Who doesn't believe their memories? Development and validation of a new Memory Distrust Scale

Robert A. Nash[1], Renan Benigno Saraiva[2], and Lorraine Hope[2]

[1] School of Psychology, Aston University, UK

[2] Department of Psychology, University of Portsmouth, UK

Data availability: https://osf.io/r8vqu/?view_only=c047e6891c0046c295035df9fd799202

Corresponding author:

Dr Robert A. Nash
School of Psychology
Aston University
Birmingham B4 7ET
United Kingdom

+44 121 204 4522
R.Nash1@aston.ac.uk

Abstract

Many people distrust their memories, because they believe themselves prone to omission errors such as forgetfulness, and/or commission errors such as misremembering. However, the most popular psychometric measure of memory distrust—the Squire Subjective Memory Questionnaire (SSMQ)—only probes people's beliefs about omission errors, not commission errors. The present research describes a new memory distrust measure that assesses the latter kind of metamemorial belief. Two studies (combined N = 797) showed that our 20-item Memory Distrust Scale (MDS) has good psychometric properties, and is correlated with—but distinct from—the SSMQ. Participants in Study 2 described eight childhood events, and rated their recollection of and belief in the occurrence of each. MDS scores were associated with the spontaneous reporting of nonbelieved memories, and predicted belief in occurrence better than did SSMQ scores. Our data suggest that the MDS and SSMQ in combination could better predict individual susceptibility to certain memory errors.

Key words: memory distrust; autobiographical belief; nonbelieved memories; psychometrics; scale development; metamemory

General Audience Summary

Not everyone always trusts their memories. You might believe you are forgetful, for example, or that you are bad at remembering people's names. Researchers have often measured memory distrust using the Squire Subjective Memory Questionnaire (SSMQ), which asks about these types of memory failures. But people might also distrust their memories for other reasons: perhaps because they know they sometimes remember things incorrectly, or remember things that never happened. The SSMQ does not focus on these types of memory failures, so we developed a new self-report scale for this purpose. In two studies (combined $N = 797$), we showed that our 20-item Memory Distrust Scale (MDS) has the characteristics of an effective individual difference measure. For example, people who scored highly on the MDS at one time tended to also score highly at a later time, and their MDS scores were related to their scores on other measures—including the SSMQ—in sensible ways. In Study 2 we asked participants to describe eight events from their childhoods, and to rate (a) how vividly they recalled each event, and (b) how sure they were that each event really happened. The SSMQ and MDS were similarly good at predicting how vividly people recalled each event, but the MDS was better at predicting their belief in the events' occurrence. This is important because we know people's beliefs about their memories can have significant real-world consequences. For example, people who distrust their memories may be vulnerable to believing misinformation, and case studies show us that memory distrust has sometimes led innocent people to confess to crimes. We suggest that the MDS could help us to better predict which people tend to make particular kinds of errors in what they believe and remember, which might help toward identifying and mitigating the consequences of those errors.

Who doesn't believe their memories? Development and validation of a new Memory Distrust Scale

Many people believe they have an unusually terrible ability to remember vital information; from facts, to faces, to past experiences. Memory scholars know how common these pessimistic beliefs are, because members of the public—and indeed our own friends and family members—are frequently eager to tell us about their 'problems' when they learn that we study memory (e.g., Storm, 2011). This apparently widespread lack of faith in one's own memories might or might not be warranted, but it might also have consequences. For example, the renowned American magician and mnemonist Harry Lorayne argued that it has self-fulfilling effects, writing that "memory likes to be trusted. The more you trust it the more reliable and useful it will become" (Lorayne, 2000, p.22). In extreme circumstances, the consequences of distrusting memory can be life-changing, as Gudjonsson and MacKeith (1982) observed when first describing the so-called 'Memory Distrust Syndrome' that led suspects to confess to serious crimes they had never committed. With these kinds of consequences in mind, in this paper we consider how effectively we can measure people's trust and distrust of their memories.

Suppose you were asked about a specific past event, but could not remember it occurring. In this situation you could first contemplate the event's memorability: if it seems easily forgettable, then you might readily accept the suggestion that it happened. But if the event seems memorable, then your lack of memory might make you certain that it cannot possibly have happened. People frequently rely on this 'memorability-based strategy' when making decisions about what they have and have not experienced (see e.g., Ghetti, 2008; Mazzoni & Kirsch, 2002; Strack & Bless, 1994). But consider a person who regularly experiences memory lapses, and who therefore tends to distrust their memory capabilities in general. Such a person, when faced with a suggestion that a memorable event occurred, might be willing to accept this suggestion even if they cannot recall the event. Indeed, in one study, participants who scored highly on trait memory distrust were more likely than their memory-trusting counterparts to accept misinformation about a mock crime, when their own witness statements were covertly altered (van Bergen et al., 2010).

Findings like these point to an important relationship between people's subjective beliefs about their memory abilities, and their susceptibility to memory errors. Yet there is uncertainty concerning this relationship. For instance, van Bergen et al.'s (2010) findings were not replicated in a similar study by the same group of authors, where memory distrust scores did not predict participants' memory suggestibility to leading questions and negative feedback, as assessed by the Gudjonsson Suggestibility Scale (van Bergen et al., 2009). Moreover, similar work by Kuczek et al. (2021) found no meaningful correlation between memory distrust and participants' susceptibility to misinformation about an educational report they had heard. These mixed findings suggest that measuring participants' memory distrust does not always successfully predict the kinds of metacognitive outcomes that we might theoretically expect it to predict, such as the reliance upon external sources of information.

Measuring Memory Distrust

To appraise some of these discrepancies between theory and empirical data, it is useful to consider how we measure memory distrust. The most popular memory distrust measure in the academic literature is the Squire Subjective Memory Questionnaire (SSMQ), first developed by Squire et al. (1979) for studying subjective memory functioning among depressed patients who were undergoing electroconvulsive therapy treatment. The questionnaire's 18 items probe a variety of everyday memory successes and failures, such as how well respondents tend to recall people's faces and names, and to recall newly-learned information and distant events from childhood (to give one example, respondents rate the statement "*My ability to remember things that have happened more than a year ago is…*", with scale anchors ranging from -4, worse than ever before, to +4, better than ever before). From its distinctive research origins, van Bergen et al. (2008, 2009, 2010a, b) repurposed the SSMQ for their detailed research programme exploring trait and state memory distrust. Their research programme has been influential in establishing a firm empirical evidence base for the conditions under which memory distrust evolves, its relationships with other personality traits and individual differences, and its consequences in important real-world contexts and cognitive tasks. Likewise, this work has inspired other researchers to conduct further investigations using the SSMQ, which enrich and expand our understanding of metamemory beliefs (e.g., Kuczek et al., 2018; Saraiva et al., 2019, 2020; Sauerland et al., 2019; Zhang et al., 2022).

Yet despite its academic currency as the leading psychometric measure of memory distrust, the SSMQ arguably only measures one aspect of memory distrust. Namely, the SSMQ seems to only assess people's tendency to think it likely they would fail to retrieve events or information. We contend that a second important aspect of memory distrust also exists: people's tendency to doubt that their successful recollections of past events are reliable or accurate (Mazzoni et al., 2010; Otgaar et al., 2014; Scoboria et al., 2014). Put differently, people might distrust their memories because (1) they do not treat their failures to remember events as 'proof' the events never happened, or because (2) they do not treat their memories of events as 'proof' the events happened. Importantly, none of the SSMQ's 18 scale items tap into the latter form of memory distrust, which we might characterise as concerning beliefs about memory *commission errors* (misremembering or falsely remembering). Instead, the SSMQ focuses on beliefs about *omission errors* (failing to remember)..[1]

Which of these kinds of memory beliefs is implicated in real-world examples of memory distrust? Most of the false confession cases described by Gudjonsson and colleagues as instances of 'Memory Distrust Syndrome' implicate suspects' beliefs about omission errors. For example, these cases include suspects who confessed despite having no recollections of committing the crime, because their history of frequent alcohol or drug intoxication made it plausible they could have committed the crimes without remembering (Gudjonsson & MacKeith, 1982; Gudjonsson

[1] It is important to consider that *beliefs about* errors are not necessarily errors themselves. Indeed, people's beliefs about their memory failures will often be entirely justified. For instance, a forgetful person would be perfectly justified in believing that they are prone to omission errors.

et al., 2014). Yet some of Gudjonsson's case studies also implicate beliefs about commission errors, noting for example that distrust "may occur when suspects' memory of their alibi and innocence is undermined by police during interrogation" (Gudjonsson, 2017, p. 157). Similarly, other experts in legal psychology and suggestibility define memory distrust as being grounded at least partly in beliefs about commission errors (e.g., Kopelman, 2013). Zhang et al. (2022), for instance, recently connected memory distrust with the concept of nonbelieved memories, whereby people maintain a vivid experience of recollecting a past event yet cease believing that the event occurred at all, or believing that their recollection is accurate. And in a completely different literature where 'memory distrust' has received significant attention, researchers of Obsessive-Compulsive Disorder (OCD) use this term in ways that seem focused towards commission errors. Specifically, clinicians and researchers note that repeated checking among people with OCD tends to be associated with a metamemory problem, with the checker experiencing such low confidence in their memories of prior checks that they feel compelled to check again, which in turn causes further memory distrust due to the recollective similarity of prior episodes (Radomsky et al., 2006; van den Hout & Kindt, 2003).

From looking across these diverse applied literatures it is clear that the term 'memory distrust' is often used to reference types of metamemory beliefs that the SSMQ does not capture. Does this discrepancy matter? Certainly, there are good reasons to distinguish omission and commission errors in memory research in general, and we know that omission errors have different properties than do commission errors. Providing false suggestions, for example, can lead people to recall fictional events quite readily (i.e., memory commissions) whereas similar suggestions about an event's non-occurrence are typically less successful in removing that subjective experience of recollection (i.e., memory omissions; Clark et al., 2012; Li et al., 2020; Nash et al., 2015). Furthermore, the misinformation effect literature shows that whereas people sometimes accept post-event information because they failed to encode the original information, other times they accept post-event information despite correctly recalling the original information, because they treat the source of the post-event information as more reliable than their own memory (e.g., Blank, 1998; Hope et al., 2008). Distinguishing—and predicting—these two different causes of misinformation acceptance is important, especially given that successful recovery and reporting of the original information will depend on its initial encoding. More generally, people are typically better aware of memory's susceptibility to omission errors than to commission errors. As mentioned above, for example, many people believe they have a dreadful memory, and yet large proportions of the general public also believe that "once you have experienced an event and formed a memory of it, that memory does not change" (Simons & Chabris, 2011). Even when challenged with the proposition that memories can be unreliable, most people believe they are good at judging the reliability of their own memories (Magnussen et al., 2006).

In sum, people's beliefs about their susceptibility to memory omission errors and to commission errors are likely to be distinct, and there is currently no evidence that a measure which indexes only the former belief—such as the SSMQ—would suffice as a proxy for the latter. In the present research, we sought to develop a new measure to tap into the aspects of memory distrust that concern commission errors, and to test the reliability and validity of this measure (Studies 1A and B). We predicted that our new scale would be correlated with the SSMQ, and other known correlates of that scale, but would be psychometrically distinct. In Study 2, we built upon the recent work of Zhang et al. (2022), who found that trait memory distrust—assessed using the SSMQ—was associated with participants' subjective ratings of how often they experience false and nonbelieved memories. Specifically, in Study 2 we used Scoboria and Talarico's (2013) memory cueing task to appraise the extent to which people's scores on our new memory distrust measure and the SSMQ would be associated with the spontaneous reporting of nonbelieved memories, and with overall belief in the cued events' occurrence.

Study 1A

Method

Scale development. To generate a pool of scale items that would tap into the commission aspects of memory distrust, we first reviewed published research literature on this topic to identify relevant themes in how the construct has been theorised and conceptualised. Key themes included self-awareness of difficulties in reality monitoring (e.g., Gudjonsson, 2017; Kopelman, 2013), a tendency to seek corroboration of one's recollections (e.g., Zhang et al., 2022), a tendency to trust external sources over one's own recollections (e.g., Gudjonsson, 2003), and a global belief that one's own recollections fail to match the objective truth (e.g., Henkel, 2014). We then constructed 20 items that mapped onto these themes, as listed in Table 1. Studies 1A and B and Study 2 served to assess the item structure, reliability, and validity of the resulting scale.

Table 1. Item-total correlations (ITC), communalities (h^2), and pattern matrix for the Memory Distrust Scale items.

Scale items	ITC	h^2	Factor Loading
1. I often look for physical evidence, such as photographs, to check whether things really happened the way I remember them	.67	.41	.64
2. I often turn to other people to help me decide whether my memories are accurate	.74	.51	.71
3. I tend to question my memories of past events if other people do not corroborate what I remember	.76	.54	.73
4. Sometimes I distrust my own memories if I cannot find any physical evidence to confirm what I remember	.81	.65	.80
5. I often have difficulty distinguishing events I remember from those I only imagined	.79	.61	.78
6. I'm often unsure whether something that I recall genuinely happened, or whether I only thought or dreamed about it	.81	.64	.80
7. I believe some of my memories may have originated entirely from my imagination	.76	.55	.74
8. 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	.74	.52	.72
9. Other people sometimes describe past events in ways that make me doubt my own recollection of those events	.81	.63	.70
10. I could be easily persuaded that an event I remember is impossible	.76	.55	.74
11. If another person contradicts my recollection of the past, they are probably correct	.75	.54	.74
12. Under the right circumstances, I could be persuaded that any one of my memories was completely false	.71	.47	.69

13. I generally have more trust in other people's recollections of events than in my own recollections	.78	.58	.76
14. I often trust other people's descriptions of a past event, even if I have a very different recollection of what happened	.69	.46	.68
15. Other people's memories are usually more accurate than my own memories	.69	.44	.66
16. My memories are rarely a very accurate reflection of what truly occurred	.69	.45	.67
17. My memories of past events are unreliable	.75	.55	.74
18. I cannot always be confident that my memories accurately reflect what really happened	.80	.62	.79
19. I have little trust that many of the events I remember did really occur	.74	.53	.73
20. I sometimes distrust that certain experiences I remember really happened at all	.77	.58	.76

Response scale: 1 – Strongly Disagree; 2 – Disagree; 3 – Slightly Disagree; 4 – Neither Agree nor Disagree; 5 – Slightly Agree; 6 – Agree; 7 – Strongly Agree.

Participants. We pre-registered our procedure and analytic plan at https://osf.io/7bz4u/?view_only=08f29688e5304bad8874c006f99252f6. Our target sample size was 400 participants following exclusions, based on a conservative respondent-to-item ratio of 20:1 for exploratory factor analysis (MacCallum et al., 1999). This sample size also provides sufficient power for the tests of convergent and divergent validity (power of .98 to detect correlations of r = .2 with alpha = .05, two-tailed). We recruited participants via two online participation panels: Qualtrics and Prolific. In Qualtrics, eligible participants were required to be nationals of the United Kingdom; in Prolific, participants were required to both be nationals of the United Kingdom, and to have a minimum Prolific approval rating of 95%.

In the Qualtrics dataset, an unknown number of participants who failed the inclusion criteria were automatically removed and replaced by the panel provider prior to the research team receiving the data; in the Prolific dataset, 27 participants who failed one or more inclusion criteria and were removed from the dataset prior to analysis. Our analysis is based on a total of 403 participants who passed all inclusion criteria, and who took part in exchange for monetary payment (Qualtrics n = 261; Prolific n = 142). Of these participants, 61% described themselves as female, 39% as male. Participants ranged in age from 18-79 years (*M* = 42 years, *SD* = 15), and participants in the Qualtrics subsample were significantly older (*M* = 45 years) than those in the Prolific subsample (*M* = 36 years; $t = 5.90$, $p < .001$). All participants were told that the study explored how people perceive their own memory functioning.

Materials.

Memory Distrust Scale (MDS). We presented the 20 MDS scale items to participants each in a random order, and asked "To what extent do you agree that each of the following statements is characteristic of you?". Each item was rated on a Likert scale ranging from 1 - *Strongly disagree* to 7 - *Strongly agree*, and for each participant a total score could range from 20 to 140, where higher scores indicated greater levels of memory distrust. We embedded three attention checks among these 20 items, each of which required participants to provide a specific answer (e.g., "Please respond to this question by selecting '*slightly agree*'"). Participants who failed any of these attention checks were excluded from analysis.

Squire Subjective Memory Questionnaire (SSMQ; Squire et al., 1979; van Bergen et al., 2010). As noted earlier, the SSMQ was originally intended as a measure of subjective memory functioning, and has been used frequently in prior research for assessing memory distrust (van Bergen et al., 2010). The scale comprises 18 items such as "My ability to remember the names and faces of people I meet is…", each of which is rated on a 9-point scale. We used the version of the SSMQ by van Bergen et al. (2010) wherein the anchors on the scale end-points were adapted to -4 - *Disastrous*, and +4 - *Perfect*. Total scores therefore can range from -72 to +72, where more-*negative* scores indicate greater levels of memory distrust. We should therefore expect the MDS and SSMQ to be negatively correlated. In the present sample, internal consistency was very high ($\alpha = .93$).

Cognitive Failures Questionnaire (CFQ; Broadbent et al., 1982). The CFQ assesses people's perceptions of the frequency with which they have experienced everyday failures and lapses in their attention and memory during the past six months. An example item is "Do you find you forget why you went from one part of the house to the other?". Participants rate each of 25 items using a 5-point scale from 1 - *Never* to 5 - *Very often*. After data collection we realised that one scale item (Item 9: "Do you fail to hear people speaking to you when you are doing something else?") had been unintentionally omitted for those participants recruited via Qualtrics. For all participants we therefore calculated total CFQ scores without this item. Total CFQ scores were based on the other 24 items, and could therefore range from 24 to 120, with higher scores indicating more perceived cognitive failures. In the present sample, internal consistency of the scale was very high (α = .93).

Gudjonsson Compliance Scale (GCS; Gudjonsson, 1989). The GCS assesses people's judgments of their susceptibility to social compliance, and includes 20 items, each of which is answered 'true' or 'false'. An example item is "I give in easily to people when I am pressured"; three of the items are reverse-scored to give a total score ranging from 0 to 20, where higher scores indicate greater compliance. In the present sample, internal consistency of the scale was high (α = .82).

Procedure. The survey was completed online at participants' own pace, with no time limit. After giving informed consent to take part, all participants reported their age and sex, before completing the MDS. Next, participants were shown a shape and single-digit number on screen, which they were asked to commit to memory; this served as an additional form of attention check to supplement the checks embedded in the MDS. Participants then completed the SSMQ, CFQ, and GCS one by one, with these three measures presented in a randomly counterbalanced order, and the order of items was also randomized within each scale. For the MDS, SSMQ, and CFQ, responses were selected using horizontal Likert scales; for the GCS participants simply clicked 'Yes' or 'No' for each scale item. Next participants completed a multiple-choice test of their memory of the shape and number shown previously (those who answered either of these latter questions incorrectly were excluded from analysis). Lastly, they read some debriefing information.

Results

The data for Study 1A and 1B can be accessed at https://osf.io/r8vqu/?view_only=c047e6891c0046c295035df9fd799202, along with the scripts for the main psychometric analysis. We did not preregister any plans to compare participants by recruitment method; nonetheless, participants recruited via Prolific vs. via Qualtrics responded comparably on the MDS (p = .70), SSMQ (p = .65), and GCS (p = .91). Prolific participants, however, scored significantly higher on the CFQ than did Qualtrics participants (p = .01, d = .27).

Exploratory factor analysis (EFA). Prior to the EFA, we screened a correlation matrix of the 20 items to identify items that were poorly correlated with the others, or items that were highly correlated and generating multicollinearity issues. This step of the analysis focused on item-to-item correlations, not to be confused with Item Total Correlations as shown in Table 1 (i.e., items' correlation with the total scale score without that item included). All items had correlations larger than r = .30 with other individual items, and only items 11 and 13 showed a particularly high Pearson correlation with each other (r = .71) so no items were removed.

Diagnostic tests were performed on the 20 items to examine the assumptions for EFA. Histograms and significant Shapiro Wilk tests for all the items indicated significant univariate non-normality, with skewness ranging from -0.01 to +0.93, and kurtosis ranging from −1.17 to +0.29. This observation was supported by the statistically significant Mardia's test, indicating that the assumption of multivariate normality was violated. Therefore, a weighted least squares extraction method for EFA was used, which provides standard errors and tests of model fit that are robust to the non-normality of the data. The items showed good factorability (Kaiser-Meyer-Olkin test = 0.97 and significant Bartlett's test) and did not present multicollinearity or singularity issues (Determinant > 0.00001).

We used parallel analysis and scree plots as factor retention criteria, which suggested the presence of one or two factors (see Supplemental Materials). A two-factor solution was extracted using oblimin rotation to allow for correlations between the factors. This solution revealed two distinguishable factors, but one of these factors had only four emerging items (items 11, 13, 14 and 15 listed in Table 1), all of which were related to memory distrust relative to other people (e.g., "If another person contradicts my recollection of the past, they are probably correct"). With the goal of achieving higher parsimony, we proceeded with the extraction of a one-factor solution, the pattern matrix for this solution is presented in Table 1. All items had factor loadings higher than 0.6 on this single MDS factor, which explained 54% of the variance in item responses. Cronbach's alpha indicated that the internal consistency of the MDS factor was very high ($\alpha = .96$). Given that assumptions for alpha are rarely met in psychometric research, we also examined the reliability of this MDS factor using omega coefficients (Dunn et al., 2014), which again showed good reliability [$\omega = .96$, 95% CI (.95, .96)]. Participants' total MDS scores averaged near the midpoint of the possible range of scores, and varied widely between participants ($M = 60.63$, Range = 20-129, $SD = 23.56$).

Convergent validity. In addition, the MDS scores correlated meaningfully with those from the other assessed measures. As Table 2 shows, and as we predicted, MDS scores were negatively associated with subjective memory functioning (as indexed by the SSMQ), and positively associated with perceived cognitive failures (indexed by the CFQ) and susceptibility to compliance (indexed by the GCS). The statistical significance of these associations held both for the overall sample as a whole, and also when analyzing the data separately for each of the two recruitment methods. Although all of these correlations were significant and somewhat stronger than we had predicted, the moderate strength of these associations suggests that the MDS measures a related but distinct construct.

We did not pre-register any predictions about how MDS scores would relate to participants' age; however, it might seem intuitively sensible to expect that older participants would report greater memory distrust. This was not the case: in fact there was a significant association in the opposite direction, $r(401) = -.26$, $p < .001$. Although this finding might initially seem to be a concern for the validity of the MDS, we found that SSMQ and CFQ scores were also correlated with age in the non-intuitive direction, $r(401) = .11$, $p = .03$, and $r(401) = -.25$, $p < .001$, respectively. These results diverge from those reported by van Bergen et al. (2010), who found a clear negative association between SSMQ and age, yet they are consistent with those of Kuczek et al. (2018), who reported a positive SSMQ—Age correlation of a similar magnitude to that found in our data ($r = .10$ in their study). In short, whereas the direction of these small correlations with age are intriguing, they do not point toward any specific issue of validity with the MDS relative to the SSMQ.

Table 2. Means, standard deviations, and Pearson correlations with confidence intervals of the MDS scores and other memory questionnaires (Study 1A)

Variable	*M*	*SD*	1	2	3
1. MDS	60.63	23.56			
2. SSMQ	22.43	20.97	-.46** [-.53, -.38]		
3. CFQ	39.01	15.15	.53** [.45, .59]	-.57** [-.63, -.50]	
4. GCS	11.34	4.33	.32** [.23, .41]	-.30** [-.38, -.21]	.40** [.31, .48]

Note. MDS = Memory Distrust Scale; SSMQ = Squire Subjective Memory Questionnaire; CFQ = Cognitive Failures Questionnaire; GCS = Gudjonsson Compliance Scale
Note. Values in square brackets indicate the 95% confidence interval for each Pearson correlation coefficient. ** indicates $p < .01$.

Study 1B

Following completion of Study 1A, we conducted two further rounds of data collection to assess the test-retest reliability of participants' MDS scores.

Method

Participants and procedure. Study 1B was not pre-registered. All 142 participants who took part in Study 1A via Prolific (but not those who took part via Qualtrics) at Time 1, and who passed the inclusion criteria, were invited to a second online study 3 weeks afterwards, in exchange for a further cash incentive. This Time 2 session was left open for one week. In the Time 2 session, participants completed the MDS only, with the items presented in a random order along with the three embedded attention check items. In total, 127 participants completed the Time 2 session within the one-week window, of whom 111 passed all the attention checks (M = 37 years, range 18-68, SD = 12 years; delay since participating in Time 1, M = 22 days, range = 21-27 days). Note that those participants who were invited and returned to take part in Study 1B did not differ significantly, in terms of any of the scale measures, from those participants who were not invited or did not return (MDS, p = .31; SSMQ, p = .84; CFQ, p = .12; GCS, p = .79).

Approximately 19 months after completion of Study 1A, these 111 people were invited again to complete an identical survey in exchange for a further cash incentive; this Time 3 session was left open for one week. In total, 57 participants completed this Time 3 session, of whom 55 passed all the attention checks (Mean age of these participants at Time 1 = 39 years, range = 21-68, SD = 12 years; delay since participating in Study 1A at Time 1, M = 586 days, range = 585-591 days).

Results

Participants' total MDS scores at Time 1 were strongly correlated with their MDS scores from 3-4 weeks later at Time 2, $r(109) = .83, p < .001$. Even after a delay of over 19 months, participants' total MDS scores at Time 3 were strongly correlated with their scores from Time 1, albeit this test-retest reliability was rather weaker than at Time 2, primarily due to a single outlier, $r(53) = .66, p < .001$ [test-retest reliability increased to $r(52) = .75, p < .001$ if this outlier was removed].

MDS scores in these follow-up samples again varied widely across participants, ranging from 20 to 120 (Time 2, M = 60.60, SD = 23.44; Time 3, M = 53.40, SD = 19.64), and the internal reliability of the MDS was excellent at both Time 2 and Time 3 (α = .96 and ω = .96 at Time 2; α = .95 and ω = .95 at Time 3).

Study 2

Based on the findings of Study 1, we were optimistic about the reliability and validity of our Memory Distrust Scale. In Study 2 we therefore set out to validate the single-factor structure of the MDS and to replicate its association with the SSMQ. In addition, we set out to trial a further test of construct validity for these memory distrust measures. Specifically, and like Zhang et al. (2022) we predicted that someone who distrusts their memories should be especially susceptible to reporting nonbelieved memories. Whereas Zhang et al. found some support for this prediction using the SSMQ, they relied upon people's subjective self-assessment of their frequency of experiencing nonbelieved memories. In Study 2, we used an arguably stronger method for assessing the occurrence of nonbelieved memories: the memory cueing method devised by Scoboria and Talarico (2013). Using this approach, participants reported a series of events from their childhoods, and then rated both their strength of belief in and recollection of those events. We predicted that participants who score more highly on the MDS would report lower ratings of belief in occurrence of their cued events, and as a consequence, we predicted that they would be more likely to report one or more nonbelieved memories. The MDS is designed to map closely onto beliefs about commission errors rather than omission errors. Therefore, we also predicted that the MDS would be associated with these belief-in-occurrence related outcomes more strongly than would the SSMQ.

Method

Participants. We pre-registered our procedure and analytic plan at https://osf.io/9gpbm/?view_only=3b5536d4f4f94cdba98000532e1a1c80 . Our target sample size was 400 participants following exclusions. As described below, Study 2 comprised two phases of data collection. A total of 662 participants ultimately completed Phase 1, of whom 268 were excluded from the dataset either because they did not return for Phase 2, or because they failed other inclusion criteria from Parts 1 and 2 of the study. Following these exclusions, 394 participants who passed all inclusion criteria were ultimately recruited (a) from the participant pools of two universities' undergraduate Psychology programmes (*n* = 206, who took part in exchange for a course credit), and (b) via Prolific (*n* = 188, who took part in exchange for a small cash incentive).[2] In Prolific, participants were required to both be nationals of the United Kingdom, and to have a minimum Prolific approval rating of 95%, and none of the participants from Study 1 were permitted to take part in Study 2.

Of our full sample of participants, 73.6% described themselves as female, 25.6% as male, 0.5% specified another gender identity, and 0.3% opted not to answer this question. By virtue of our primary reliance on student respondents in this study, participants were less diverse in age than in Study 1 (M = 22 years, range = 18-65, SD = 6 years), but the Prolific subsample was significantly older (M = 25 years) than the university student subsample (M = 19 years; t = 10.4, p < .001). All participants were told that the study explored how people perceive their own memory functioning.

Materials and Procedure.

Phase 1. In the first phase of the study, participants first provided demographic information and then saw the shape-and-number attention check used in Study 1A, which they were asked to remember for later. To begin the main part of this phase, participants were given the following instructions, adapted from Scoboria and Talarico (2013):

> *"Next you will be asked to think of a number of events from your childhood before the age of 11. By "Event" we mean any story or scene that occurred at a particular time and in a specific place. Events usually last for minutes or hours and no more than a single day."*

Participants were given eight blank text boxes, and were asked to type a sentence into each box to briefly describe eight different childhood events that matched these simple criteria.

[2] We followed our pre-registered approach to data collection, but because our sample after a second wave of online data collection was only six participants short of the target sample size, we opted not to commence a further wave.

They were not given any specific event cues or prompts; rather, they were instructed to simply choose the first events that came to mind, whether or not those events were personally significant or had been previously discussed with other people.

Once participants had described these events, they were presented at random with one of their event descriptions, and were asked four questions about that event, which appeared in a random order. Two questions assessed *belief in occurrence*, namely "How likely is it that you personally did in fact experience this event?" (1 - *Definitely did not happen*; 8 - *Definitely happened*), and "How strong is your belief that this event actually occurred (whether or not you remember it)?" (1 - *No belief*; 8 - *Strong belief*). The other two questions assessed *recollection*, namely "Do you actually remember experiencing this event?" (1 - *No memory of event at all*; 8 - *Clear and complete memory of event*), and "How strong is your memory for this event (whether or not you believe it occurred)?" (1 - *No memory*; 8 - *Strong memory*). After rating each of these four questions, participants were shown another random exemplar of their self-reported events, and this process was repeated until they had rated all eight events. Finally, they completed a multiple-choice test for their memory of the shape and number shown at the start of the study (those who misremembered both details were excluded from the dataset), and they read some interim debriefing information that referred generally to our interest in memory beliefs, but made no specific reference to memory distrust.

Phase 2. Two weeks after completing Phase 1, participants were invited to a second study. Astute participants might have noticed and recalled that the researcher's name was the same as for Phase 1, but we did not otherwise make any explicit connection with the study that participants had completed beforehand. After consenting to take part, participants simply completed the MDS (with the same three embedded attention check items as were used in Study 1), followed by the SSMQ, and were then debriefed.

Coding of event data. The lead author coded each of the eight autobiographical events described by each participant based on our pre-registered criteria. Specifically, responses were coded as valid if they described an event that would have ostensibly occurred on a single specific day, before the participant was aged 11, and did not repeat an event already described by the same participant. For example, when a participant wrote "Going to my aunt and uncle's wedding when I was 2," this was coded as a valid response. But when participants wrote "Waiting for my dad to get home from work," "My grandparents used to have a dog," or simply "Birthday," these were coded as invalid responses because they apparently concerned repeated or extended events, or because it was unclear they were referring to a personal event at all.

The second author also coded the data from a random sample of 10% of those participants who completed both Phase 1 and Phase 2 (i.e. n = 47 of the 467 participants who completed both phases, x 8 responses = 376 judgments), and he agreed with the first author for 88% of the total judgments (Gwet's AC1 = .83). Therefore, the first author's judgments were considered reliable and used as the basis of all analyses.

Per our pre-registration, all participants who were judged to have reported fewer than five valid events were removed from further analyses, along with those who failed other attention checks. For all other 394 participants, we calculated our key dependent variables based on participants' ratings of their valid events only. That is, we excluded the rating data for any individual events/responses that were not deemed valid. For each participant we then calculated (1) a mean *belief in occurrence* score, calculated from the mean ratings of the two belief in occurrence items, combined across all valid events; and (2) a mean *recollection* score, calculated from the mean rating of the two recollection items, combined across all valid events. We also calculated (3) a mean *belief discrepancy* score, calculated as the mean belief in occurrence score minus the mean recollection score. Finally, we calculated (4) the proportion of the total valid events reported by the participant that could be classified as 'nonbelieved memories'. Given that many self-reported nonbelieved memories are rated only fractionally lower on belief in occurrence than in recollection (Scoboria et al., 2017), here we defined nonbelieved memories simply as events for which the mean recollection rating exceeded the mean belief in occurrence rating.

Results

The data for Study 2 can be accessed at https://osf.io/r8vqu/?view_only=c047e6891c0046c295035df9fd799202. We did not preregister any plans to compare participants by recruitment method; nonetheless, participants recruited via Prolific vs. via university participant pools behaved comparably on the SSMQ (p = .47), and in all dependent measures from the autobiographical events task (all ps > .14). However, Prolific participants scored significantly lower on the MDS (M = 59.0, SD = 20.1) than did those from university participant pools (M = 66.8, SD = 19.3), p < .001. We return to further consider this interesting insight from exploratory analysis below.

Confirmatory factor analysis. The purpose of the subsequent analysis was to confirm the factor structure for the 20-item MDS scale. The results from the EFA in Study 1A indicated that a one-factor solution was appropriate to describe the MDS, so we conducted a confirmatory factor analysis to test data fit to this model based on the Study 2 data. Goodness of fit was evaluated using the robust root mean square error of approximation (RMSEA) and its 90% confidence interval, robust comparative fit index (CFI), and robust Tucker-Lewis index (TLI). These fit indices provide different types of information (i.e., absolute fit, fit adjusting for model parsimony, fit relative to a null model), and when combined they provide a reliable and conservative evaluation of model fit (Schreiber et al., 2006). The chi-square test is reported, but not relied upon to evaluate model fit due to its oversensitivity to sample size and the fact that it tests for perfect fit. We assessed the evaluation of the model based on conventional criteria for acceptable model fit (RMSEA < .08, CFI > .90, TLI > .90)

We used diagnostic tests on the 20 items to examine the assumptions for CFA and indicated that the assumption of multivariate normality was violated. Therefore, the parameters in CFA were estimated using a maximum likelihood estimation with robust standard errors, which provides tests of model fit that are robust to the non-normality of the data (Finney & DiStefano, 2006). Figure 1 presents the model specification and goodness of fit indices for the one-factor model. Although the CFI and TLI goodness of fit measures were fractionally below conventional criteria for acceptable fit, we determined that overall, the one-factor solution achieved an acceptable fit to the data.

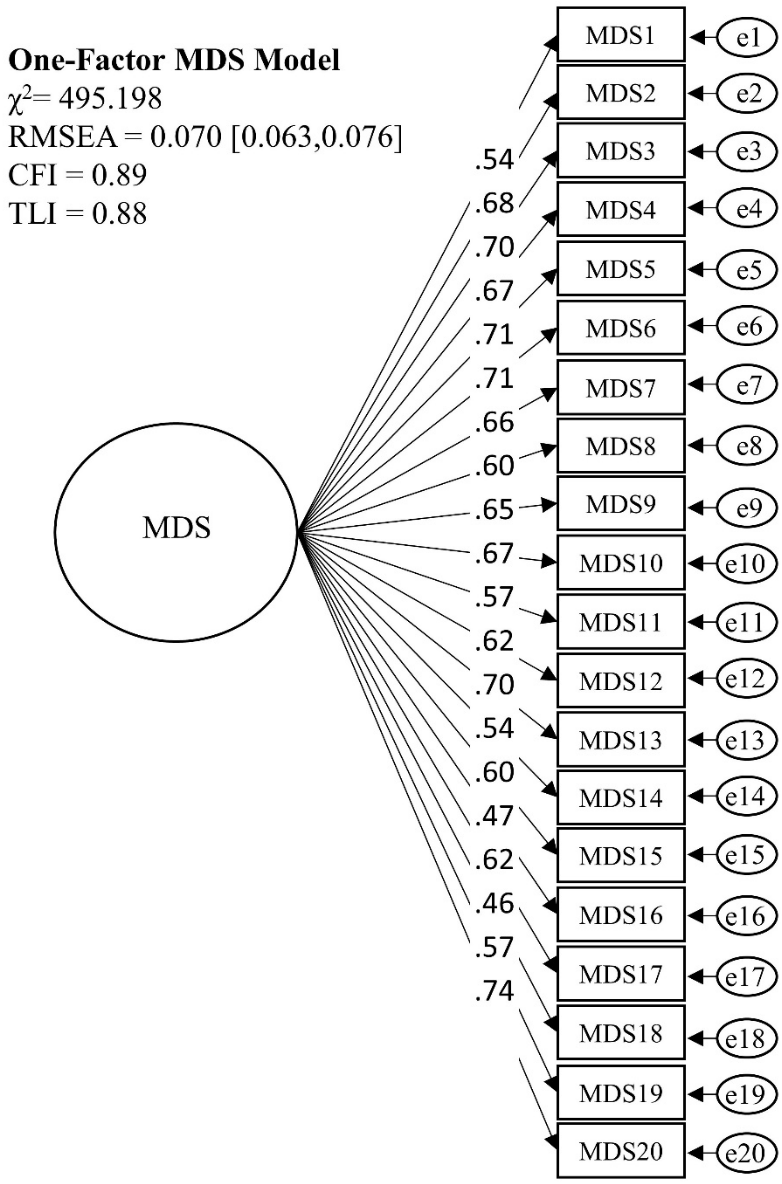

Figure 1. *Model specification and model fit indices for the one-factor solution of the Memory Distrust Scale (MDS). CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation.*

Descriptive statistics and convergent validity. As in Study 1A, participants' total MDS scores averaged near the midpoint of the possible range of scores, and varied widely between participants (*M* = 63.06, Range = 20-140, *SD* = 20.05). Ratings of the individual MDS items showed very high internal consistency in this sample, α = .93, ω = .93. Importantly, in line with Study 1A and our predictions, the Pearson correlation between total scores on the MDS and the SSMQ was moderate and negative, $r\,(392) = -.49$, $p < .001$ – a significant correlation that also held when analyzing the data separately for each of the two recruitment methods.

We registered no predictions about associations with participants' age, but exploratory analysis revealed the same counterintuitive patterns of findings that we observed in Study 1A; namely that age was associated negatively with MDS scores, $r\,(392) = -.27$, $p < .001$, and positively with SSMQ scores, $r\,(392) = .09$, $p = .09$. Recall that unlike Study 1, this study involved both a sample recruited from university participant pools, as well as a sample recruited online, and that the student sample scored significantly higher on the MDS. It is noteworthy that the student sample was also significantly younger than the Prolific sample, and so we wanted to rule out the possibility that the student respondents' data could have distorted these findings. That is to say, it is plausible that the negative correlation between MDS scores and age in Study 2 was an artefact of the younger participants being better educated about the fallibility of memory, as a consequence of them having studied psychology. This explanation is appealing, but it gained minimal support from the data. In fact, even when we focused our analyses solely on the Prolific subsample—that is, excluding all of the student participants from our analyses—the associations of age with MDS scores ($r = -.28$) and SSMQ scores ($r = .13$) remained comparable to the full-sample data.

Autobiographical event task performance. On average, participants' belief in the occurrence of the events they reported was extremely high (*M* = 7.49 out of 8, *SD* = 0.63; Median = 7.68); as were their recollection ratings to a lesser extent (*M* = 6.48 out of 8, *SD* = 1.06; Median = 6.61). Whereas these high ratings should be unsurprising, a consequence was that these variables were skewed from normality, especially the belief in recollection ratings, which were heavily skewed. We decided to proceed with our pre-registered parametric analyses irrespective of this issue, but in Table 3 we also report Spearman correlations between all variables, and in drawing conclusions we refer to these nonparametric statistics alongside those obtained from the pre-registered analyses.

Our first regression model predicted participants' mean belief in occurrence scores from their MDS and SSMQ scores. The overall model fit was significant, albeit the effect sizes were very small by conventional standards, $F(2, 391) = 9.15$, $p < .001$, $R^2 = .04$. We had predicted that MDS scores would be a better predictor than would SSMQ scores, and this indeed appeared to be the case: when both scores were entered in parallel into the model, the MDS emerged as a significant predictor, $\beta = -.19$, $t = 3.31$, $p = .001$, whereas the SSMQ did not, $\beta = .04$, $t = 0.76$, $p = .45$. Likewise, in our nonparametric analyses the MDS scores were a stronger predictor of belief in occurrence than were SSMQ scores (see Table 3).

The second regression model predicted participants' mean recollections scores, again using MDS and SSMQ scores as predictors. The overall model fit was significant, $F(2, 391) = 21.31$, $p < .001$, $R^2 = .10$. We had predicted that MDS scores and SSMQ scores would be associated with recollection similarly, and again this prediction was supported: when both scores were entered in parallel into the model, both MDS and SSMQ scores emerged as significant predictors of a similar magnitude (MDS, $\beta = -.16$, $t = 2.92$, $p < .01$; SSMQ, $\beta = .20$, $t = 3.68$, $p < .001$. In the nonparametric tests, both MDS and SSMQ scores were associated with recollection to a similar extent.

The third regression model predicted participants' mean belief discrepancy scores from their MDS and SSMQ scores. After data collection, we concluded that this outcome variable was less informative than we had anticipated, because the belief in occurrence scores being so close to ceiling meant that these belief discrepancy scores largely corresponded to a noisier measure of recollection rather than being highly informative about belief in recollection. Nonetheless, because we preregistered this analysis we report it regardless. The overall model fit was significant, $F(2, 391) = 7.93$, $p < .001$, $R^2 = .04$. We had predicted that MDS scores would be a better predictor than would SSMQ scores, but this prediction was not supported: when both scores were entered in parallel into the model, the MDS was not a significant predictor, $\beta = 0.05$, $t = 0.93$, $p = .35$, whereas the SSMQ was a significant predictor, $\beta = -.17$, $t = 2.93$, $p < .01$.

Finally, we had predicted that the MDS, but not the SSMQ, would correlate significantly with the proportion of nonbelieved memories reported. In total, 5.0% of participants' self-selected events were nonbelieved memories, which is consistent with reports from prior studies in which participants responded to general event cues rather than being specifically asked to report nonbelieved events (e.g., Scoboria & Talarico, 2013). Our prediction about these proportions was technically supported, as the Spearman correlations in Table 3 were statistically significant for the MDS and not for the SSMQ. However, it is important to note that the magnitudes of both correlation coefficients were very similar, and we are therefore reluctant to draw inferences from the significance of one versus non-significance of the other. Per our pre-registration, we also grouped participants into those who reported one or more nonbelieved memory ($n = 94$), and those who did not ($n = 300$). Our analyses showed that people who reported a nonbelieved memory scored higher on the MDS ($M = 66.94$, $SD = 20.08$) than participants who did not ($M = 61.84$, $SD = 19.91$), $t(392) = 2.16$, $p = .03$, $d = .26$. The same was not true of the SSMQ, where the scores of the nonbelieved memory group ($M = 22.11$, $SD = 22.01$) did not differ significantly from those of the remaining participants ($M = 25.79$, $SD = 20.80$), $t(392) = 1.48$, $p = .14$, $d = .17$. As above, we note that the standardized effect sizes here were similar for the MDS and the SSMQ; the support for our specific predictions about the comparison of the MDS and SSMQ should therefore not be over-stated, but these results do nevertheless provide some evidence of construct validity for the MDS.

Table 3. Spearman correlation coefficients of the associations between participants' total scores on the scale measures and event cueing task measures used in Study 2.

Variable	*M*	*SD*	1	2	3	4	5
1. MDS	63.06	20.05					
2. SSMQ	24.91	21.13	-.48**				
3. Belief in occurrence	7.49	0.63	-.27**	.17**			
4. Recollection	6.48	1.06	-.24**	.26**	.48**		
5. Belief discrepancy	1.04	0.96	.13*	-.17**	.02	-.76**	
6. Proportion of nonbelieved memories	0.05	0.11	.10*	-.08	-.48**	-.01	-.37**

Note. MDS = Memory Distrust Scale; SSMQ = Squire Subjective Memory Questionnaire
* indicates $p < .05$, ** indicates $p < .01$.

Discussion

People are sometimes unsure whether their memories of past events are reliable. These feelings of memory distrust are particularly relevant in a wide variety of contexts where memory recollections are of practical value and where errors can have serious consequences, such as in forensic and clinical settings. Yet the leading measure used for assessing memory distrust (i.e., the SSMQ) does not clearly index this important form of distrust. In the present research, we aimed to develop and test the validity of a new memory distrust scale to fill this important gap. Psychometric analysis of the Memory Distrust Scale (MDS) showed that the 20 items load onto a single memory distrust factor. Two studies showed that the MDS has strong internal reliability and good test-retest reliability over a span from several weeks to more than one and a half years. Moreover, MDS scores correlated in predictable ways with other key measures (i.e., the SSMQ, CFQ, and GCS), showing good evidence of convergent validity.

We also found that the MDS was associated with theoretically relevant metacognitive variables, particularly people's strength of belief in the occurrence of past events, their strength of recollection of those events, and their resulting tendency to report nonbelieved memories.

Like in Scoboria and Talarico's (2013) studies, participants here reported relatively few nonbelieved memories. As per our predictions, MDS scores were significantly associated with the reporting of nonbelieved memories, whereas based on null hypothesis significance testing criteria alone, SSMQ scores were not. It would not be safe to conclude from these data that the MDS predicts nonbelieved memories better than does the SSMQ, because the two effect sizes were, in fact, very similar. Nevertheless, these data do provide the first behavioral evidence of a link between memory distrust and nonbelieved memories, complementing the self-report evidence described by Zhang et al. (2022). More generally, the data point to a trait tendency for some people to disbelieve their memories more than others, a trait that can be assessed using the MDS. Study 2, for example, suggests that the MDS was particularly informative in predicting people's appraisals of the truth-value of their memories. That is to say, whereas both the MDS and the SSMQ were similarly—and significantly—related to recollective experience, the MDS significantly outweighed the SSMQ in its ability to predict belief in occurrence. This outcome aligns with our preregistered predictions, but we might expect that the predictive performance of the MDS and SSMQ would vary for different types of autobiographical events. For example, most of our participants chose to describe events they remembered vividly, but we might speculate that the SSMQ would outperform the MDS when predicting people's belief in the occurrence of events they remember poorly. The reason for expecting this pattern is that in those circumstances, beliefs in occurrence should be informed by judgments of the plausibility of having forgotten the events, which relate more closely to the SSMQ.

Although it was not surprising, the fact that Study 2's participants overwhelmingly described events that they believed in firmly, and recalled well, complicates our findings because it meant there was relatively little variability in the belief in occurrence and recollection ratings. Consequently, the results from our regression analyses must be interpreted cautiously due to the skewed data distributions, although we note that the results from these analyses are in line with those of our nonparametric correlational analyses. Future research may benefit from using different memory tasks to avoid these statistical issues, but also to address further interesting questions that the MDS can shed light upon. We propose that such research should use the MDS and SSMQ in combination, to better account for the causes and consequences of certain memory errors and metacognitive judgments. In particular, it would be valuable to explore the extent to which the MDS—individually and alongside the SSMQ—can predict other outcomes of theoretical and applied importance, such as the eyewitness misinformation effect (e.g., van Bergen et al., 2010), and the abandoning of true and false memories following social influence (e.g., Clark et al., 2012; Mazzoni et al., 2014).

Whereas the effect sizes of the relationships between the MDS and metamemory judgments in Study 2 were small by common conventions, several were larger than is typical for individual differences research (e.g., Gignac & Szodorai, 2016, suggest that the median correlation coefficient in such research falls in the order of $r = .20$). Moreover, it is notable that these relationships were detected using a procedure in which (1) the psychometric and memory cueing data were collected a week apart, in ostensibly unrelated studies; and (2) people spontaneously reported whichever autobiographical events came to mind, rather than being prompted to recall events they were uncertain about, or to recall specific public events. We would argue that observing meaningful statistical relationships in these circumstances provides additional evidence in support for the trait stability and validity of memory distrust, at least within this one cultural context. That said, given the interest in memory distrust among researchers in the OCD literature, it is important to emphasize that we have not tested or validated the MDS in clinical samples, and it should not be treated as clinically informative at present.

One curious finding from our exploratory analyses was that in both Studies 1 and 2, increasing age was associated with lesser memory distrust, rather than with greater memory distrust as intuition and some (but not all) prior research would suggest. In Study 2, this pattern held even after we excluded our statistically younger sample of Psychology students from analyses, whose specialist education might have equipped them with greater awareness of the fallibility of memory, and who indeed scored higher on the MDS than did the statistically older Prolific sample. It would be of interest for future research to directly explore how different kinds of memory trust/distrust beliefs vary with age, using non-online sampling to examine generalizability of these findings, and taking into account educational background as a potential confound variable.

Finally, our findings clearly highlight how common it is for people to have doubts about the accuracy or reliability of their own memories. Indeed, around one in three participants scored above the midpoint of the MDS, indicating overall tendencies toward memory distrust (Study 1A, 34.3%; Study 2, 32.5%). In contrast, far fewer participants had negative SSMQ scores that would indicate overall memory distrust on that scale (Study 1A, 11.9%; Study 2, 13.2%). Although MDS and SSMQ scores were correlated, these overall results support our initial premise that beliefs about personal susceptibility to omission errors and to commission errors do not always go hand in hand. With the important real-world consequences of these different metamemory beliefs in mind, we must appreciate that a measure of one clearly does not suffice as a proxy for the other.

Data availability

Data can be downloaded from https://osf.io/r8vqu/?view_only=c047e6891c0046c295035df9fd799202

References

Blank, H. (1998). Memory states and memory tasks: An integrative framework for eyewitness memory and suggestibility. *Memory, 6*, 481-529. https://doi.org/10.1080/741943086

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21*, 1-16. https://doi.org/10.1111/j.2044-8260.1982.tb01421.x

Clark, A., Nash, R. A., Fincham, G., & Mazzoni, G. (2012). Creating non-believed memories for recent autobiographical events. *PLoS One, 7*, e32998. https://doi.org/10.1371/journal.pone.0032998

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105,* 399-412. https://doi.org/10.1111/bjop.12046

Finney, S., & DiStefano, C. (2006). Nonnormal and categorical data in structural equation modeling. In G. Hancock & R. Mueller (Eds.), *Structural equation modeling: A second course.* Information Age Publishing.

Ghetti, S. (2003). Memory for nonoccurrences: The role of metacognition. *Journal of Memory and Language, 48,* 722-739. https://doi.org/10.1016/S0749-596X(03)00005-6

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74-78.

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. John Wiley & Sons. https://doi.org/10.1002/9780470713297

Gudjonsson, G. H. (2017). Memory distrust syndrome, confabulation and false confession. *Cortex, 87,* 156-165. https://doi.org/10.1016/j.cortex.2016.06.013

Gudjonsson, G. H. (1989). Compliance in an interrogative situation: A new scale. *Personality and Individual Differences, 10*, 535-540. https://doi.org/10.1016/0191-8869(89)90035-4

Gudjonsson, G. H., & MacKeith, J. (1982). False confessions: Psychological effect of interrogation: A discussion paper. In A. Trankell (Ed.), *Reconstructing the past: The role of psychologists in criminal trials* (pp. 253-269). Kluwer.

Gudjonsson, G. H., Sigurdsson, J. F., Sigurdardottir, A. S., Steinthorsson, H., & Sigurdardottir, V. M. (2014). The role of memory distrust in cases of internalised false confession. *Applied Cognitive Psychology, 28*, 336-348. https://doi.org/10.1002/acp.3002

Henkel, L. A. (2014). Memory trust and distrust in elderly eyewitnesses: To what extent do older adults doubt their memories? In M. P. Toglia, D. F. Ross, J. Pozzulo, & E. Pica (Eds.), *The elderly eyewitness in court* (pp. 232-262). Psychology Press.

Hope, L., Ost, J., Gabbert, F., Healey, S., & Lenton, E. (2008). "With a little help from my friends…": The role of co-witness relationship in susceptibility to misinformation. *Acta Psychologica, 127*, 476-484.

Kopelman, M. (2013). Memory disorders in the law courts. *Medico-Legal Journal, 81*, 18-28. https://doi.org/10.1177/0025817213477010

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. https://doi.org/10.1016/j.paid.2016.05.021

Kuczek, M., Szpitalak, M., & Polczyk, R. (2021). Anxious and distrustful-How do state anxiety and memory distrust influence the misinformation effect? *Polish Psychological Bulletin, 52*, 341-348.

Li, C., Wang, J., & Otgaar, H. (2020). Creating nonbelieved memories for bizarre actions using an imagination inflation procedure. *Applied Cognitive Psychology, 34*, 1277-1286. https://doi.org/10.1002/acp.3722

Lorayne, H. (2000). *How to develop a super power memory*. Frederick Fell Publishers.

Magnussen, S., Andersson, J., Cornoldi, C., De Beni, R., Endestad, T., Goodman, G. S., […] & Zimmer, H. (2006). What people believe about memory. *Memory, 14*, 595-613. https://doi.org/10.1080/09658210600646716

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-99. https://doi.org/10.1037/1082-989X.4.1.84

Mazzoni, G., & Kirsch, I. (2002). Autobiographical memories and beliefs: A preliminary metacognitive model. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (pp.

121-146). Cambridge University Press. https://doi.org/10.1017/CBO9780511489976.007

Mazzoni, G., Clark, A., & Nash, R. A. (2014). Disowned recollections: Denying true experiences undermines belief in occurrence but not judgments of remembering. *Acta Psychologica, 145*, 139-146. https://doi.org/10.1016/j.actpsy.2013.11.007

Mazzoni, G., Scoboria, A., & Harvey, L. (2010). Nonbelieved memories. *Psychological Science, 21,* 1334-1340. https://doi.org/10.1177/0956797610379865

Nash, R. A., Wheeler, R. L., & Hope, L. (2015). On the persuadability of memory: Is changing people's memories no more than changing their minds? *British Journal of Psychology, 10*, 308-326. https://doi.org/10.1111/bjop.12074

Otgaar, H., Scoboria, A., & Mazzoni, G. (2014). On the existence and implications of nonbelieved memories. *Current Directions in Psychological Science, 23,* 349-354. https://doi.org/10.1177/0963721414542102

Radomsky, A. S., Gilchrist, P. T., & Dussault, D. (2006). Repeated checking really does cause memory distrust. *Behaviour Research and Therapy, 44*, 305-316. https://doi.org/10.1016/j.brat.2005.02.005

Saraiva, R. B., Hope, L., Horselenberg, R., Ost, J., Sauer, J. D., & van Koppen, P. J. (2020). Using metamemory measures and memory tests to estimate eyewitness free recall performance. *Memory, 28*, 94-106. https://doi.org/10.1080/09658211.2019.1688835

Saraiva, R. B., van Boeijen, I. M., Hope, L., Horselenberg, R., Sauerland, M., & van Koppen, P. J. (2019). Development and validation of the Eyewitness Metamemory Scale. *Applied Cognitive Psychology, 33,* 964-973. https://doi.org/10.1002/acp.3588

Sauerland, M., Krix, A. C., & Sagana, A. (2019). Deceiving suspects about their alibi is equally harmful to the innocent and guilty. *Applied Cognitive Psychology, 33*, 1238-1246. https://doi.org/10.1002/acp.3577

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research, 99*, 323-338. https://doi.org/10.3200/JOER.99.6.323-338

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*, 1242-1258. https://doi.org/10.1037/a0034110

Scoboria, A., Nash, R. A., & Mazzoni, G. (2017). Sub-types of nonbelieved memories reveal differential outcomes of challenges to memories. *Memory, 25*, 876-889.

Scoboria, A., & Talarico, J. M. (2013). Indirect cueing elicits distinct types of autobiographical event representations. *Consciousness and Cognition, 22*, 1495-1509. https://doi.org/10.1016/j.concog.2013.09.007

Simons, D. J., & Chabris, C. F. (2011). What people believe about how memory works: A representative survey of the US population. *PLoS One, 6*, e22757. https://doi.org/10.1371/journal.pone.0022757

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.

Storm, B. C. (2011). The benefit of forgetting in thinking and remembering. *Current Directions in Psychological Science, 20*, 291-295. https://doi.org/10.1177/0963721411418469

Strack, F., & Bless, H. (1994). Memory for nonoccurrences: Metacognitive and presuppositional strategies. *Journal of Memory and Language, 33*, 203-217. https://doi.org/10.1006/jmla.1994.1010

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*, 373-384. https://doi.org/10.1348/135532509X471960

Van Bergen, S., Horselenberg, R., Merckelbach, H., Jelicic, M., & Beckers, R. (2010). Memory distrust and acceptance of misinformation. *Applied Cognitive Psychology, 24*, 885-896. https://doi.org/10.1002/acp.1595

Van Bergen, S., Jelicic, M., & Merckelbach, H. (2008). Interrogation techniques and memory distrust. *Psychology, Crime & Law, 14*, 425-434. https://doi.org/10.1080/10683160701822533

Van Bergen, S., Jelicic, M., & Merckelbach, H. (2009). Are subjective memory problems related to suggestibility, compliance, false memories, and objective memory performance? *American Journal of Psychology, 122*, 249-257.

Van den Hout, M., & Kindt, M. (2003). Repeated checking causes memory distrust. *Behaviour Research and Therapy, 41*, 301-316. https://doi.org/10.1016/S0005-7967(02)00012-8

Zhang, Y., Otgaar, H., & Wang, J. (2022). Memory distrust is related to memory errors, self-esteem, and personality. *Applied Cognitive Psychology, 36*, 283-292. https://doi.org/10.1002/acp.3917