A Tale of Two Distrusts: Memory Distrust towards Commission and Omission Errors in the Chinese Context

Yikang Zhang[1, 2], Fangzhu Qi[1], Henry Otgaar[2, 3], Robert A. Nash[4], and Marko Jelicic[2]

[1] Institute of Psychology and Behavior, Henan University, China

[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

In Press, Journal of Applied Research in Memory and Cognition

Correspondence

Correspondence should be addressed to Fangzhu Qi, Institute of Psychology and Behavior, Henan University, Jinming Avenue, 475001, Kaifeng, Henan Province (fzqi@henu.edu.cn) or Yikang Zhang, Faculty of Psychology and Neuroscience, Maastricht University, Universiteitssingel 40, 6229 ER, Maastricht (yikang.zhang@maastrichtuniversity.nl).

Author Note.

Yikang Zhang played the lead role in the conceptualization, project administration, data curation, data analysis, and writing of the original draft. Fangzhu Qi played the lead role in funding acquisition and a supporting role in data curation, project administration, and the original draft writing. Henry Otgaar, Robert Nash, and Marko Jelicic played supporting roles in the conceptualization, methodology, original draft writing, and review/ editing.

Abstract

People differ in their skepticism toward their own memories, which is called memory distrust and is measured by the Squire Subjective Memory Questionnaire (SSMQ) and the Memory Distrust Scale (MDS). In Study 1 (N = 458), we translated the MDS into Chinese and found that MDS scores were correlated with self-reported memory errors, compliance, and life habits impacting source monitoring, and had acceptable test-retest reliability after four weeks. In Study 2, participants (N = 383) completed a recognition task and received false feedback, then they completed the recognition task again, and completed the MDS and SSMQ three days later. High (versus low) memory distrust people were more likely to accept the false feedback and change their memory afterward. The present research confirms the validity of the Chinese MDS, advancing the theoretical understanding of the interplay between meta-memorial beliefs and social influence on memory reconstruction.

Keywords: memory distrust, memory errors, compliance

General Audience Summary

The experience of sometimes finding it difficult to trust one's own memory is widely shared. Moreover, some people are more skeptical about their memories while others are less. This individual difference is referred to as trait memory distrust. Memory distrust has been measured with two scales, the Squire Subjective Memory Questionnaire (SSMQ) and the more recent Memory Distrust Scale (MDS). The former emphasizes people's concerns about forgetting one's previous experience while the latter asks how concerned people are about mistakenly remembering something that did not really happen.

In the present studies (N = 841), we translated the MDS into Chinese and proved it to be effective in measuring memory distrust. We found that people who score high on the MDS reported having more memory errors and being more compliant with authorities. We also found that people with high memory distrust as measured by the MDS were more likely than people with low memory distrust to accept false feedback, by changing their prior answers in a memory test.

In forensic settings where the completeness and veracity of memory reports are crucial, memory distrust could lead to severe consequences, such as wasting public resources if people do not report what they remember, and even miscarriages of justice. Our study provides the first preliminary evidence that memory distrust as measured by the MDS affects how people react to suggestive information, and could therefore be of interest when evaluating eyewitnesses' or suspects' statements. The validation of the Chinese MDS also invites clinical cognition research in the Chinese context, such as studies of the relationship between memory distrust and repeated checking. We encourage researchers to use the tool and further examine the role of

memory distrust in forensic and clinical research in China.

Introduction

The experience of sometimes finding it difficult to trust one's memory is widely shared (e.g., Kuczek et al., 2018; Nash et al., 2022; Otgaar et al., 2019; van Bergen et al., 2010a; Zhang et al., 2022a). Sometimes, as a result of suggestive questioning during police interrogations it can lead to egregious outcomes such as false confessions (i.e., Memory Distrust Syndrome; see Gudjonsson & MacKeith, 1982; Gudjonsson et al., 2014). Moreover, individuals differ in the extent to which they are more vs. less skeptical towards their own memories (Nash et al., 2022; van Bergen et al., 2010a). To measure individual differences in this meta-memorial appraisal, also conceptualized as trait memory distrust, van Bergen et al. (2010a) adapted and validated the Squire Subjective Memory Questionnaire (SSMQ, Squire et al., 1979), which has been employed widely by subsequent research on memory distrust (e.g., Kuczek et al., 2021; Saraiva et al., 2020; van Bergen et al., 2010b; Zhang et al., 2022a).

Recently, researchers argued that the SSMQ alone does not fully capture the construct of trait memory distrust (Nash et al., 2022). More specifically, the SSMQ's eighteen items (e.g., "My ability to remember things that have happened more than a year ago is") only tap into people's distrust insofar that they make memory omission errors, that is, failing to retrieve memories of experiences. However, people can sometimes also have distrust insofar that they make memory commission errors, such as mistaking imagination or dreams as reality (i.e., source monitoring, Johnson & Raye, 1981). People's beliefs about these types of errors are, however, not captured by the SSMQ. To address this issue, Nash et al. (2022) developed and validated the Memory Distrust Scale (MDS), a new measurement tool that focuses on people's

distrust toward commission errors (e.g.,"I am sometimes uncertain whether an event that I recall happened to me, or whether I saw it on TV or in a movie"). Nash et al. (2022) showed that the MDS and SSMQ were both correlated with other meta-cognitive measures such as the Cognitive Failures Questionnaire (CFQ, Broadbent et al., 1982) and that they were only moderately correlated with one another. Moreover, the authors demonstrated that compared with the SSMQ, MDS was a better predictor of people's ratings of autobiographical belief (i.e., their belief of specific autobiographical events having happened), with people who scored high on the MDS being more likely to report events with lower belief ratings, as compared with their counterparts who scored low on MDS. Nash et al. (2022) recommended that when examining the relationship between memory distrust and other memory phenomena (e.g., the misinformation effect, Loftus et al., 1992; van Bergen et al., 2010b), researchers should use both the SSMQ and the MDS in tandem.

Cultural Differences

Most research on memory distrust has been conducted in Western, Educated, Industrialized, Rich, and Democratic (WEIRD; Henrich et al., 2010) populations (e.g., Nash et al., 2022; van Bergen et al., 2009, 2010a, 2010b; Zhang et al., 2022a). Yet memory, like other psychological phenomena, is shaped by culture (e.g., Ross & Wang, 2010; Wang, 2021). More specifically, according to the cultural dynamic theory of autobiographical memory, culture influences the encoding, retention, and retrieval of our memories as well as the functions of memory sharing (Wang, 2016). As a consequence, cultural differences extend to the formation of false memories as well. For example, using the Deese-Roediger/McDermott (DRM) paradigm, Wang et al. (2021) showed that European participants formed more self-related false memories

than Chinese participants, possibly due to cultural differences in independent versus interdependent self-construal (Markus & Kitayama, 1991). Like false memories more generally, memory distrust has been suggested to be related to people's susceptibility to social influence (Zhang et al., 2022b), which has been found to be shaped by self-construal, with people who are more interdependent exhibiting greater compliance to others (Oeberst & Wu, 2015). Furthermore, it has been argued that judgments of mnemicity—that is, the attributions of mental representations as being memories—is a result of metacognitive and social construction processes that are influenced by collective norms (Mahr, 2023; Mahr et al., 2022). That is to say, the 'criteria' for what counts as memory may differ across cultures. In short, it may be unreasonable to attempt to apply theory and evidence on memory distrust across diverse cultures without undertaking empirical validation.

Beyond the theoretical merits of cross-cultural replications and translation, studying memory distrust in a Chinese context is also important for practical reasons. A recent review on forensic practice in Asia (Le et al., 2023) showed that false confessions and eyewitness identification errors—both plausible consequences of memory distrust—are important causes of wrongful convictions in Asian countries, just as they are in Western countries. Yet Asian countries lack scholarly work on issues pertaining to forensic interviewing and memory (but see Sumampouw et al., 2022), even in China and Japan where such research does exist but remains very limited. The same is true in other applied domains. In the clinical domain, for instance, the role of memory distrust in psychopathology in general (e.g., depression, Schweizer et al., 2018; distress, Mewton et al., 2014), and in obsessive-compulsive disorder in particular (e.g., Coles et al., 2006; Radomsky & Alcolado, 2010; Strauss et al., 2020), has been well studied in WEIRD contexts, yet similar research is only

emerging in other cultures such as China (Wong et al., 2022). Given the theoretical and practical importance of studying memory cross-culturally, and the sparseness of applied forensic and clinical research in China, the validation of a Chinese MDS stands to offer insights into how culture shapes the remembering processes, and to provide a large number of non-English-speaking researchers with access to the MDS as a research tool.

Zhang et al. (2022b) previously translated the SSMQ into Chinese, which showed good internal consistency and criterion validity. However, to our knowledge, the MDS has yet to be translated into Chinese and validated for relevant research. The present research represents the first effort to translate and validate the Chinese version of the MDS.

Theoretical and Empirical Correlates of Memory Distrust

According to the socio-cognitive model of memory proposed by Scoboria and Henkel (2020), when receiving negative social feedback that contradicts their recollections (e.g., being told that something did not happen), people weigh both the qualities of their internal representations and the qualities of the external feedback, which then determines whether or not they reduce/relinquish their autobiographical beliefs. Building on this model, Zhang et al. (2022b) argued that trait memory distrust could moderate this weighing process, with people who are more skeptical about their memory functioning placing less trust in their internal representations and greater trust in the feedback, as compared with their low memory distrust counterparts. Research has shown that people who score highly on memory distrust—assessed using the SSMQ—exhibit a greater misinformation effect (i.e., committing memory errors after receiving false information) compared with people who score low on

memory distrust (van Bergen et al., 2010a; although this effect was not found by Kuczek et al., 2021). Therefore, we expect that people who have high (versus low) memory distrust would be more likely to accept and be influenced by negative feedback that contradicts their memories. Further, as the SSMQ and MDS measure two distinct aspects of memory distrust, we expect that SSMQ scores would moderate people's likelihood of accepting negative feedback about making omission errors (i.e., when it is suggested that they have forgotten something) whereas MDS scores would moderate the likelihood of accepting negative feedback about making commission errors (i.e. when it is suggested that they have misremembered or falsely remembered something).

Previous research conducted in Western cultures showed a positive correlation between memory distrust and susceptibility to compliance (e.g., Nash et al., 2022; van Bergen et al., 2010a). As compliance is more prominent in an interdependent culture such as Chinese culture (Oeberst &Wu, 2015), we expect that both measures of memory distrust would likewise correlate with compliance in the Chinese population. Previous research also showed that memory distrust positively correlates with self-reported memory errors (e.g., Britain: Nash et al., 2022; China: Zhang et al., 2022b). We thus expected that both measures of memory distrust would also be related with self-reported memory errors.

The Present Research

In Studies 1a and 1b, we translated the MDS into Chinese and examined its internal consistency, test-retest reliability, and criterion validity. Moreover, we explored whether memory distrust was related to certain life habits (e.g., TV consumption) that in theory could influence source monitoring (e.g., distinguishing

memories from imagination). Taking into consideration that memory is broadly influenced by culture (e.g., Ross & Wang, 2010; Wang, 2021) and that compliance is stronger in cultures with interdependent self-construal (Oeberst &Wu, 2015), we used the data from Nash et al. (2022) to explore potential differences in the factor structure of the MDS in two different cultural contexts (i.e., Britain and China).

In Study 2, to test whether memory distrust would predispose individuals to accepting negative social feedback about their memories, we asked participants to complete an online memory task, gave them false feedback on some of their recognition responses, and then asked them to complete the recognition task again. Then we measured their memory distrust three days later to examine whether people who are high on memory distrust would be more likely to accept the false feedback, as compared with people who are low on memory distrust. Both studies were pre-registered (https://osf.io/m9skg and https://osf.io/gmye8).

Study 1a

Method

Ethical Approval

Studies 1a, 1b, and 2 were reviewed and approved by the Institutional Review Board of [Masked Institution] (reference: 20221110001).

Participants

We recruited participants from university participant pools and via social media. To participate in the study, participants had to read the information letter introducing the aim, tasks, and compensation scheme of the study, and provide informed consent. In addition, they also had to be 18 years old or older. Participants were compensated with 4 RMB[1] for the first session and an additional 4 RMB for completing the follow-up session.

Following Nash et al. (2022), we planned to recruit at least 400 participants after exclusions, based on a conservative respondent-to-item ratio of 20:1 for exploratory factor analysis (MacCallum et al., 1999). Using G*Power 3.1 (Faul et al., 2009), a sensitivity analysis for bivariate correlation showed that with $\alpha = .05$ and $1\text{-}\beta = .95$, a sample of 400 would be needed to reliably detect an effect no smaller than $\rho = .18$.

A total of 533 participants completed the first survey and, in accordance with our pre-registration, 74 were excluded for failing at least one attention check (i.e., participants did not select the required answer when responding to the attention checks). One additional participant was removed due to duplicated responses, leaving the final sample being 458 ($n_{\text{women}} = 254$, $n_{\text{men}} = 203$, $n_{\text{no disclosure}} = 1$). We did not use

[1] RMB is the official currency in the People's Republic of China

specific stopping rules and ended the data collection when the number of responses were close to the planned sample size and there were few new signups daily. The average age of the sample following exclusions was 22.55 years (*SD* = 3.53). Nearly all participants had an education level of a college degree or above (97.6%).

Materials

Squire Subjective Memory Questionnaire (SSMQ, Squire et al., 1979). The SSMQ as adapted by van Bergen et al. (2010a) is the most widely used measure of memory distrust. It comprises 18 items (e.g., "my ability to pay attention to what goes on around me is" from -4 = *Disastrous* to 4 = *Excellent*) that tap into one single underlying factor about one's subjective appraisal of one's memory functioning. The current study employed the Chinese version translated by Zhang et al. (2022b). The scale showed good internal reliability in the current sample (Cronbach's $\alpha = .92$; McDonald's $\omega = .94$). Note that a higher score on the SSMQ would indicate a lower level of memory distrust.

Memory Distrust Scale. To address the issue that the SSMQ's items focus only on concerns over omission errors (e.g., failing to recall past events), Nash et al. (2022) developed the Memory Distrust Scale (MDS), which emphasizes memory appraisals over commission errors. The MDS consists of 20 items in which participants rate to what extent the 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") are characteristic of themselves (1 = *Strongly disagree* to 7 = *Strongly agree*). Note that a higher score on the MDS would indicate a greater level of memory distrust. The questionnaire was translated into Chinese and translated back to English by two fluent English-Chinese speakers (Y.Z. and M.Z.) to ensure equivalent meaning of the items. There was no

significant difference in the meanings of the original and the back-translated version. Therefore, we used the Chinese version (see Table A1 in appendix) translated by Y.Z. in the current studies.

Gudjonsson Compliance Scale. The Gudjonsson Compliance Scale (GCS, Gudjonsson, 1989) assesses people's self-reported susceptibility to social compliance, comprising 20 items with dichotomous response options ("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 study, we used the Chinese version translated and validated by Oeberst and Wu (2015). The scale showed good internal reliability in the current sample (Cronbach's $\alpha = .80$; McDonald's $\omega = .83$).

Prospective and Retrospective Memory Questionnaire. The Prospective and Retrospective Memory Questionnaire (PRMQ) is a 16-item validated questionnaire measuring people's self-reported susceptibility to prospective and retrospective memory failures in daily life (Smith et al., 2000; Yang et al., 2022). Participants responded to PRMQ items (e.g., "Do you fail to recognize a place you have visited before?") on a 5-point Likert scale from 1= *Never* to 5 = *Often*. In the current study, we used the Chinese version that has been validated by Yang et al. (2022). The questionnaire showed good internal reliability in the current sample (Cronbach's $\alpha = .92$; McDonald's $\omega = .94$). Higher scores in the PRMQ correspond to self-reported higher frequencies of memory errors.

Procedure

The study was hosted on Qualtrics (https://www.qualtrics.com/). After reading the information letter and giving informed consent, participants answered

demographic questions including age, gender, and education. Next, they completed the GCS, SSMQ, MDS, and PRMQ, with both the order of these four measures and the order of the items within each measure being random. Subsequently, participants rated four statements assessing their sociometric status (Cronbach's $\alpha = .88$; McDonald's $\omega = .92$) using a 7-point scale (1 = Strongly Disagree, 7 = Strongly Agree); these were 1) I have a high level of respect in others' eyes, 2) Others admire me, 3) I have high social standing, and 4) Others look up to me (Anderson et al., 2012). Finally, participants left their email addresses for receiving the follow-up questionnaire.

One month after completing the first survey, participants received the link for the follow-up questionnaire, and completed only the MDS for a second time.

Data Analysis Overview

All data analyses were performed in R (version 4. 1.2, R Core Team, 2021). All anonymized datasets and coding scripts are on OSF (https://osf.io/p49yz/?view_only=b1f3cc0996d74822a7696041977a8da5). First, we conducted exploratory factor analyses (EFA) to examine the factor structure of the MDS. After establishing that the Chinese version of the MDS had good reliability and construct validity, we conducted correlational analyses to examine its convergent validity and test-retest validity. For exploratory purposes, we also used the data of Nash et al. (2022) to compare the factor structure and correlational patterns between their British sample and our Chinese sample, to examine potential cultural differences.

Results

Exploratory Factor Analysis

Item-item correlations of the Chinese version of the MDS were all statistically significant and ranged from r = .29 to .70, suggesting that there was no issue of poorly correlated items nor severe multi-collinearity. Item-total correlations ranged from r = .49 to .77, suggesting that all items had responses that varied in line with those for all other items, across the population of items. Therefore, no items needed to be removed.

Univariate normality tests (Anderson-Darling Test) showed that all 20 items of the MDS violated the normality assumption (ps < .001), with skewness ranging from -0.40 to 0.68 and kurtosis ranging from -1.29 to -0.50. A Henze-Zirkler test (HZ = 1.29, p <.001) and Mardia test of multivariate skew and kurtosis (skew = 3444.34, p < .001; kurtosis = 36.24, p < .001) also indicated multivariate non-normality. The Kaiser-Meyer-Olkin criterion (KMO = .97) and Bartlett's test of sphericity, χ^2 (190) = 6058.11, p < .001, showed that the data were suitable for factor analyses. Taking into consideration of the above-mentioned results, we proceeded with exploratory factor analyses (EFA) using Weighted Least Square estimator with robust standard errors (WLSMV), which is more appropriate for non-normal data than other approaches such as Maximum Likelihood (Sellbom & Tellegen, 2019).

We performed a series of tests (e.g., Empirical Kaiser criterion and Parallel analysis) to identify an appropriate number of retained factors (for details, see analysis output –factor selection), with a one-factor solution being recommended most often (5 out of 12 tests), followed by the three-factor solution (4 out of 12). First, a three-factor solution was extracted with oblimin rotation, which allows correlations between factors. The result showed that only Items 1 and 2 had a loading larger

than .30 on the third factor, which explained 3% of the variance. Therefore, the three-factor solution was deemed not practically meaningful.

The two-factor oblimin solution revealed two meaningful factors with a correlation of .74. Seventeen items had loadings greater than .30 on Factor 1, which explained 40.9% of the total variance. Nine items had loadings greater than .30 on Factor 2, explaining 15.2% of the variance (See Table A1 in the Appendix). A closer examination of the items suggested that Factor 1 taps more into the social aspects of memory distrust (e.g., "Other people's memories are usually more accurate than my own memories.") and Factor 2 taps more into source monitoring (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."). Notably, our result differs from Nash et al. (2022), whose analyses led them to prefer a one-factor solution, but whose initial examination of a two-factor solution indicated that the more social aspects of memory distrust loaded more onto the second, minor factor. This difference between samples may hint at cultural differences in memory distrust between the Chinese and the British samples; we return to this possibility shortly.

Finally, we extracted the one-factor solution. All items showed adequate loading (from .49 to .79) on the factor, which explained 52.5% of the variance. Taking into consideration the principle of parsimony, and that 5 out of 12 factor selection tests suggested the one-factor solution (similar to Nash et al., 2022), we decided to adopt the one-factor solution for subsequent analyses, the pattern matrix for which is presented in Table 1. The MDS showed great internal reliability (Cronbach's $\alpha = .96$; McDonald's $\omega = .96$).

Table 1

Item-Total Correlations (ITC), Communalities (h^2), and Pattern Matrix for the Memory Distrust Scale Item

	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.	.49	.24	.49
2	I often turn to other people to help me decide whether my memories are accurate.	.66	.44	.67
3	I tend to question my memories of past events if other people do not corroborate what I remember.	.76	.60	.78
4	Sometimes I distrust my own memories if I cannot find any physical evidence to confirm what I remember.	.72	.55	.74
5	I often have difficulty distinguishing events I remember from those I only imagined.	.77	.63	.79
6	I am often unsure whether something that I recall genuinely happened, or whether I only thought or dreamed about it.	.70	.52	.72
7	I believe some of my memories may have originated entirely from my imagination.	.66	.46	.68
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.	.65	.44	.66
9	Other people sometimes describe past events in ways that make me doubt my own recollection of those events.	.65	.43	.66
10	I could be easily persuaded that an event I remember is impossible.	.75	.59	.77
11	If another person contradicts my recollection of the past, they are probably correct.	.75	.59	.77
12	Under the right circumstances, I could be persuaded that any one of my memories was completely false.	.67	.48	.69
13	I generally have more trust in other people's recollections of events than in my own recollections.	.75	.60	.78
14	I often trust other people's descriptions of a past event, even if I have a very different recollection of what happened.	.74	.58	.76
15	Other people's memories are usually more accurate than my own memories.	.77	.63	.79
16	My memories are rarely a very accurate reflection of what truly occurred.	.72	.55	.74
17	My memories of past events are unreliable.	.71	.53	.73
18	I cannot always be confident that my memories accurately reflect what really happened.	.72	.55	.74
19	I have little trust that many of the events I remember did really occur.	.76	.62	.79
20	I sometimes distrust that certain experiences I remember really happened at all.	.69	.50	.71

Criterion Validity

As shown in Table 2, the MDS had statistically significant correlations with the PRMQ, GCS, and SSMQ. Moreover, the correlation between the MDS and SSMQ was moderate, supporting the notion that these two tests measure distinct aspects of memory distrust. The SSMQ but not the MDS had a moderate correlation with sociometric status, possibly because SSMQ also taps into a more general self-efficacy by using phrases like "my ability to remember". Consistent with Nash et al. (2022), the MDS had a weak-to-moderate negative correlation with age. However, we did not detect a statistically significant correlation between SSMQ and age. One possible explanation is that the current sample lacked variation in the age range. In sum, the results showed that the MDS had good criterion validity.

Table 2

Means, standard deviations, and correlations with 95% confidence intervals

Variable	*M*	*SD*	1	2	3	4	5
1. MDS	3.38	1.20					
2. SSMQ	1.62	1.09	-.35** [-.43, -.27]				
3. PRMQ	2.24	0.67	.61** [.55, .67]	-.48** [-.55, -.40]			
4. GCS	0.65	0.21	.33** [.24, .41]	-.28** [-.37, -.20]	.22** [.13, .31]		
5. Social Status	4.23	1.14	-.05 [-.14, .05]	.46** [.38, .53]	-.20** [-.28, -.11]	-.23** [-.32, -.15]	
6. Age	22.56	3.53	-.17** [-.25, -.08]	.05 [-.04, .14]	-.05 [-.14, .04]	-.01 [-.10, .08]	-.01 [-.10, .08]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. PRMQ = Prospective and Retrospective Memory Questionnaire; GCS = Gudjonsson Compliance Scale; SSMQ = Squire Subjective Memory Questionnaire; MDS = Memory Distrust Scale.
* indicates $p < .05$. ** indicates $p < .01$.

Exploratory Analysis: Measurement Invariance Testing

As highlighted by the difference in item loadings in the two-factor EFA between the Chinese sample and the British sample, there could be cultural differences in the construct of memory distrust between China and the UK. Since the one-factor model will likely be most commonly used by researchers, and because it is important to examine whether the scores can be directly compared cross cultures, we performed measurement invariance (MI) testing for the selected one-factor model to examine whether and to what extent the MDS is interpreted in the same way across different groups of individuals (Putnick & Bornstein, 2016). For the first model (M1), all the loadings and intercepts were freely estimated in both groups. The second model restricted the item loadings to be equal, testing metric invariance (i.e., Do the items load on to the latent construct in the same way across groups?), while the third model restricted both loadings and intercepts to be equal between the two groups, testing scalar invariance (i.e., Do the items have same intercepts across groups?).

The results are presented in Table 3. M1 showed an acceptable fit, with CFI = .931, RMSEA = .058, and SRMR = .042, indicating configural invariance, that is, the organization of the construct was same across groups. All items showed adequate loading on the factor, ranging from .493 to .805 (standardized) across groups. According to the recommendation of Chen (2007), for analyses with adequate sample size (total N > 300) and equal sample size across groups, a $\Delta CFI \leq -.010$, supplemented by a $\Delta RMSEA \geq .015$ or a $\Delta SRMR \geq .030$ would indicate metric non-invariance. For testing scalar invariance, a $\Delta CFI \leq -.010$, supplemented by a $\Delta RMSEA \geq .015$ or a $\Delta SRMR \geq .010$ would indicate scalar non-invariance. Based on our analyses then, we concluded that MDS achieved metric invariance in the Chinese and UK sample, but scalar invariance was rejected. This means that the underlying

391 factor across groups had the same unit (i.e., an increase of 1 in the MDS items has the
392 same meaning across groups) but the intercepts of the items were different. For
393 example, this could mean that participants in China tend to react with higher
394 agreement to some items compared with British participants.

Table 3

Measurement Invariance Testing across Chinese and UK Samples

Model	df	χ^2	$\Delta\chi^2$	Δdf	*p*	CFI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
M1	340	318.71	-			.931	.059	.042	-	-	-
M2	359	715.62	106.10	19	<.001	.923	.060	.063	−.008	.001	.021
M3	378	916.91	389.47	19	<.001	.896	.068	.071	−.027	.008	.008

Study 1b

Participants and Procedure

Four weeks after the data collection of Study 1a, we sent the follow-up survey via email to all participants whose data were included in Study 1a, and received 301 responses in total. Three people failed the attention check and were left out of the analyses, leaving a sample of 298 participants (n_{woman} = 178, 59.7%; M_{age} = 22.56, SD_{age} = 3.37). In the follow-up survey, we included the MDS as well as several ad-hoc questions about participants' daily habits for exploratory purposes only (see Table 4). Participants responded to all items on a 7-point scale from 1 = strongly disagree to 7 = strongly agree.

Results

The internal consistency of the Chinese MDS was again very good (Cronbach's α = .95, McDonald's ω = .96). Correlation analyses showed that the Chinese version of the MDS had adequate test-retest reliability, r (297) = .70, 95% CI [.64, .75], similar to the results of Nash et al. (2022). Moreover, as shown in Table 5, memory distrust measured by the MDS had stable correlations with both the SSMQ and age. Overall, the Chinese MDS has good internal reliability, adequate test-retest reliability, and good criterion validity.

Exploratory analyses

Exploratory analyses showed that certain life habits were related to memory distrust. Participants who reported that they had a more organized daily life had lower memory distrust, as indicated by the lower MDS score and higher SSMQ scores, than

participants reporting a less organized daily life. On the other hand, people who consumed more TV or movie products had higher scores on the MDS, but not on the SSMQ, than their counterparts with lower TV consumption, which given that we might expect TV consumption to be associated with source monitoring errors, is consistent with the fact that the MDS measures memory distrust toward making commission errors. Finally, the affordance of corroborative memory cues (i.e., being able to find corroborative cues for one's memories) was associated with lower memory distrust.

One puzzling result is that SSMQ was positively related to all habits except for the consumption of movies or TV products. However, this does not mean that these habits are necessarily related to memory distrust. Since sociometric status was also positively associated with SSMQ and the endorsement of the habits with the exception of consuming TV (see https://osf.io/jgnez?view_only=b1f3cc0996d74822a7696041977a8da5), we performed regression analyses examining whether these habits predicted SSMQ after controlling for sociometric status. When status was controlled for, only being organized (B = 0.12, SE = 0.037, p = .001) and the affordance of memory cues (B = 0.24, SE = 0.055, p < .001) were associated with higher SSMQ scores (i.e., greater memory trust). Taken together, the analyses suggested that memory distrust could be influenced by a person's habits and life structures. However, due to the exploratory nature of these analyses, we caution against drawing strong conclusions from the current results.

Table 4

Exploratory Items Probing Life Habits

Items Assessing Participants' Habits for Exploratory Purpose

1 Organized	My daily life is routinized and organized
2 Interact with people	I interact with many different people daily
3 Reading	I spend lots of time reading literature
4 TV or movie	I spend lots of time watching TV series or movies
5 Social Media	I tend to use social media to record my life
6 Memory Cues	Generally, I can find cues that corroborate my daily experiences

444

Table 5

Means, standard deviations, and correlations with confidence intervals

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. 2nd MDS	3.31	1.10									
2. MDS	3.33	1.10	.70** [.64, .75]								
3. SSMQ	1.60	1.00	-.43** [-.52, -.33]	-.41** [-.50, -.31]							
4. Age	22.56	3.37	-.18** [-.29, -.07]	-.17** [-.28, -.06]	.12* [.01, .24]						
5. Organized	4.87	1.50	-.12* [-.23, -.00]	-.01 [-.12, .11]	.28** [.17, .38]	.08 [-.04, .19]					
6. Interact with people	4.16	1.58	.02 [-.09, .14]	.11 [-.00, .22]	.22** [.11, .33]	-.05 [-.16, .06]	.20** [.09, .31]				
7. Reading	3.94	1.54	-.12* [-.23, -.00]	-.03 [-.14, .08]	.23** [.12, .34]	-.07 [-.18, .04]	.13* [.01, .24]	.32** [.22, .42]			
8. TV or movie	4.13	1.64	.17** [.05, .28]	.14* [.02, .25]	.05 [-.07, .16]	-.12* [-.24, -.01]	.08 [-.04, .19]	.06 [-.05, .18]	.14* [.03, .25]		
9. Social Media	4.37	1.73	.04 [-.08, .15]	.11 [-.01, .22]	.13* [.02, .24]	-.13* [-.24, -.01]	.18** [.07, .29]	.27** [.17, .38]	.23** [.12, .33]	.16** [.04, .26]	
10. Memory Cues	5.10	1.02	-.23** [-.34, -.12]	-.20** [-.30, -.08]	.36** [.26, .46]	.09 [-.03, .20]	.16** [.05, .27]	.19** [.08, .30]	.21** [.10, .32]	.02 [-.09, .13]	.28** [.17, .38]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. * indicates $p < .05$. ** indicates $p < .01$.

Study 2

Study 1 established that the Chinese MDS has good internal consistency and adequate test-retest reliability across a one-month interval. In Study 2, we further tested the criterion validity (convergent and discriminant) of the MDS, by asking three questions. First, is memory distrust (measured by both the MDS and the SSMQ) related to the acceptance of false social feedback about making memory commission/omission errors? Second, does the MDS have a stronger association (compared with the SSMQ) with the acceptance of feedback about making commission errors? And third, does the SSMQ have a stronger association (compared with the MDS) with the acceptance of feedback about making omission errors?

Method

Participants

We recruited Chinese participants from university participant pools and via social media. The study was hosted on Qualtrics. To participate in the study, participants were required to read the information letter introducing the aim, tasks, and compensation scheme of the study, and to provide informed consent. Furthermore, they had to be over 18 years old and not from a psychology major. Participants were compensated with 10 RMB. To incentivize participants to perform well in the recognition tasks, we also offered a 5-RMB bonus for participants who ranked in the top 10% of scorers in these tasks.

Following Nash et al. (2022), we planned to recruit at least 400 participants after exclusions. We planned to oversample during the first part of the experiment with the expectation that there would be enough valid entries after dropouts and exclusions in

the second part of the experiment. A total of 475 participants completed the memory task, the first part of the experiment (see Procedure for details) but only 407 completed the trait measures. Among the 407 participants who completed both parts of the experiment, 24 failed at least one attention check. Thus, the sample that contains both memory task results and trait measures comprised 383 individual data entries (n_{women} = 204, n_{men} = 176, $n_{\text{no disclosure}}$ = 3). Using G*Power 3.1 (Faul et al., 2009), a sensitivity analysis for bivariate correlation showed that with α = .05 and 1-β = .95, a sample of 383 could reliably detect an effect no smaller than ρ = .18. The average age of the complete sample following exclusions was 22.2 (*SD* = 4.33). Nearly all participants had an education level of a college degree or above (96%).

Materials

Stimuli for the Memory Task. A total of 40 mildly positively valenced color images were selected from the OASIS, an open-access stimulus set with normative ratings for valence and arousal on 7-point scales (Kurd et al., 2017). The stimuli from OASIS depict a broad spectrum of natural or social situations (e.g., buildings or car accidents). Twenty scenes were used in Session 1 for encoding (hereafter referred to as old scenes), as described below. For the recognition tasks in Sessions 2 and 3, 40 scenes (20 old ones and 20 new ones) were presented to participants. One-way ANOVAs showed that the old and new stimuli did not differ on valence, M_{old} = 4.95, SD_{old} = 0.80; M_{new} = 4.69, SD_{new} = 0.62, F (1, 38) = 1.26, p = .268, η_p^2 = .03, or arousal scores, M_{old} = 2.95, SD_{old} = 0.87; M_{new} = 2.86, SD_{new} = 0.66, F (1, 38) = 0.12, p = .734, η_p^2 = .003.

Memory Distrust. We employed the Chinese MDS and Chinese SSMQ to measure memory distrust, as in Study 1. Both scales showed good internal reliability

in the current sample (SSMQ: Cronbach's α = .94, McDonald's ω = .95; MDS: Cronbach's α = .95, McDonald's ω = .96). For each scale, we calculated the mean score of all items and used it as the index of memory distrust.

Procedure

After reading the information letter and giving informed consent, participants viewed 20 scene images, one at a time and in randomized order. Each scene was presented for 5 seconds. After all scenes had been viewed, participants completed 20 addition/subtraction problems[2], and they then moved on to the first recognition task in which 20 old scenes and 20 new scenes were presented one at a time in a random order without a time limit. In the recognition task, participants were first asked to indicate whether a scene was old or new, after which they were shown feedback on the screen for 5 seconds, which supposedly communicated the recognition response given to that same image by another, randomly selected participant from the study. Participants were told that all participants saw the same set of scenes, but that their responses might differ from one another and they were instructed to pay attention to the feedback. In reality, all the feedback was pre-determined with a probability of 25% being false. That is, for old scenes, the algorithm had a 25% probability of falsely advising participants that the previous participant had judged the stimulus as new. For new scenes, the algorithm had a 25% probability of falsely advising participants that the previous participant had judged the stimulus as old.

Immediately after the first recognition task, participants completed another 20 math problems as a distraction, and then completed the second recognition task. This second recognition task was the same as the first, except that participants only made

[2] We did not record the duration of the distraction tasks. Estimated completion time is 3-5 minutes.

old-new judgments without receiving feedback.

Three days after the memory tasks, participants received the trait measures survey via email. After completing the MDS and SSMQ, participants answered demographic questions about their age, gender, and education level. Then they answered four questions: 1) “To what extent did you find the experiment procedures difficult to understand?” (1 = not difficult at all, to 7 = very difficult); 2) “How serious were you when completing the experiment<” (1 = not serious at all, to 5 = very serious)[3]; 3) “What do you think is the purpose of the experiment?” (Open-ended); 4) “Did you notice any errors in the experiment or do you have any suggestions to improve the experiment?” (open-ended).

After the data collection was completed, we calculated and ranked the accuracy for each participant based on the first recognition task, since their initial old-new responses were made prior to the feedback. We then paid participants their compensation and the bonus where applicable, and debriefed them.

Data Analysis Overview

First, we ran a confirmatory factor analysis to validate the one-factor solution of the MDS in a second sample. Then we examined whether providing false feedback in the first recognition task would increase the chance of participants making errors of commission (in the case of identifying new stimuli) or omission (in the case of identifying old stimuli) in the second recognition task. More specifically, for the latter we used the lme4 package (Bates et al., 2014) to run generalized linear mixed models (GLMM) with the recognition outcome (correct vs. incorrect) as the dependent variable, and whether false feedback had been provided on that item (yes vs. no) as

[3] No participant reported being not serious (all scores $\geqslant$3)

the fixed effect. For random effects, we included random intercepts for Participant ID and Scene ID.

To test our hypotheses that memory distrust is related to the acceptance of false feedback, we ran additional separate generalized linear mixed models for recognition outcomes of old vs. new stimuli, with false feedback, memory distrust (either SSMQ or the MDS), and their interaction terms as fixed effects, and we included random intercepts for Participant ID and Scene ID in both models. Furthermore, we then compared the power of the SSMQ and MDS to predict commission errors and omission errors.

Less central to the main purpose of the study, we also calculated participants' accuracy in the first recognition task for both old and new stimuli, and used pairwise correlation and multivariate regression analyses to examine the associations of these accuracy scores with memory distrust.

Results

Confirmatory Factor Analysis

As shown in the EFA in Study 1, the adopted one-factor solution showed adequate item loadings for all 20 items. Using the data from Study 2, we conducted a confirmatory factor analysis for this one-factor model based to further test its validity. Similar to Study 1, we used the WLSMV estimator since the Study 2 data also violated the multivariate normality assumption (Mardia test: skew = 3792.20, $p < .001$; kurtosis = 37.49, $p < .001$). The chi-squared test for the one factor model was statistically significant, χ^2 (190) = 309.35, $p < .001$. However, due to its oversensitivity to sample size, the chi-squared test was not used to evaluate the model fit. Model fit indices (CFI = .952, TLI = .947, RMSEA = .046, SRMR = .043)

suggested that the model had good fit (Hu & Bentler, 1999). We thus conclude that the one-factor model was validated in an independent sample.

Correlation between Memory Distrust and Memory Performance

As shown in Table 7, we replicated the moderate correlation between MDS and SSMQ, as well as the negative correlation between MDS and age. Participants' accuracy was high across both recognition tasks. Moreover, the recognition accuracy of old stimuli (i.e., the Hit rate) and of new stimuli (i.e., the Correct Rejection rate) were only moderately correlated. The MDS had a negative weak-to-moderate correlation with the recognition of old stimuli, but not with the recognition of new stimuli.

To further examine the association between memory distrust and memory performance, we computed the Signal Detection Theory (SDT, Green & Swets, 1966) indices d' and β. Note, however, that the SDT analysis was not pre-registered and is thus exploratory. A higher *d'* would indicate higher sensitivity when distinguishing old scenes from new scenes, while β reflects an observer's bias to say 'old' vs. 'new', with the unbiased observer having a value around 1.0. A higher β would indicate a more conservative criterion biased toward saying 'new'. Results showed that the MDS was negatively correlated with d' (r = -.18, 95%CI [-.27, -.08]) but positively correlated with β (r = .19, 95%CI [.09, .28]) in the first recognition test. The opposite pattern was found for the SSMQ, d': r = .07, 95%CI [-.03, .17]; β: r = -.16, 95%CI [-.25, -.06]). Regression analysis showed that when both the MDS and SSMQ were entered into the model, only the MDS (B = 0.19, SE = 0.07, p = 0.009) but not the SSMQ (B = -0.13, SE = 0.07, p = 0.090) was a significant predictor of β. We found similar results for the regression analysis for d', with only MDS being a statistically

590 significant predictor (B = -0.16, SE = 0.05, p = 0.002). The results lend support to the
591 idea that people who are more concerned with making commission errors may be
592 biased to say ‘new’ in memory recognition tests.

Table 7

Accuracy and Memory Distrust

Variable	*M*	*SD*	1	2	3	4	5	6
1. SSMQ	1.64	1.14						
2. MDS	3.38	1.19	-.42** [-.50, -.34]					
3. Hits_1st	0.87	0.15	.09 [-.01, .19]	-.23** [-.32, -.13]				
4. Correct_Rejections_1st	0.91	0.15	-.02 [-.12, .08]	-.05 [-.15, .05]	.43** [.35, .50]			
5. Hits_2nd	0.89	0.13	.04 [-.06, .14]	-.17** [-.27, -.07]	.83** [.80, .85]	.42** [.35, .49]		
6. Correct_Rejections_2nd	0.81	0.23	.03 [-.07, .13]	-.09 [-.19, .01]	.45** [.38, .52]	.71** [.66, .75]	.31** [.22, .39]	
7. Age	22.20	4.33	-.05 [-.15, .05]	-.13* [-.22, -.03]	.04 [-.06, .14]	.03 [-.07, .13]	.02 [-.08, .12]	-.04 [-.14, .06]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. * indicates $p < .05$. ** indicates $p < .01$.

The Effect of False Feedback on Memory Omission and Commission Errors

Generalized linear mixed models (GLMM) showed that false feedback had a negative impact on the correct recognition of old stimuli in the second recognition task, B = -0.24, SE = 0.08, p =.002. However, it only explained very limited variation (Pseudo-R^2 = .002). False feedback also had a negative effect on the correct recognition (i.e., rejection) of new stimuli, B = -0.32, SE = 0.07, p <.001 (Pseudo-R^2 = .003). To test the robustness of these results, we included the recognition outcome of the same stimulus in the first recognition task as a control variable**,** and ran two additional models. The two models revealed similar results that participants were more likely to make an incorrect recognition response after being exposed to false feedback, regardless whether or not they made the correct response in the first test (Old Stimuli: B = -0.46, SE = 0.10, p <.001; New Stimuli: B = -0.34, SE = 0.07, p <.001).These results showed that the experimental manipulation of false feedback was successful.

The Moderation effect of Memory Distrust

We next proceeded with examining whether memory distrust (measured by the Chinese MDS and the Chinese SSMQ) moderated the effects of false feedback on errors of omission and commission. Per our preregistration, we included false feedback, memory distrust (either the MDS or the SSMQ), and their interaction terms as fixed effects and we included random intercepts for Participant ID and Scene ID in all GLMMs. Results showed that for the recognition of old stimuli, the interaction between false feedback and MDS was not significant, albeit was in the predicted direction, B = -0.16, SE = 0.08, p =.065. When the recognition result of the first task was controlled for (exploratory), the interaction term became significant, B = -0.26,

SE = 0.10, *p* =.015, with the effect of false feedback on omission errors being greater among participants with higher MDS scores. As for the SSMQ, the interaction term did not reach the conventional significance level whether the first recognition outcome was controlled for (*B* = 0.18, *SE* = 0.10, *p* =.079) or not (*B* = 0.11, *SE* = 0.08, *p* =.197), thus rejecting the possibility that the effect of false feedback on omission errors differed among participants who scored high or low on SSMQ.

As for the recognition outcome of new stimuli, neither the interaction of false feedback with MDS score (*B* = 0.05, *SE* = 0.08, *p* =.502) nor with SSMQ score (*B* = 0.11, *SE* = 0.08, *p* =.197) was a significant predictor. The same pattern held when the first recognition outcome was controlled for (MDS: *B* = 0.08, *SE* = 0.08, *p* = .322; SSMQ: *B* = -0.12, *SE* = 0.08, *p* =.122). That is, the effect of false feedback on commission errors in the second test did not differ between people who were either high or low on memory distrust.

Exploratory Analysis. To further test the potential moderation effect of memory distrust, we excluded 30 participants who correctly guessed the purpose of the experiment (e.g., "to examine the effect of others' memory on one's memory report") and re-ran the above GLMM analyses. Results again showed that the interaction between MDS and false feedback was a significant negative predictor of correct recognition of old stimuli (*B* = -0.18, *SE* = 0.09, *p* =.039). That is, the effect of false feedback on omission errors in the second test was greater among participants who scored higher on MDS than their counterparts who scored lower on MDS. We did not detect a significant interaction between MDS and false feedback in the case of recognition of new stimuli (*B* = 0.04, *SE* = 0.08, *p* =.597). For the models examining the moderating effect of the SSMQ, we did not detect significant interactions in either case (Old stimuli: *B* = 0.11, *SE* = 0.09, *p* =.193; New stimuli: *B* = -0.08, *SE* = 0.08, *p*

=.335). Analyses with the recognition result of the first test as a control variable showed similar results that only the interaction between MDS and false feedback was a significant negative predictor of correct recognition of old stimuli (B = -0.27, SE = 0.11, p =.012).

Taken together, these results suggest 1) that high memory distrust as measured by MDS might predispose individuals to accepting false feedback about having made commission errors, and thus to make omission errors in subsequent memory tasks as a result, and 2) that compared with the SSMQ, MDS was a better predictor of accepting the false feedback that one has made commission errors. However, we failed to find support that high memory distrust as measured by SSMQ predisposed people to accepting false feedback about having made omission errors.

General Discussion

How people view, reconstruct, and report their memories can be shaped by their relatively stable beliefs about their own memory functioning, whether these beliefs are accurate or not. The current studies aimed to translate and validate a Chinese version of the MDS, and to empirically examine how memory distrust contributes to the occurrence of commission and omission errors in people's recognition memory reports.

First and foremost, the current study provides evidence that the Chinese version of the MDS has excellent internal consistency and adequate test-retest reliability across four weeks, comparable to the English version reported by Nash et al. (2022). As for the criterion validity of the MDS, consistent with Nash et al. (2022), the MDS had a moderate correlation with SSMQ, supporting that both the MDS and the SSMQ measure two related but distinct aspects of memory distrust. Both MDS and SSMQ

were also moderately related to the PRMQ, a self-report measure of memory functioning, suggesting that people who are high on memory distrust also tend to report more memory errors. Furthermore, both the MDS and SSMQ had a moderate correlation with the GCS, further establishing the link of memory distrust with the susceptibility to social influence. Interestingly, we also discovered that, unlike the SSMQ, the MDS was unrelated to self-reported social status. One potential explanation is that the items in the SSMQ are framed in terms of ability, and thus are influenced by the general appraisal of one's self-efficacy. We speculate that this unexpected difference could suggest that when examining the relationships between memory distrust and other psychological phenomena, MDS could introduce fewer confounds than the SSMQ.

Besides the similarities across the British and Chinese population, we also noticed potential cultural differences in the construct of memory distrust. First, in the two-factor solution of the Chinese MDS, many items tapping into the social aspect (e.g., being persuaded by others' memory) loaded onto the first factor while only a few items emphasizing source monitoring loaded onto the second factor. This result is inconsistent with the result of Nash et al. (2022) that the second smaller factor was associated with memory distrust related to social influence, thus hinting that social influence may play a more important role in memory distrust in the Chinese population. Further measurement invariance tests showed that although all items loaded on the latent construct similarly across the two populations, the intercepts of the items were different. This means that one group tends to agree with (some of) the statements more than the other group, and that it is therefore not appropriate to compare latent means between groups.

Exploratory analyses also revealed that memory distrust, despite being a stable

individual difference, can be associated with certain daily habits or life structures. More-organized people have lower memory distrust than people with a less-organized daily life. On the other hand, being able to find evidence that corroborate one's memory was associated with lower memory distrust. Of theoretical relevance, TV/movie consumption was positively related to MDS scores but not to SSMQ scores, which may fit with the claim that the MDS measures memory distrust specifically toward making commission errors.

Study 2 further examined the validity of the MDS as well as the SSMQ, using an experimental false-feedback design. We found that participants who had higher memory distrust performed worse in the memory tasks, as compared with those who had lower memory trust. Further analysis under the framework of Signal Detection Theory (Green & Swets, 1966) showed that when compared with people who were lower on memory distrust, people who were higher on memory distrust had both a lower sensitivity and a response bias toward saying "new". Regression analyses revealed that the MDS was more closely associated with response bias to say "new" than was the SSMQ, consistent once again with the idea that the MDS taps into concerns about making commission errors.

After establishing that our false-feedback manipulation was successful in inducing commission and omission errors, we examined whether memory distrust would moderate the relationship as hypothesized. Participants who were more concerned about making commission errors (as measured by the MDS) were more likely to accept the false feedback that they could be wrong, and to therefore make omission errors in subsequent tests, than were those who were less concerned about making commission errors. However, no significant results were found regarding the interaction between SSMQ (tapping into distrust toward omission errors) and the

false-feedback. Further inspection of the data revealed that in total, 16.08% (n = 1504) of the recognition outcomes for new scenes changed from the first to the second test. In the meanwhile, the rate of changing recognition answers was 8.34% (n = 791) for the old scenes. This could mean that participants with different levels of memory distrust were equally more likely to change their recognition decisions for the new scenes and the analyses suffered from a ceiling effect. Further, in both tests, we used the same set of filler scenes. As a result, in the second test, participants needed to distinguish scenes they saw during the encoding and the scenes they saw only in the first test. Participants who were more confident in their memories might have mistakenly attributed the memory of new (i.e., fillers) scenes to the encoding task instead of the first recognition task. Therefore, until additional studies explore these relationships further, we caution against the interpretation that the SSMQ is a poorer predictor of memory errors than is the MDS.

Theoretical and Practical Implications

Across two studies, we found evidence that social influence plays an important part in our remembering processes. People who are more skeptical about their memories were more likely to accept false feedback and change their memory reports accordingly. This is consistent with the conjectures of Nash et al. (2022) as well as Zhang et al. (2022b). When people are confronted that their memory might be false, they engage in a weighing process comparing their internal representations and the external information (Scoboria & Henkel, 2020). In this process, people's meta-memorial beliefs could impact how they evaluate their specific memories, leading to either sticking with one's prior belief or accepting the external information (Zhang et al., 2022b). This is also corroborated by our results in Study 1 as well as that of Nash

et al. (2022) that people who are more skeptical about their memories also reported to be more compliant. In forensic settings where the completeness and veracity of memory reports are crucial, either the withholding of information or the acceptance of external information due to one's memory distrust can have severe implications. For example, withholding information could lead to failures to prosecute due to lack of evidence, while acceptance of false external information could lead to even more severe outcomes such as prosecuting the wrong person, resulting in the miscarriage of justice. Although our methodology is a far stretch from how police interviews are conducted, our results offer tentative evidence that people (e.g., suspects and/or witnesses) might react to suggestive feedback differently based on their metacognitive appraisals. Hence, measuring trait memory distrust could be of interest when evaluating witnesses' statements. Moreover, for clinical researchers, although the current research did not examine the relationship between memory distrust and clinical symptoms such as checking behavior (see for example, Coles et al., 2006 and Wong et al., 2022), we do believe that the MDS would be a useful tool for future research in this area.

Limitations and Future Directions

It is important to convey the limitations of the present work. First, several of the analyses in the current research were not preregistered. Although the exploratory nature of these analyses (e.g., life habit measure and SDT analyses) has been emphasized throughout the manuscript, our exploratory findings merit further investigation to confirm their replicability. Second, the items measuring life habits and structures were created ad-hoc and unlikely to represent valid or complete measures of those constructs. Future studies could build on this exploratory analysis

using more validated measures (e.g., Creature of Habit Scale, Ersche et al., 2017; TV consumption, Seabrook et al., 2016). Finally, and more importantly, the experimental false-feedback manipulation in Study 2, although successful, had a relatively weak effect on participants' responses. For example, in the second test, participants' accuracy for the old stimuli with correct feedback was 89.48%, while their accuracy for the old stimuli with false feedback was 87.27%. Similar results were found for the new stimuli (with correct feedback: 81.84%, with false feedback: 77.79%). These effects might have been so small for several reasons. The short interval between encoding and recognition and the uniqueness of the stimuli may have made the feedback not very believable. Furthermore, participants' attentiveness to the feedback in an online setting may have not been optimal. Finally, whereas we told participants what answers another participant had supposedly given, there was no reason for them to treat this other participant as especially credible or reliable. These problems, as well as the high overall rates of changing responses from the 1st to the 2nd recognition test, may have contributed to the non-significant results regarding the SSMQ. Therefore, further studies could include stronger experimental manipulations (e.g., delivering credible feedback in person) and extend the interval between encoding and testing to have a more robust test on the effect of memory trust on the susceptibility of accepting suggestions.

Conclusion

The present study validated a Chinese version of the MDS with good internal consistency, test-retest reliability, and criterion validity. Life habits such as consuming TV or movie products could influence memory distrust measured by MDS. Moreover, people with high level of memory distrust were more likely to

accept false feedback and make omission errors in subsequent recognition task. Our research takes an important step in developing theory and evidence on memory distrust in non-WEIRD cultural contexts, and the Chinese version of the MDS could be an effective tool for measuring memory distrust in Chinese populations.

Acknowledgements

We thank Mengying Zhang for helping with the translation process.

Conflict of Interests

The authors declare that they have no conflicts of interest to disclose.

References

Anderson, C., Kraus, M. W., Galinsky, A. D., & Keltner, D. (2012). The Local-Ladder Effect: Social status and subjective well-being. *Psychological Science*, *23*(7), 764–771. https://doi.org/10.1177/0956797611434537

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

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), 1-16.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: a Multidisciplinary Journal*, *14*(3), 464-504.

Coles, M. E., Radomsky, A. S., & Horng, B. (2006). Exploring the boundaries of memory distrust from repeated checking: Increasing external validity and examining thresholds. *Behaviour Research and Therapy*, *44*(7), 995-1006.

Ersche, K. D., Lim, T. V., Ward, L. H., Robbins, T. W., & Stochl, J. (2017). Creature of habit: A self-report measure of habitual routines and automatic tendencies in everyday life. *Personality and Individual Differences*, *116*, 73-85.

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. (1989). Compliance in an interrogative situation: A new

scale. *Personality and Individual Differences*, *10*(5), 535-540.

Gudjonsson, G. H., & MacKeith, J. A. C. (1982). False confessions. Psychological effects 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*(3), 336-348.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and Brain Sciences*, *33*(2-3), 61-83.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal*, *6*(1), 1-55.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, 88(1), 67– 85. https://doi.org/10.1037/0033-295X.88.1.67

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.

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*(4), 341-348.

Le, L. C., Hoang, Y. H., Bui, H. T., Nguyen, D. Q., Mai, S. T., & Luong, H. T. (2023). Wrongful convictions in Asian countries: A systematic literature review. *International Journal of Comparative and Applied Criminal Justice*, 1-17.

Loftus, E. F., Hoffman, H. G., & Wagenaar, W. A. (1992). The misinformation effect: Transformations in memory induced by postevent information. In M.L. Howe,

C.J. Brainerd & V.F. Reyna (Eds.), *Development of long-term retention* (pp. 159-183). Springer.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99.

Mahr, J.B. (2023), How to Become a Memory: The individual and collective aspects of mnemicity. *Topics in Cognitive Science*. https://doi.org/10.1111/tops.12646

Mahr, J. B., Van Bergen, P., Sutton, J., Schacter, D. L., & Heyes, C. (2022). Mnemicity: A cognitive gadget?. *Perspectives on Psychological Science*, 17456916221141352.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224-253.

Mewton, L., Sachdev, P., Anderson, T., Sunderland, M., & Andrews, G. (2014). Demographic, clinical, and lifestyle correlates of subjective memory complaints in the Australian population. *The American Journal of Geriatric Psychiatry*, *22*(11), 1222-1232.

Nash, R. A., Saraiva, R. B., & Hope, L. (2022). Who doesn't believe their memories? Development and validation of a new Memory Distrust Scale. *Journal of Applied Research in Memory and Cognition.* Advance online publication. https://doi.org/10.1037/mac0000061

Oeberst, A., & Wu, S. (2015). Independent vs. interdependent self-construal and interrogative compliance: Intra-and cross-cultural evidence. *Personality and Individual Differences*, *85*, 50-55.

Otgaar, H., Bücken, C., Bogaard, G., Wade, K. A., Hopwood, A. R., Scoboria, A., & Howe, M. L. (2019). Nonbelieved memories in the false memory archive. *Journal of Applied Research in Memory and Cognition*, *8*(4), 429-438.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71-90.

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2012.

Radomsky, A. S., & Alcolado, G. M. (2010). Don't even think about checking: Mental checking causes memory distrust. *Journal of Behavior Therapy and Experimental Psychiatry*, *41*(4), 345-351.

Ross, M., & Wang, Q. (2010). Why we remember and what we remember: Culture and autobiographical memory. *Perspectives on Psychological Science*, *5*(4), 401-409.

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*(1), 94-106.

Schweizer, S., Kievit, R. A., Emery, T., & Henson, R. N. (2018). Symptoms of depression in a large healthy population cohort are related to subjective memory complaints and memory performance in negative contexts. *Psychological Medicine*, *48*(1), 104-114.

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.

Seabrook, R. C., Ward, L. M., Reed, L., Manago, A., Giaccardi, S., & Lippman, J. R. (2016). Our scripted sexuality: The development and validation of a measure of the heterosexual script and its relation to television consumption. *Emerging Adulthood, 4(5*), 338-355.

Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment, 31*(12), 1428–1441. https://doi.org/10.1037/pas0000623

Smith, G., Del Sala, S., Logie, R. H., & Maylor, E. A. (2000). Prospective and retrospective memory in normal ageing and dementia: A questionnaire study. *Memory (Hove, England), 8*(5), 311–321.

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.

Strauss, A. Y., Fradkin, I., McNally, R. J., Linkovski, O., Anholt, G. E., & Huppert, J. D. (2020). Why check? A meta-analysis of checking in obsessive-compulsive disorder: Threat vs. distrust of senses. *Clinical Psychology Review*, *75*, 101807.

Sumampouw, N., Bjørndal, L. D., Magnussen, S., Otgaar, H., & Brennen, T. (2022). Knowledge about eyewitness testimony: a survey of Indonesian police officers and psychologists. *Psychology, Crime & Law*, *28*(8), 763–777.

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

van Bergen, S., Horselenberg, R., Merckelbach, H., Jelicic, M., & Beckers, R. (2010b). Memory distrust and acceptance of misinformation. *Applied Cognitive Psychology*, *24*(6), 885-896.

van Bergen, S., Jelicic, M., & Merckelbach, H. (2009). Are subjective memory problems related to suggestibility, compliance, false memories, and objective memory performance?. *The American Journal of Psychology, 122*(2), 249-257.

Wang, J., Otgaar, H., Santtila, P., Shen, X., & Zhou, C. (2021). How culture shapes

constructive false memory. *Journal of Applied Research in Memory and Cognition, 10*(1), 24–32. https://doi.org/10.1037/h0101792

Wang, Q. (2016). Remembering the self in cultural contexts: A cultural dynamic theory of autobiographical memory. *Memory Studies*, 9(3), 295-304.

Wang, Q. (2021). The cultural foundation of human memory. *Annual Review of Psychology*, *72*, 151-179.

Watkins, M. W. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Wong, M. L., Leung, C. N. W., Lau, K. N. T., Chung, K. F., & Lau, E. Y. Y. (2022). The relationships among sleep problems, anxiety, memory complaints and compulsive checking behaviours. *Journal of Obsessive-Compulsive and Related Disorders*, *34*, 100728.

Yang, T. X., Wang, Y., Wang, Y., Su, X. M., Ni, K., Lui, S. S., & Chan, R. C. (2022). Validity and normative data of the Chinese Prospective and Retrospective Memory Questionnaire (PRMQ) across adolescence, adults and elderly people. *Memory*, *30*(3), 344-353.

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.*

Zhang, Y., Otgaar, H., & Wang, J. (2022b). Memory distrust is related to memory errors, self-esteem, and personality. *Applied Cognitive Psychology*, *36*(2), 283-292.

Appendix

Table A1

Factor Loading in the Two-Factor EFA model

Scale Items		F1	F2
1	我经常通过寻找物证（例如照片）来确认事情是否真的以我所记忆的方式发生。 I often look for physical evidence, such as photographs, to check whether things really happened the way I remember them.		.56
2	我经常求助于他人来帮助我确定自己的记忆是否准确。 I often turn to other people to help me decide whether my memories are accurate.	.36	.36
3	如果他人不能证实我对过去事件的记忆，我倾向于对自己的记忆产生质疑。 I tend to question my memories of past events if other people do not corroborate what I remember.	.73	
4	有时，如果我找不到任何物证来佐证自己的记忆，就会不信任自己的记忆。 Sometimes I distrust my own memories if I cannot find any physical evidence to confirm what I remember.	.64	
5	我经常难以区分自己的回忆和单纯的想象。 I often have difficulty distinguishing events I remember from those I only imagined.	.48	.37
6	我经常不确定自己记得的事情是真的发生过还是仅仅想过或梦到过。 I am often unsure whether something that I recall genuinely happened, or whether I only thought or dreamed about it.		.68
7	我相信我的一些记忆可能完全来自想象。 I believe some of my memories may have originated entirely from my imagination.	.38	.35
8	我有时不确定我记得的事件是真的发生在我身上，又或者只是我在电视上或电影中看到过。 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.		.80
9	有时，他人对过往的描述会使我怀疑自己的对那些事件的回忆。 Other people sometimes describe past events in ways that make me doubt my own recollection of those events.	.34	.37
10	我很容易被说服自己记得的事件不可能发生过。 I could be easily persuaded that an event I remember is impossible.	.83	
11	如果我和另一个人对过去的回忆相矛盾，那很可能他/她的回忆是正确的。 If another person contradicts my recollection of the past, they are probably correct.	.78	
12	在适当的情况下，我可以被说服，自己的任何一个记忆都是完全错误的。 Under the right circumstances, I could be persuaded that any one of my memories was completely false.	.76	
13	通常而言，相比于自己的回忆，我对他人的回忆持有更多的信任。	.75	

	I generally have more trust in other people's recollections of events than in my own recollections.		
14	我经常相信他人对某个事件的描述，即便我自己的回忆与其相差甚远。 I often trust other people's descriptions of a past event, even if I have a very different recollection of what happened.	.73	
15	相比于我自己的记忆，他人的记忆通常而言更准确。 Other people's memories are usually more accurate than my own memories.	.92	
16	我的记忆很少能非常准确地反映真正发生的事情。 My memories are rarely a very accurate reflection of what truly occurred.	.83	
17	我对过去事件的记忆不可靠。 My memories of past events are unreliable.	.69	
18	我难以总是确信自己的记忆准确地反映了事情的真正经过。 I cannot always be confident that my memories accurately reflect what really happened.	.46	.34
19	我对许多自己记得的事情的确发生过没有什么信心。 I have little trust that many of the events I remember did really occur.	.73	
20	有时我不相信自己记得的经历实实在在地发生过。 I sometimes distrust that certain experiences I remember really happened at all.	.36	.41

Note. Response scale: 1 = 非常不同意; 2 = 不同意; 3 = 略不同意; 4 = 中立; 5 = 略同意; 6 = 同意; 7 = 非常同意.

Response scale: 1 = strongly disagree; 2 = disagree; 3 = slightly disagree; 4 = neither agree nor disagree; 5 = slightly agree; 6 = agree; 7 = strongly agree.

For the English version of the Memory Distrust Scale, please cite the original article:

Nash, R. A., Saraiva, R. B., & Hope, L. (2022). Who doesn't believe their memories? Development and validation of a new Memory Distrust Scale. *Journal of Applied Research in Memory and Cognition.* Advance online publication. https://doi.org/10.1037/mac0000

959

960