(7), 922−933.

Otgaar, H., Scoboria, A., & Mazzoni, G. (2014). On the existence and implications of nonbelieved memories. *Current Directions in Psychological Science*, *23*(5), 349−354.

Otgaar, H., Scoboria, A., & Smeets, T. (2013). Experimentally evoking nonbelieved memories for childhood events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 717–730. [https://doi.org/10.1037/a0029668](https://psycnet.apa.org/doi/10.1037/a0029668)

Riesthuis, P. (2024). Simulation-Based Power Analyses for the Smallest Effect Size of Interest: A Confidence Interval Approach for Minimum-Effect and Equivalence Testing. *Advances in Methods and Practices in Psychological Science*.

Pezdek, K., Finger, K., & Hodge, D. (1997). Planting false childhood memories: The role of event plausibility. *Psychological Science*, *8*, 437–441.

R Core Team. (2021). R: a language and environment for statistical computing. R Foundation for Statistical Computing; 2020.

Scoboria, A., & Henkel, L. (2020). Defending or relinquishing belief in occurrence for remembered events that are challenged: A social‐cognitive model. *Applied Cognitive Psychology*, *34*(6), 1243-1252.

Scoboria, A., Jackson, D. L., Talarico, J., Hanczakowski, M., Wysman, L., & Mazzoni, G. (2014). The role of belief in occurrence within autobiographical memory.Journal of Experimental Psychology: General, 143(3), 1242–1258. <https://doi.org/10.1037/a0034110>

Scoboria, A., Otgaar, H., & Mazzoni, G. (2018). Defending and reducing belief in memories: An experimental laboratory analogue. *Memory & Cognition*, *46*(5), 770-786.

Squire, L. R., Wetzel, C. D., & Slater, P. C. (1979). Memory complaint after electroconvulsive therapy: assessment with a new self-rating instrument. *Biological Psychiatry*. *14*, 791–801.

van Bergen, S., Brands, I., Jelicic, M., & Merckelbach, H. (2010). Assessing trait memory distrust: Psychometric properties of the Squire Subjective Memory Questionnaire. *Legal and Criminological Psychology*, *15*(2), 373-384.

Wade, K. A., Nash, R. A., & Garry, M. (2014). People consider reliability and cost when verifying their autobiographical memories. *Acta Psychologica*, *146*, 28-34.

Wang, J., Otgaar, H., Howe, M. L., Smeets, T., Merckelbach, H., & Nahouli, Z. (2017). Undermining belief in false memories leads to less efficient problem−solving behaviour. *Memory*, *25*(7), 910−921.

Zhang, Y., Battista, F., Thissen, D., Otgaar, H., Wang, J., & Jelicic, M. (2022a). Examining the associations between nonbelieved memories and memory distrust, self-esteem, and rumination. *Psychology of Consciousness: Theory, Research, and Practice.* Advance online publication. [https://doi.org/10.1037/cns0000344](https://psycnet.apa.org/doi/10.1037/cns0000344)

Zhang, Y., Nash, R. A., & Otgaar, H. (2023a). Preference for cheap-and-easy memory verification strategies is strongest among people with high memory distrust. *Memory*, 1-11.

Zhang, Y., Otgaar, H., & Wang, J. (2022b). Memory distrust is related to memory errors, self-esteem, and personality. *Applied Cognitive Psychology*, 1– 10. <https://doi.org/10.1002/acp.3917>

Zhang, Y., Otgaar, H., Nash, R. A., & Rosar, L. (2024). Memory distrust shapes the dynamics of recollection and belief in occurrence over time. Memory.

Zhang, Y., Qi, F., Otgaar, H., Nash, R. A., & Jelicic, M. (2023c). A Tale of Two Distrusts: Memory Distrust towards Commission and Omission Errors in the Chinese Context. *Journal of Applied Research in Memory and Cognition.* DOI: 10.1037/mac0000134

# Appendices

## Table A1

***Descriptives of the Synthetic Dataset***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Condition | d’ | c | *β* | *n*hit | *n*false alarm |
|  | *M (SD)* | *M (SD)* | *M (SD)* | *M (SD)* | *M (SD)* |
| Commission | 1.93 (0.69) | 0.15 (0.31) | 1.71 (1.37) | 15.73 (2.61) | 2.83 (2.42) |
| Control | 1.87 (0.64) | 0.08 (0.29) | 1.30 (0.64) | 15.94 (2.61) | 3.15 (2.38) |
| Omission | 1.86 (0.69) | 0.03 (0.30) | 1.18 (0.62) | 16.14 (2.64) | 3.49 (2.58) |

## Sensitivity Analysis Protocol

**t tests -** Linear bivariate regression: One group, size of slope

**Analysis:** Sensitivity: Compute required effect size

**Input:** Tail(s) = One

Effect direction = Slope > H0

α err prob = 0.05

Power (1-β err prob) = 0.8

Total sample size = 630

Slope H0 = 0

Std dev σ\_x = 2

Std dev σ\_y = 0.3

**Output:** Noncentrality parameter δ = 2.4891574

Critical t = 1.6472836

Df = 628

Slope H1 = 0.0148029

## Table A2

***Analyses on Stimuli Valence and Arousal Ratings***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Effect | *F* statistics | p-value | *η*2 |
| Arousal | Old vs. new | *F* (1, 76) = 1.27 | .263 | =. 016 |
| Block | *F* (1, 76) = 0.04 | .850 | <.001 |
| Old vs. new \* Block | *F* (1, 76) = 0.70 | .407 | =.009 |
| Valence | Old vs. new | *F* (1, 76) = 0.01 | .910 | < .001 |
| Block | *F* (1, 76) = 1.25 | .267 | =.016 |
| Old vs. new \* Block | *F* (1, 76) = 0.17 | .679 | =.002 |

## Feedback Script

### Commission Error Condition

Thank you for completing the first memory test. We have collated your responses and evaluated your performance. In the previous test, while you were pretty good at correctly identifying old photos, in the meantime, you mistakenly recognized several new photos as old, indicating that these new photos were present in the first part of the study while in reality, they were not. Your accuracy at ***correctly identifying*** ***new photos*** is among the lowest **20%** among all participants so far. That is, on average you made more such errors than other participants.

It is possible that you sometimes have difficulty distinguishing events you remember from those you only imagined. You sometimes may also become uncertain whether a recalled event really happened, or whether you saw it on TV or in a movie. Your ability to remember the source of information and the context of events may not be complete.

It's important to note that participants who rank in the **top 10%** in the second memory test will receive a $ 3 bonus and it is important to complete the experiment seriously.

### Omission Error Condition

Thank you for completing the first memory test. We have collated your responses and evaluated your performance. In the previous test, while you were pretty good at correctly identifying new photos, in the meantime, you mistakenly recognized several old photos as new too, indicating that these old photos were not present in the first part of the study while in reality, they were. Your accuracy at ***correctly identifying old photos*** is among the lowest **20%** among all participants so far. That is, on average you made more such errors than other participants.

It is possible that you often experience the Tip-of-the-tongue phenomenon when you know that you remember something but are unable to express it. Sometimes, you may also find it difficult to recall what you were doing after you have taken your mind off it for a few minutes, especially under stress and pressure. In the future, memorizing things and successfully recalling them after a while may become more difficult.

It's important to note that participants who rank in the **top 10%** in the second memory test will receive a $ 3 bonus and it is important to complete the experiment seriously.

### Control Condition

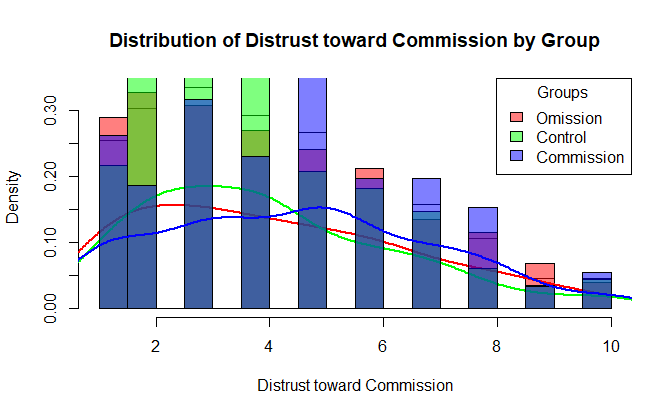
Thank you for completing the first memory test. We have gathered your responses.

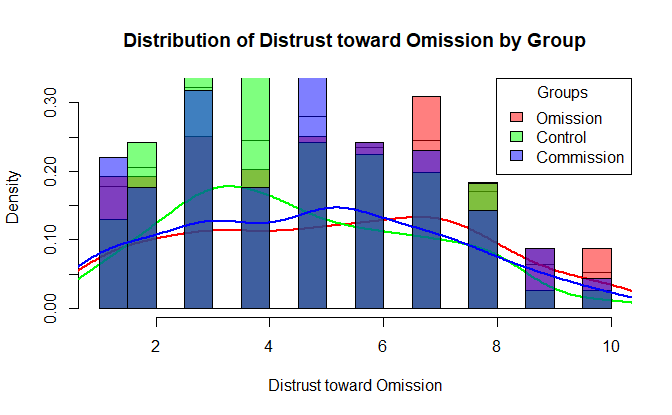
This is a 1-minute break before the second memory test commences.

Please keep in mind that people who rank **top 10%** in the second memory test will get a $ 3 bonus and it is important to complete the experiment seriously.

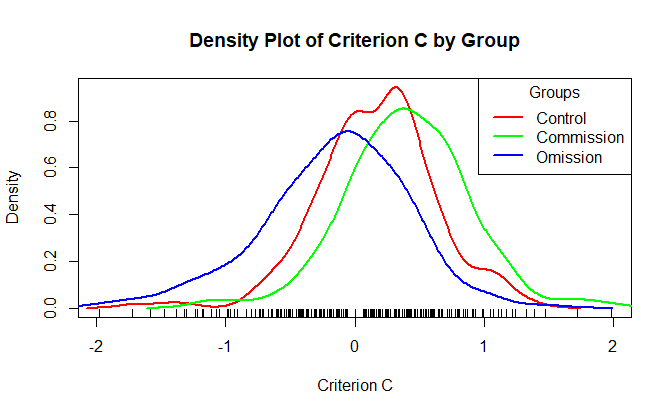
# Supplementary Materials

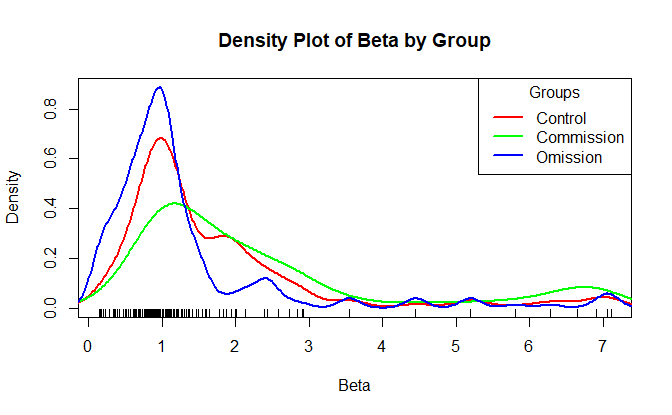
## Plots for State Memory Distrust





## Plots for Response Criterion in the Second Test





## Effect of Feedback on *d*’

We ran a one-way ANOVA to examine the effect of experimental manipulation on sensitivity *d*’ as exploratory analyses. Results showed that there was a small but significant effect, *F* (2, 619) = 3.15, *p* = .043, *η*2 = .01, 95% [.00, .03]. Sensitivity was lower in the feedback-omission condition, compared to the control and the feedback-commission condition (see Table 2), but the pairwise differences did not reach statistical significance (*p*s > .070). Thirty-five participants (5.7%) had negative *d*’ in the second test. Exploratory analyses showed that the results after excluding their data were very similar to the results without exclusion.

## Effect of Feedback on Belief-Recollection Association

In Zhang et al. (2024), the association between recollection and belief ratings were moderated by memory distrust levels with the association being weaker among people who distrust their memory more. We, therefore, explored if the experimental manipulation had a similar effect with a linear mixed model (LMM). The model included belief ratings as the dependent variable, recollection ratings, experimental conditions, and their interaction as fixed effects, and random intercepts for stimuli as well as participants. Both belief and recollection ratings were standardized. Results are presented in Table 3. We noticed that compared to the control condition, the feedback-commission condition had a lower intercept by -0.06 and feedback-omission condition had a higher intercept by 0.01. which was not statistically significant. We also found that the interaction terms were significant: the association was the strongest in the feedback-commission condition, followed by the control condition, and then the feedback-omission condition. The stronger association in the feedback commission compared to the control condition seems to be driven by the lower intercept. On the other hand, the weaker association in the omission condition, in the absence of a higher intercept, seems to be consistent with the previous finding in Zhang et al. (2024), indicating a weaker correspondence between these two metamemory judgments. A second model with recognition as the dependent variable showed that the predictive power of recollection ratings on recognition (i.e., old judgment) was smaller in the two feedback conditions compared to the control condition, as indexed by the negative interaction terms.

***Effect of Feedback on Belief-Recollection Association***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Belief | | | Recognition | | |
| **Fixed effects** | *B* | *SE* | *p* | *B* | *SE* | *p* |
| Intercept | 0.02 | 0.01 | 0.194 | -0.04 | 0.15 | 0.803 |
| Recollection | 0.97 | 0.00 | <.001 | 5.79 | 0.11 | <.001 |
| Condition-Commission | -0.06 | 0.02 | .001 | -0.33 | 0.21 | 0.111 |
| Condition-Omission | 0.01 | 0.02 | .662 | 0.25 | 0.20 | 0.208 |
| Recollection\*Condition-Commission | 0.01 | 0.00 | .021 | -0.79 | 0.14 | <.001 |
| Recollection\*Condition-Omission | -0.02 | 0.00 | <.001 | -0.67 | 0.14 | <.001 |
| **Random effects** |  |  |  |  |  |  |
|  |  | *SD* | *ICC* |  | *SD* | *ICC* |
|  | Participants | 0.19 | .29 | Participants | 1.96 | .53 |
|  | Stimuli | 0.02 | .004 | Stimuli | 0.36 | .02 |
| Pseudo-R² (fixed / total) | 0.88/0.92 | | | 0.80/ 0.91 | | |

1. In reality, the probabilities will likely not be symmetrical. However, we consider the potential effect of the differences of probabilities on estimates insubstantial given that we set a rather conservative SESOI. [↑](#footnote-ref-1)
2. In Zhang et al., (2023), the trait memory distrust measure and memory task were measured three days apart and their correlation was *r* = .19. In the current study, the association is expected to be stronger given that we will measure state memory distrust right before or after the memory task, and plan to increase participants’ motivation to be accurate by raising the performance bonus to $3, 100% of the basic experiment payoff. [↑](#footnote-ref-2)
3. In each data collection cycle, the time window for participating in session 1 was 3 hours. For the second session, we expected a long time (e.g., 6 hours) to complete data collection given the pool would be restricted to participants who have signed up for Session 1. The potential range of gap between Sessions 1 and 2 therefore was 21 (24-3) hours to 30 (24+6) hours. [↑](#footnote-ref-3)
4. We ran an exploratory analysis with manipulation check order as a factor, which did not find evidence that the timing of the manipulation check had a main effect, *F* (1, 616) = 0.84, *p* = .361, η2 = .001, 95% [.00, .01] or that it interacted with experimental conditions, *F* (2, 616) = 1.03, *p* = .359, η2 = .003, 95% [.00, .02]. [↑](#footnote-ref-4)
5. Exploratory analysis with manipulation check order as a factor did not show evidence that the timing of the manipulation check had a main effect, *F* (1, 614) = 0.48, *p* = .488, η2 = .001, 95% [.00, .01] or that it interacted with experimental conditions, *F* (2, 614) = 1.32, *p* = .267, η2 = .004, 95% [.00, .02]. [↑](#footnote-ref-5)
6. Exploratory analysis showed that effects of feedback manipulation on response bias were larger among conditions where participants completed the manipulation check after (vs. before) the second recognition test. However, the moderation effect did not reach statistical significance, *F* (2, 616) = 2.84, *p* = .059, *η*2 = .009, 95% CI [.00, .03]. [↑](#footnote-ref-6)