

The anchoring effect in metamemory monitoring

Chunliang Yang^{1,2} · Bukuan Sun³ · David R. Shanks^{1,2}

Published online: 21 November 2017
© Psychonomic Society, Inc. 2017

Abstract Judgments about future memory performance (*metamemory judgments*) are known to be susceptible to illusions and bias. Here we asked whether metamemory judgments are affected, like many other forms of judgment, by numerical anchors. Experiment 1 confirmed previous research showing an effect of informative anchors (e.g., past peer performance) on metamemory monitoring. In four further experiments, we then explored the effects of uninformative anchors. All of the experiments obtained significant anchoring effects on metamemory monitoring; in contrast, the anchors had no effect on recall itself. We also explored the anchoring effect on metamemory control (restudy choices) in Experiment 4. The results suggested that anchors can affect metamemory monitoring, which in turn affects metamemory control. The present research reveals that informative and, more importantly, uninformative numbers that have no influence on recall itself can bias metamemory judgments. On the basis of the current theoretical understanding of the anchoring effect and metamemory monitoring, these results offer insight into the processes that trigger metacognitive biases.

Keywords Anchoring · Informative · Uninformative · Metamemory monitoring · Metamemory control

✉ Chunliang Yang
chunliang.yang.14@ucl.ac.uk

¹ Department of Experimental Psychology, University College London, London, UK

² Division of Psychology and Language Sciences, University College London, 26 Bedford Way, London WC1H 0AP, UK

³ School of Education, Fuqing Branch of Fujian Normal University, Fuqing, China

A large body of previous studies has shown that people's metamemory monitoring (i.e., metacognitive judgments about memory) can often be relatively accurate (for a review, see Rhodes, 2016). For example, people predict that difficult items are less likely to be remembered, and their predictions about their future memory performance are positively correlated with their actual test performance (Scheck, Meeter, & Nelson, 2004; Yang, Potts, & Shanks, 2017a, 2017b; Yang & Shanks, 2017). Nonetheless, recent research has revealed that a range of factors may bias people's metamemory monitoring. For instance, although the font size of study words has no effect on later recall, people give higher judgments of learning (JOLs; i.e., the judged likelihood that a given item will be remembered at a later test) to large- than to small-font-size words (Hu et al., 2015; Hu, Liu, Li, & Luo, 2016; Rhodes & Castel, 2008; Yang, Huang, & Shanks, 2017). Along the same lines, although loudness has no effect on memory retention, people give higher JOLs to loud than to quiet words (Rhodes & Castel, 2009). Finn (2008) found a framing effect in metamemory monitoring: People give higher memory performance predictions when they are asked to predict the likelihood that they will *remember* a given item than when they are asked to predict the likelihood that they will *forget* the item, although the frame has no effect on actual recall.

Besides miscalibration of metamemory monitoring, people's metamemory control (i.e., strategic choices such as restudy decisions and study time allocation) is intimately related to metamemory monitoring (Metcalf & Finn, 2008; Rhodes & Castel, 2009; Yang et al., 2017a), and miscalibrated metamemory monitoring often leads to inefficient metamemory control (Kornell & Bjork, 2008b; Soderstrom & Bjork, 2014; Yang et al., 2017a). If a learner believes, for instance, that massed learning is more effective than spaced learning (metamemory monitoring; Kornell & Bjork, 2008a; Yan, Bjork, & Bjork, 2016), she is likely to structure her

learning in an inefficient way by exposing herself to massed rather than to spaced materials (metamemory control; Tauber, Dunlosky, Rawson, Wahlheim, & Jacoby, 2013). Therefore, how people monitor their learning is a critical question for learners, educators, and psychologists, and understanding the processes involved in metamemory monitoring is central to understanding metamemory control.

Theories of metamemory monitoring

Although metamemory monitoring has been extensively explored for over 50 years (Hart, 1965), the underlying causal mechanisms are still unclear. There is substantial evidence that people's beliefs affect their JOLs (Hu et al., 2015; Mueller & Dunlosky, 2017; Mueller, Dunlosky, Tauber, & Rhodes, 2014). Hu et al. (2015), for instance, found that people hold a priori beliefs that large words are more memorable than small words and they incorporate these beliefs into the judgment process. Another possible cue is processing fluency: Fluent processing experience produces a feeling of knowing, which can form a basis for JOLs (Undorf & Erdfelder, 2015; Undorf, Zimdahl, & Bernstein, 2017; Yang et al., 2017). For example, Yang et al. (2017) found that perceptual fluency during encoding affects JOLs, and Undorf and Erdfelder found that conceptual fluency affects JOLs as well. These findings, jointly, support a *dual-process* theory which proposes that people's JOLs are based on both processing fluency and beliefs about remembering difficulty. However, other studies have found that, in some situations, processing fluency is unlikely to contribute to JOLs (Susser, Jin, & Mulligan, 2016; Witherby & Tauber, 2017). Therefore, Mueller, Dunlosky, and Tauber (2016) proposed an *analytic-processing* theory, which hypothesizes that people's beliefs about remembering difficulty play the major role in constructing JOLs and that processing fluency plays a less important or even no role (for detailed description and comparison of the dual-process and analytic-processing theories, see Mueller & Dunlosky, 2017, p. 246).

Unlike both the dual-process and analytic-processing theories, which focus on the roles of people's beliefs and/or processing fluency, a *cue-utilization* theory postulates that JOLs are based on a set of different cues or heuristics (Koriat, 1997, pp. 350–352). This theory highlights that, when making a JOL, people search for and incorporate a variety of diagnostic cues/heuristics to reduce prediction uncertainty and to reach a reasonable judgment. Specifically, this theory hypothesizes that three kinds of heuristics contribute to JOLs: intrinsic heuristics (natural characteristics of the study items; e.g., semantic relatedness of word pairs), extrinsic heuristics (the conditions of learning; e.g., study times, study methods), and mnemonic heuristics (subjective processing experiences; e.g., processing fluency, cue familiarity).

Below, we consider how these theories might accommodate the effects of numerical anchors on judgments of learning and on metamemory control.

Anchoring effects and the underlying mechanisms

The anchoring effect refers to the fact that numeric judgments tend to assimilate toward previously encountered numerical anchors. For example, Tversky and Kahneman (1974) first asked participants to judge whether the proportion of African nations in the UN is higher or lower than an arbitrary anchor (high anchor: 65%; low anchor: 10%). Then all participants made an absolute estimate. Those who encountered the high anchor gave higher estimates (45%) than did those who encountered the low anchor (15%). The anchoring effect, which has been explored for decades (Slovic, 1967), is a pervasive judgment bias, evident for example in legal judgments (Englich, Mussweiler, & Strack, 2006), willingness to pay and accept in purchasing (Simonson & Drolet, 2004), negotiations (Galinsky & Mussweiler, 2001), work performance ratings (Thorsteinson, Breier, Atwell, Hamilton, & Privette, 2008), and judgments of confidence (S. R. Carroll, Petrusic, & Leth-Steensen, 2009; for reviews, see Furnham & Boo, 2011; Newell & Shanks, 2014).

It is widely assumed that anchoring can be caused by several different mechanisms—there is not just a single “anchoring effect,” but several (S. R. Carroll et al., 2009; Furnham & Boo, 2011; Newell & Shanks, 2014). Two proposals have been particularly popular. The first is the *confirmatory hypothesis testing* or *selective-accessibility* theory (Chapman & Johnson, 1999; Strack & Mussweiler, 1997). This hypothesizes that encountering an anchor induces people to construct a mental model in which the target property is aligned with the anchor, and then test whether the anchor is the possible target property. Some supporting evidence for selective accessibility comes from a study by Mussweiler and Strack (2000). In their Experiment 2, Mussweiler and Strack asked participants to indicate whether a German car costs more or less than €40,000 (high anchor) or €20,000 (low anchor). The selective-accessibility theory assumes that the anchor semantically activates features of the target that are consistent with it. Thus in the presence of a high anchor, for example, features such as leather upholstery and a powerful engine will be activated. When subsequently asked to judge the car's cost, these activated features serve to bias participants' estimates upwards. To test the theory, Mussweiler and Strack asked their participants to complete a lexical decision task to measure attribute activation following the comparative question but prior to making their absolute judgment. They found that the high-anchor group responded faster to words associated with expensive cars (e.g., *Mercedes*) than to words associated with inexpensive cars (e.g., *Golf*), whereas the low-

anchor group showed the reverse pattern. These results imply that, consistent with the selective-accessibility theory, anchors semantically activate consistent features of the target.

The second is the *scale distortion* theory, recently proposed by Frederick and Mochon (2012). They suggested that the anchoring effect is triggered by a distortion of the psychological scale. This theory is based on the contrast effect (e.g., warm water is perceived to be warmer if people have previously placed their hand in cold water). Imagine that people think that 35% is a reasonable estimate of the proportion of African countries in the UN when no anchor is shown. A low anchor (10%) causes 35% to appear much larger than it otherwise would, which in turn leads people to give estimates lower than 35% and assimilate to the low anchor. In contrast, a high anchor (65%) makes 35% appear small, which leads to estimates higher than 35% and assimilation to the high anchor. In summary, the selective-accessibility theory postulates that anchors activate cues consistent with them and change the mental representation of the target property. In contrast, the scale distortion theory hypothesizes that anchors do not affect the representation of the target property, but only the psychological interpretation of the response scale.

Anchoring effect in metamemory monitoring

It is surprising that little research has been conducted to directly explore the anchoring effect in metamemory monitoring, because any detectable effects of anchors may provide important insight into the processes that underlie metacognitive judgments. One possible reason for this lacuna is that researchers tend to treat metacognition and judgment and decision-making (JDM) research as distinct. The present research is designed to bridge these lines of research by investigating the anchoring effect in metamemory monitoring. Specifically, we asked whether metamemory monitoring, like other kinds of metacognitive judgment, is susceptible to anchors.

Many previous studies have proposed that the anchoring effect may account for some metamemory illusions (i.e., disparities between JOLs and actual memory performance). For example, Scheck et al. (2004) and Finn and Metcalfe (2007) proposed that anchoring helps to explain the *underconfidence-with-practice effect* (Koriat, Sheffer, & Ma'ayan, 2002). Koriat (1997) instructed participants to study some word pairs across two study–test cycles. In the first cycle, participants studied each pair for 5 s and made a JOL to predict the likelihood that they would remember the pair, and then took a test on all pairs. In the second cycle, participants restudied all pairs, again made item-by-item JOLs, and took a final recall test. Koriat found a relatively good degree of agreement between participants' JOLs and recall in the first cycle. Nonetheless, in the second cycle, participants showed an

underconfidence-with-practice effect: Their JOLs were significantly lower than their recall. Finn and Metcalfe and Scheck et al. proposed that the JOLs in the second cycle were heavily based on test performance in the first cycle. Participants performed poorly in the first cycle, which set a low performance anchor. The additional exposure (restudy opportunities) in the second cycle significantly enhanced subsequent test performance, but participants' JOLs in the second cycle were biased by their previous test performance (anchors), and therefore the JOLs were insufficient to reflect the true level of recall improvement.

Scheck et al. (2004) further proposed that anchoring can account for the insensitivity of metamemory judgments to recall difficulty. They found that both participants' recall and JOLs increased as recall became easier, but that JOLs increased more gradually than did recall. Scheck et al. proposed that participants might base their JOLs on their beliefs about the percentage of items they can usually remember (e.g., 50%; M. Carroll, Nelson, & Kirwan, 1997), and this “anchor” then biased their JOLs. Anchoring has also been proposed as a possible cause of other metamemory illusions (e.g., Dunlosky & Matvey, 2001; Yang et al., 2017a).

The anchoring effect proposed as an explanation of the underconfidence-with-practice effect makes reference to an internal performance estimate as the anchor, whereas the effect proposed as an explanation of the insensitivity of metamemory judgments to recall difficulty makes reference to people's beliefs about the percentage of items they can usually remember as the self-generated anchor. Establishing a causal influence of these anchors in judgments is difficult, because in the experimental procedures the experimenter cannot directly manipulate the anchor value. Only a handful of studies have explored the possibility that external numerical anchors might causally influence metamemory judgments.

Zhao and Linderholm (2011) and Zhao (2012) asked two groups of participants to study some texts. The participants then received an anchor instruction. In a high-anchor group, participants were told that “Based on a previous study, college students' test performance ranged from 90% to 100%, with a mean of 95%.” In a low-anchor group, in contrast, participants were told that “Based on a previous study, college students' test performance ranged from 50% to 60%, with a mean of 55%.” Then participants were asked to make a JOL to predict their future test performance. Zhao and colleagues found that the JOLs in the high-anchor group were on average higher than those in the low-anchor group, but there was no difference in recall between the groups. A similar finding was obtained by England and Serra (2012) using word pair stimuli and with JOLs made on each pair. Overall, these three studies clearly showed anchoring effects in metamemory monitoring. However, one important question has yet to be explored.

Motivation of the current research

All three previous studies (England & Serra, 2012; Zhao, 2012; Zhao & Linderholm, 2011) investigated the effect of *informative* anchors in metamemory monitoring. Prior to making their JOLs, participants were explicitly told about past peer performance. The fact that past peer performance (high/low) affects JOLs can be explained in terms of rational inference: If the experimenter is informing the learner that in a relevant cohort of other participants, certain materials proved easy or difficult to learn, then any rational (Bayesian) learner would incorporate that information into their JOLs. Previous studies have not asked whether *uninformative* anchors, which a rational judge should ignore, can similarly influence people's metamemory monitoring. Therefore, the primary aim of the present research is to explore this question.

Would we expect metamemory monitoring to be susceptible to uninformative anchors? A voluminous literature has established that the anchoring effect is a pervasive phenomenon across judgment domains (Furnham & Boo, 2011). Intuitively, we would thus expect that metamemory monitoring would be susceptible to anchoring. Indeed, the scale distortion theory of the anchoring effect straightforwardly predicts that JOLs will be biased by uninformative anchors: The anchors should distort the uses of the response scale, and JOLs should thus assimilate to them. In addition, the selective-accessibility theory of the anchoring effect and the cue-utilization theory of metamemory monitoring may combine to predict an anchoring effect: Uninformative anchors might induce people to search for cues to test whether the anchor is the possible target property (selective accessibility), and then people might use these cues to form their JOLs (cue utilization).

However, two metamemory theories (the dual-process and analytic-processing theories) predict no such effect. First, uninformative anchors cannot affect processing fluency, because they are always presented after the target item. Second, there is little reason to expect that people would have a priori beliefs that uninformative anchors correlate with memory, or would develop such beliefs online across a study phase, because, in our Experiments 3B and 4, the anchors were randomly selected by the participants and completely uninformative. Therefore, both the dual-process and analytic-processing theories predict no effect of uninformative anchors on JOLs. Thus, the theories reviewed above make different predictions about the effects of uninformative anchors on metamemory monitoring. Although the present research was designed to empirically explore the anchoring effect on metamemory monitoring, it also has theoretical implications that we will return to in the General Discussion.

Experiment 1

Experiment 1 was designed to conceptually replicate the effect of informative anchors on metamemory monitoring (England & Serra, 2012; Zhao, 2012; Zhao & Linderholm, 2011).

Method

Participants Given that no previous research had explored the anchoring effect in metamemory monitoring in a within-subjects design, we assumed that the difference in JOLs between high- and low-anchor pairs would be a medium-sized (Cohen's $d = 0.50$) effect. On the basis of this estimate, we computed the required sample size using the G*Power program (Faul, Erdfelder, Lang, & Buchner, 2007), which showed that the required sample size to observe a significant ($\alpha = .05$) difference in JOLs at .8 power would be 35. Forty participants with an average age of 20.18 years ($SD = 1.38$; 31 females, nine males) were recruited from the Fuqing Branch of Fujian Normal University. They took part in the experiment for course credit.

Materials, design, and procedure The stimuli were 50 weakly associated Chinese word pairs developed by Hu et al. (2016). These 50 pairs were separated into five sets, one for each block of the study phase. The assignment of sets to each block was counterbalanced across participants using a Latin square design. In each block, for each participant, five pairs were randomly selected by the computer as high-anchor pairs, and the remainder as low-anchor pairs.

The experiment employed a within-subjects design (anchor: low/high). Participants were informed that after studying each pair they would be asked to predict the likelihood that they would remember that pair in 5 min. In addition, they were told that, to improve their prediction accuracy, the computer would tell them what proportion of individuals were able to remember that pair in our previous studies. Participants were informed that they would be asked to recall the second (target) word in response to the first (cue) word at the final test.

The main task consisted of three phases: *encoding*, *distraction*, and *test*. In the encoding phase, each of the ten pairs in each block was randomly presented, one by one, for 5 s. Blocks were not marked in any way to participants. After studying each pair, that word pair disappeared and an anchor instruction was shown at the upper center of the screen. For the high-anchor pairs, the anchor instruction was "According to our previous studies, about 90% of individuals could remember this pair in 5 min." For the low-anchor pairs, the anchor instruction was "According to our previous studies, about 10% of individuals could remember this pair in 5 min." Participants were asked to make a JOL to predict the likelihood that they would remember that pair in 5 min. The JOLs were made on a slider displayed at the middle of the screen,

labeled from 0 (*I'm sure I'll not remember it*) to 100 (*I'll definitely remember it*), and responses were self-paced. Following the encoding phase, a distraction task (solving math problems for 5 min) was administered. Finally, participants took a cued recall test on all 50 pairs. In the test, all 50 cue words were randomly presented, one at a time. Participants attempted to recall the corresponding targets and enter them via the keyboard. There was no feedback in the test.

Results

Figure 1 depicts the JOLs and recall for low- and high-anchor pairs. We found no difference in final recall between the high-anchor ($M = 40.3\%$, $SD = 19.84$) and low-anchor ($M = 38.9\%$, $SD = 20.01$) pairs, percentage difference = 1.3%, 95% confidence interval (CI) = $[-4.19, 6.89]$, Cohen's $d = 0.08$ (Faul et al., 2007; Morris & DeShon, 2002). In contrast, higher JOLs were given to high-anchor ($M = 61.3$, $SD = 19.45$) than to low-anchor ($M = 54.1$, $SD = 19.31$) pairs, difference = 7.25, 95% CI = $[2.83, 11.67]$, Cohen's $d = 0.52$, indicating an effect of informative anchors on metamemory monitoring. Twenty-nine participants gave higher JOLs to high-anchor pairs, and 11 showed the reverse pattern. We calculated a Gamma correlation value between JOLs and recall for each participant, and used Fisher's Z transformation in this and all subsequent experiments to calculate an estimate of the mean correlation (Silver & Dunlap, 1987). A significantly positive correlation was apparent between JOLs and recall (see Table 1). Thus, over and above the effect of the anchors, participants were able to some extent to accurately report their likelihood of correct recall.

Discussion

Informative anchors exerted a significant effect on metamemory monitoring when the anchors were manipulated within subjects. In contrast, the anchors did not affect recall itself. The latter result reveals a sharp dissociation between metamemory judgments and memory. It is important to emphasize that an anchoring effect on recall could easily be anticipated. Suppose that a participant, shown a low anchor, makes a correspondingly low JOL. On the basis of that judgment, she could then allocate more attention and effort to rehearsing that word pair than she would to a word pair associated with a high anchor, leading to better final recall for the low-anchor items. We explored the relationship between metamemory monitoring and control in Experiment 4.

Experiment 2

In Experiment 1, we found an effect of informative anchors on metamemory monitoring. As was discussed above, rational

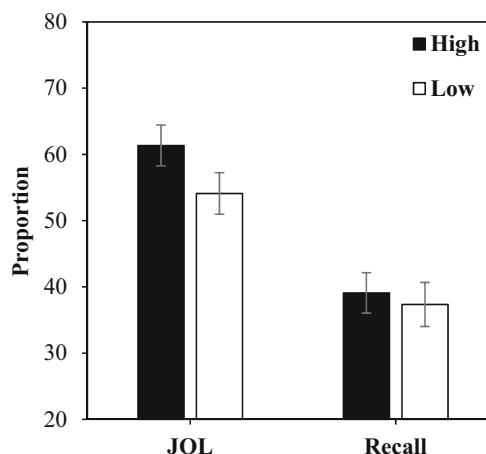


Fig. 1 Experiment 1: JOLs and recall for high- and low-anchor pairs. Error bars represent ± 1 standard error

participants might explicitly integrate the value of an informative anchor (e.g., past peer performance) with their judgment about the difficulty of remembering the current word pair to generate their JOL. In Experiment 2, we explored the effect of uninformative anchors on metamemory monitoring, where such integration would not be rationally justifiable in this way. The classic paradigm for exploring the effect of uninformative anchors on judgments involves a comparative question (i.e., whether a target value is greater or smaller than an anchor value) prior to an absolute estimate of the target value. For example, Jacowitz and Kahneman (1995) asked participants whether the length of the Mississippi River is longer or shorter than 2,000 (high anchor) or 70 (low anchor) miles before asking them to estimate the length of the Mississippi River. Following this method, we employed a comparative question to explore the effect of uninformative anchors in metamemory monitoring.

Table 1 Gamma correlations across experiments

	Correlation Values	95% CI
Experiment 1		
JOLs \times Recall	.22	[.15, .30]
Experiment 2		
JOLs \times Recall	.24	[.14, .33]
Experiment 3A		
JOLs \times Recall	.27	[.18, .36]
Experiment 3B		
JOLs \times Recall	.23	[.12, .34]
Experiment 4		
JOLs \times Recall	.30	[.20, .39]
JOLs \times Restudy choices	-.70	[-.80, -.58]
Restudy choices \times Recall	-.39	[-.55, -.21]

Method

Participants The sample size was determined according to a power analysis based on the results of Experiment 1. The sample size to observe a significant ($\alpha = .05$) difference in JOLs between high- and low-anchor pairs at .8 power was 32 participants. The data from two participants were lost due to a computer failure, leaving the data from 30 participants. All participants were undergraduates from the Fuqing Branch of Fujian Normal University, with an average age of 19.93 years ($SD = 1.74$; 21 females, nine males). They took part for course credit.

Materials, design, and procedure Forty-eight word pairs from Experiment 1 were employed, randomly divided into four sets, each comprising 12 pairs. The assignment of sets to each block was counterbalanced across participants. Again, anchor (low/high) was manipulated within subjects. In comparison to Experiment 1, we changed the anchor values to decrease their informativeness. The low anchors consisted of the numbers 10%, 20%, and 30%, whereas the high anchors consisted of the numbers 70%, 80%, and 90%. In each block, for each participant, two of the 12 word pairs were randomly selected and were assigned to each anchor value.

The procedure was the same as in Experiment 1, with the following exceptions. After studying each pair, participants were asked “Is the likelihood you would be able to remember the preceding word pair in 5 min higher or lower than [10%/20%/30%/70%/80%/90%]?” Then they were asked to predict the likelihood that they would be able to remember the pair in 5 min by dragging and clicking a scale ranging from 0 to 100.

Results and discussion

Again, we found a significant anchoring effect in metamemory monitoring. Participants gave significantly higher JOLs to high-anchor ($M = 66.99$, $SD = 16.60$) than to low-anchor ($M = 62.48$, $SD = 20.44$) pairs, difference = 4.51, 95% CI = [1.52, 7.50], Cohen’s $d = 0.56$ (see Fig. 2). Nineteen participants gave higher JOLs to high-anchor pairs, and 11 showed the reverse pattern. Yet there was no difference in recall itself between the high-anchor ($M = 43.1\%$, $SD = 21.42$) and low-anchor ($M = 43.3\%$, $SD = 24.94$) pairs, difference = -0.3% , 95% CI = [-6.80, 6.24], Cohen’s $d = -0.02$. Moreover we found a positive correlation between JOLs and recall (see Table 1), revealing some evidence of metacognitive accuracy aside from the anchoring effect.

Experiment 3A

In Experiment 2, a significant effect of uninformative anchors emerged in metamemory monitoring. However, although the

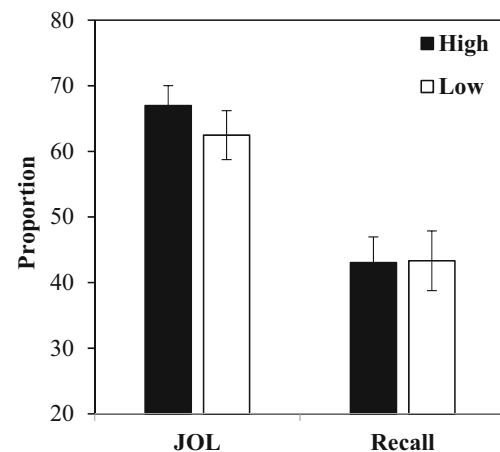


Fig. 2 Experiment 2: JOLs and recall for high- and low-anchor pairs. Error bars represent ± 1 standard error

comparative questions simply presented a number against which participants compared their estimate, it is possible that they were interpreted as informative. For example, participants might think that a low anchor implied that the preceding pair was a difficult one.

Previous studies exploring the effect of uninformative anchors have asked participants to randomly select anchor values themselves to emphasize their lack of relevance to the judgment question. For instance, Tversky and Kahneman (1974) asked participants to randomly select a number (65 or 10) by spinning a wheel of fortune, and then asked participants to judge whether the proportion of African countries in the UN is higher or lower than 65% or 10%. Following Tversky and Kahneman, in Experiment 3A we instructed participants to randomly select a number and then to report whether the likelihood they would remember a given word pair was higher or lower than the number they selected. In this way, we explored whether uninformative anchors have an effect on metamemory monitoring when people believe the anchors are randomly selected. Chapman and Johnson (1999) noted that the anchors might be interpreted as nonrandom and hence informative in Tversky and Kahneman’s study, because there were only two numbers (65 and 10) on the fortune wheel. We therefore employed nine anchor values (10%, 20%, 30%, . . . , 90%).

Method

Participants The sample size was determined according to a power analysis of the results of Experiment 2. The sample size to observe a significant ($\alpha = .05$) difference in JOLs between high- and low-anchor pairs at .8 power was 28 participants. In total, we recruited 28 participants, with an average age of 20.89 years ($SD = 1.20$; 22 females, six males) from the Fuqing Branch of Fujian Normal University. They participated for course credit.

Materials, design, and procedure Forty-five word pairs from the previous experiments were chosen and randomly divided into five sets, each consisting of nine pairs. The assignment of sets to each block was counterbalanced across participants. For each participant in each block, a pair was randomly assigned to an anchor. The low anchors were 10%, 20%, and 30%; the medium anchors were 40%, 50%, and 60%; and the high anchors were 70%, 80%, and 90%.

The procedure was the same as in Experiment 2, with the following exceptions. After participants had studied each pair, a box appeared on the screen containing rapidly changing numbers (10%, 20%, . . . 90%), randomly generated and displayed superimposed for 20 ms each. Participants were asked to press the “b” key to select a number. After the keypress, the anchor number assigned to that pair was shown on the screen, and participants were asked “Is the likelihood you would remember the preceding word pair in 5 min higher or lower than [10%/20%/30%/ . . . 90%]?” Then the participants were asked to predict the likelihood that they would be able to remember the preceding pair in 5 min, by dragging and clicking a scale ranging from 0 to 100.

We instructed participants to randomly select a number, but in fact, the anchor number was decided by the computer. In this way, we ensured that in each block one word pair was assigned to each anchor value. We hypothesized that because the numbers changed very rapidly (50 per second), participants would not realize that the anchor number was chosen independently of their action.

Results

There was no main effect of uninformative anchors on recall, $F(2, 54) = 0.10, p = .91, \eta_p^2 < .01$, in a repeated measures analysis of variance (ANOVA) with anchor value (low/medium/high) as the factor. In contrast, we found that JOLs linearly decreased across high-, medium-, and low-anchor pairs, $F(1, 27) = 18.87, p < .001, \eta_p^2 = .41$ (see Fig. 3A). Participants gave significantly higher JOLs to high-anchor ($M = 55.80, SD = 16.47$) than to low-anchor ($M = 49.96, SD = 17.46$) pairs, difference = 5.83, 95% CI = [3.08, 8.59], Cohen’s $d = 0.82$. Twenty participants gave higher JOLs to high-anchor pairs, and eight showed the reverse pattern. Participants also gave higher JOLs to high- than to medium-anchor pairs ($M = 52.44, SD = 16.63$), difference = 3.35, 95% CI = [0.73, 5.97], Cohen’s $d = 0.50$, and marginally higher JOLs to medium- than to low-anchor pairs, difference = 2.48, 95% CI = [−0.24, 5.20], Cohen’s $d = 0.35$. To further explore the anchoring effect on metamemory monitoring, we ran a repeated measures ANOVA with anchor value (10%, 20%, . . . , 90%) as a within-subjects variable. The results indicated that JOLs linearly increased from low- to high-anchor pairs, $F(1, 27) = 16.85, p < .001, \eta_p^2 = .38$ (see Fig. 3B). We again found a positive correlation between JOLs and recall (see Table 1).

Experiment 3B

In Experiment 3A, the anchor assignment was randomly determined by the computer rather than by the participant. Although the pace at which numbers changed (20 ms each) was very fast, it is conceivable that participants suspected that the anchors—as a result of being selected by the computer rather than by themselves—were in fact nonrandom, contrary to our intentions. Therefore, in Experiment 3B we made a further change to reduce participants’ suspicion: The anchor numbers were truly selected by the participants.

To establish that the selected numbers continued to be random, we carried out a procedural check. Ten University College London (UCL) participants were recruited to determine whether they could select an anchor number under intentional instructions to do so. Each participant performed 45 trials. On each trial, constantly changing numbers (10%, 20%, . . . , 90%) were randomly generated by the computer and presented on the display for 20 ms each, in the same manner as in Experiment 3A. Participants were encouraged to do their best to press the “b” key to select only the target number 50%. The results showed no difference between the average frequency that participants successfully caught 50% ($M = 5.00, SD = 2.54$) and the level expected by chance (5). Hence, this procedure ensured that the anchor was both truly chosen by the participant and random.

Participants On the basis of the data from Experiment 3A, we would need 26 participants in order to observe a significant anchoring effect ($\alpha = .05$) on JOLs at .8 power. The data from two participants were lost due to computer failure, leaving the data from 24 participants. All participants were undergraduates from the Fuqing Branch of Fujian Normal University, with an average age of 20.75 years ($SD = 0.85$; 20 female, four male). They participated for course credit.

Materials, design, and procedure The materials, design, and procedure were the same as in Experiment 3A, with one exception. After participants pressed the “b” key, the number they selected was shown at the middle of the screen, and participants were asked whether or not the likelihood they would be able to remember the preceding pair was higher or lower than the number they had selected.

Results

Consistent with Experiment 3A, we observed no main effect of uninformative anchors in recall, $F(2, 46) = 0.04, p = .96, \eta_p^2 < .01$, but the JOLs linearly decreased across high-, medium-, and low-anchor pairs, $F(1, 23) = 5.29, p = .03, \eta_p^2 = .19$ (see

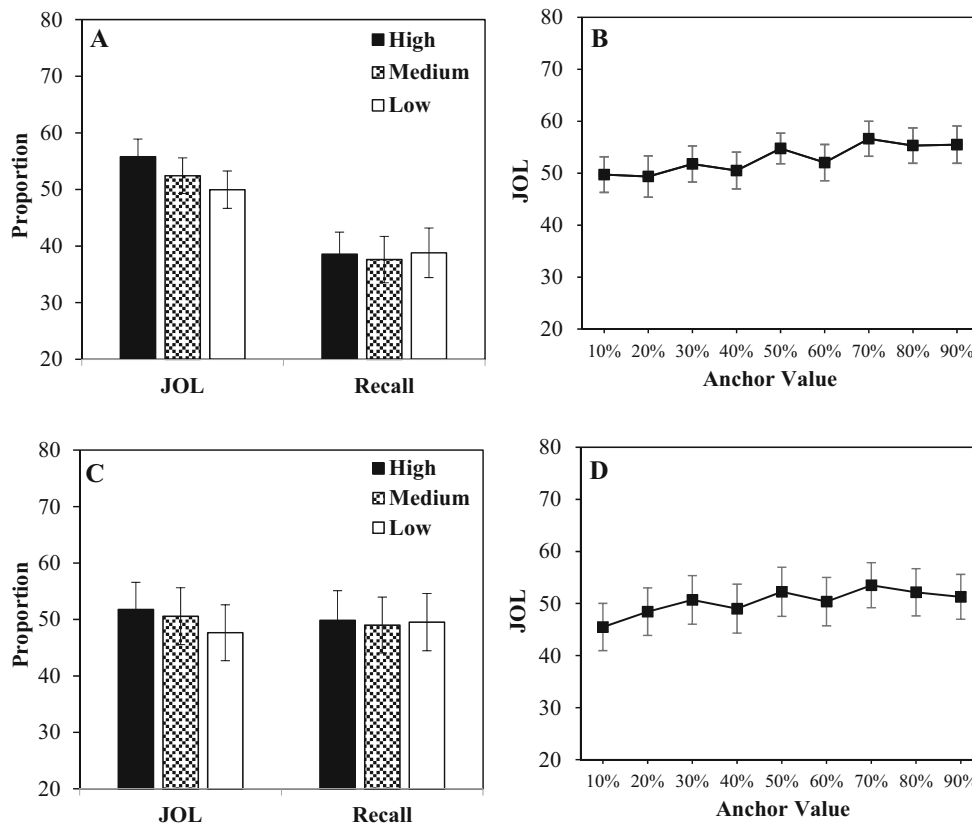


Fig. 3 Experiment 3: (A) JOLs and recall for high-, medium-, and low-anchor pairs in Experiment 3A. (B) JOLs across different anchor values in Experiment 3A. (C) JOLs and recall for high-, medium-, and low-anchor

pairs in Experiment 3B. (D) JOLs across different anchor values in Experiment 3B. Error bars represent ± 1 standard error

Fig. 3C). Participants gave significantly higher JOLs to high-anchor ($M = 51.78, SD = 23.55$) than to low-anchor ($M = 47.66, SD = 24.27$) pairs, difference = 4.13, 95% CI = [0.42, 7.83], Cohen’s $d = 0.47$. Sixteen participants gave higher JOLs to high-anchor pairs, and eight showed the reverse pattern. Participants also gave numerically higher JOLs to high- than to medium-anchor pairs ($M = 50.58, SD = 24.70$), difference = 1.20, 95% CI = [− 1.83, 4.23], Cohen’s $d = 0.28$, and marginally higher JOLs to high- than to low-anchor pairs, difference = 2.93, 95% CI = [− 0.11, 5.96], Cohen’s $d = 0.41$. A repeated measures ANOVA with all anchor values as a within-subjects variable revealed that JOLs increased linearly from low- to high-anchor pairs, $F(1, 23) = 8.63, p = .007, \eta_p^2 = .27$ (see Fig. 3D). Again we found a positive correlation between JOLs and recall (see Table 1).

Discussion

Both Experiments 3A and 3B clearly revealed significant effects of uninformative anchors on metamemory monitoring, regardless of whether or not the anchor values were truly selected by the participants.

Experiment 4

Across Experiments 2–3B we obtained effects of uninformative anchors on metamemory monitoring. From a theoretical perspective, these effects could be explained by scale distortion (the anchors would bias the psychological interpretation of the JOL scale), or alternatively, the selective-accessibility and cue-utilization theories might combine to explain them (anchors activate cues, which in turn form the basis for JOLs).

In Experiment 4, we asked whether anchoring has an effect on metamemory control (restudy choices). As was noted by Brewer, Chapman, Schwartz, and Bergus (2007), although the anchoring effect in metacognitive judgments has been established as a robust phenomenon, surprisingly little research has explored the effect in decisions (action choices). Brewer et al. conducted the first study to explore this issue. They asked HIV-positive patients to imagine that they were having sex with an HIV-negative partner, and that the condom broke during intercourse. The patients were then asked to answer either a high- or a low-anchor question: Would the risk of infection be higher or lower than 99% (high anchor) or 1% (low anchor)? (Note that these may have been interpreted as informative anchors.) All patients were then asked to make

their best estimates of the risk. Following the estimation, they made a choice whether to recommend that their partner take a treatment (e.g., see a doctor, have an HIV test) or not (e.g., do nothing, wait and see). Brewer et al.'s results showed a significant anchoring effect in the estimates of the risk of transmission: Patients in the high-anchor group reported significantly higher estimates (64%) than did those in the low-anchor group (43%). However, anchoring had no effect on the treatment choices: There was no difference in treatment choices between the high- and low-anchor groups. Brewer and colleagues suggested that their results supported the scale distortion theory (i.e., that anchors affect the uses of the response scale but do not change the mental representation of the target property).

Following Brewer et al. (2007), in Experiment 4 we asked participants to answer a comparative (anchor) question, then to make a JOL, and then to make a restudy choice. We used uninformative anchors. Scale distortion theory predicts an anchoring effect on JOLs but no effect on restudy choices, because anchors (10%–90%) and restudy choices (YES/NO) are expressed on different scales (Frederick & Mochon, 2012). Another aim of Experiment 4 was to conceptually replicate Experiment 3B's findings with a UK rather than a Chinese sample. Experiment 4 was fully preregistered at the Open Science Framework (OSF; <https://osf.io/d73a5/registrations/>).

Method

Participants The sample size was determined according to the results of Experiment 3B. To observe a significant anchoring effect in JOLs at .8 power, we required at least 31 participants. Therefore, 31 participants, with an average age of 22.26 years ($SD = 6.03$; 21 females, ten males) were recruited from the UCL participant pool. They received £3 or course credit as compensation.

Materials, design, and procedure Forty-five weakly associated English word pairs (e.g., *family–doctor*) were selected from the South Florida Free Association norms developed by D. L. Nelson, McEvoy, and Schreiber (1998). Their forward semantic association strength ranged from .010 to .020, and the backward strength was 0.

The design and procedure were the same as in Experiment 3B, with one exception. After making each JOL, participants were asked whether they wanted to restudy the preceding pair after seeing all 45 pairs. They were told that they would be able to restudy the pair after they had seen all 45 word pairs if they selected *YES*. If they selected *NO*, the pair would not be shown again. In fact, no pairs were restudied prior to the final recall test, regardless of participants' restudy choices.

Results and discussion

The study proceeded in accordance with the preregistration plan. Figure 4A depicts participants' JOLs, restudy choices, and recall for the low-, medium-, and high-anchor pairs. We observed no main effect of uninformative anchors on recall, $F(2, 60) = 0.34, p = .72, \eta_p^2 = .01$, but JOLs linearly decreased across high-, medium-, and low-anchor pairs, $F(1, 30) = 9.91, p = .004, \eta_p^2 = .25$. Participants gave higher JOLs to high-anchor ($M = 57.93, SD = 12.32$) than to low-anchor ($M = 52.26, SD = 15.06$) pairs, difference = 5.67, 95% CI = [1.99, 9.35], Cohen's $d = 0.57$. Twenty-two participants gave higher JOLs to high- than to low-anchor pairs, and nine showed the reverse pattern. Participants also gave numerically higher JOLs to high- than to medium-anchor pairs ($M = 56.50, SD = 12.10$), difference = 1.43, 95% CI = [− 1.17, 4.03], Cohen's $d = 0.20$, as well as higher JOLs to medium- than to low-anchor pairs, difference = 4.25, 95% CI = [0.96, 7.53], Cohen's $d = 0.47$.

What was the effect of the anchors on restudy choices—the key novel question addressed in this experiment? These linearly increased across high-, medium-, and low-anchor pairs, $F(1, 30) = 13.20, p = .001, \eta_p^2 = .31$. Participants selected a higher proportion of low-anchor ($M = 40.9, SD = 22.46$) than of high-anchor ($M = 31.6, SD = 23.14$) pairs for restudy, difference = 9.2%, 95% CI = [4.04, 14.13], Cohen's $d = 0.65$. Twenty-three participants showed this pattern, and five the reverse pattern (two participants did not select any pairs for restudy, and one participant selected all pairs for restudy). Participants also selected a numerically higher proportion of low- than of medium-anchor pairs ($M = 34.6, SD = 24.91$) for restudy, difference = 6.3, 95% CI = [− 0.33, 12.89], Cohen's $d = 0.35$, and a numerically higher proportion of medium- than of high-anchor pairs, difference = 3.0, 95% CI = [− 2.39, 8.30], Cohen's $d = 0.20$.

To further explore the effect of uninformative anchors in metamemory monitoring, we ran a repeated measures ANOVA with all anchor values as the independent variable and with JOLs as the dependent variable. The results indicated that JOLs increased linearly from low- to high-anchor pairs, $F(1, 30) = 10.95, p = .002, \eta_p^2 = .27$ (see Fig. 4B). Next, for each participant and each anchor value, we calculated the proportions of pairs that were selected for restudy (see Fig. 4C), and ran a repeated measures ANOVA with all anchor values as the independent variable and restudy choices as the dependent variable (this analysis was not part of the preregistration). The results indicated that restudy choices decreased linearly from low- to high-anchor pairs, $F(1, 30) = 7.58, p = .01, \eta_p^2 = .20$ (see Fig. 4C).

Next, we conducted a multilevel mediation analysis to explore whether the anchoring effect in metamemory control (restudy choices) is mediated by its effect on metamemory monitoring (JOLs) using the `blm` R program (Vuorre &

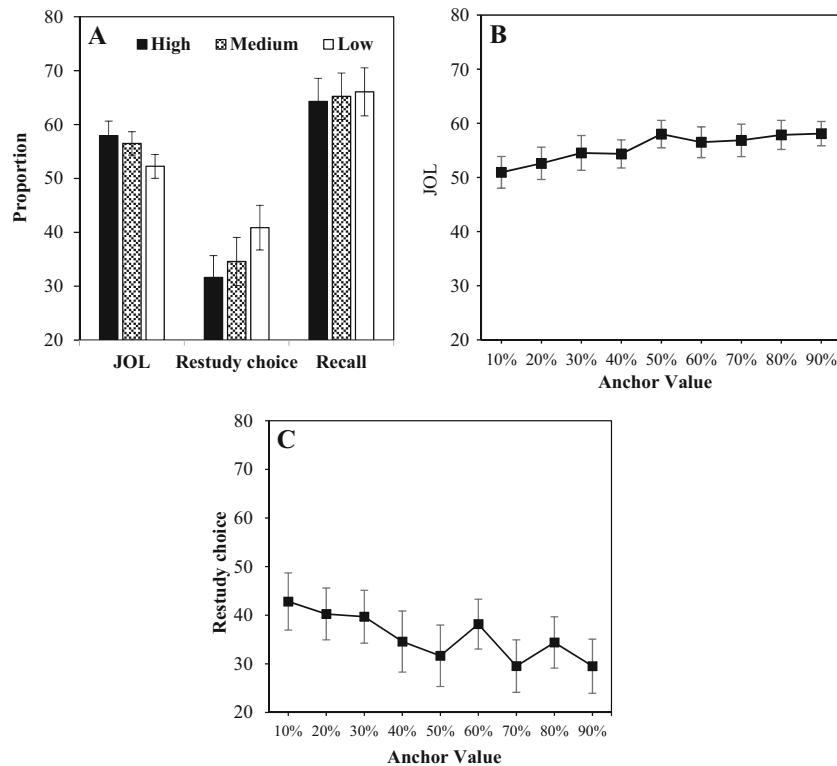


Fig. 4 Experiment 4: (A) JOLs, restudy choices, and recall for high-, medium-, and low-anchor pairs. (B) JOLs across different anchor values. (C) Restudy choices across different anchor values. Error bars represent ± 1 standard error

Bolger, 2017) (this analysis was also not part of the preregistration). The results showed that JOLs significantly mediated the anchoring effect on restudy choices, mediation effect parameter = -0.09 , 95% CI = $[-0.16, -0.03]$ (for the detailed results of the mediation analysis, see Table 2). The anchoring effect on metamemory control is at least partly mediated by its effect on metamemory monitoring. We found a significant positive correlation between JOLs and recall, a significant negative correlation between JOLs and restudy choices, and a significant negative correlation between restudy choices and recall (see Table 1).

General discussion

In the present research, we first explored the effect of informative anchors on metamemory monitoring. In Experiment 1, informing participants of past peer performance in a within-subjects design biased their subsequent JOLs. More importantly, we found a significant effect of uninformative anchors on metamemory monitoring across Experiments 2–4, even when the anchors were self-selected to emphasize their uninformativity (Exps. 3B and 4). Furthermore, in Experiment 4, we explored the anchoring effect in metamemory control. A higher proportion of low-anchor than of high-anchor pairs was selected for restudy. These results indicate that metamemory control is more closely related to

metamemory monitoring than it is to actual memory status (Metcalf & Finn, 2008; Yang et al., 2017a): Participants chose to restudy items for which their JOLs were biased downward by a low anchor, even though the anchor had no effect on their actual likelihood of recalling the item.

We observed significant anchoring effects on metamemory monitoring but no significant effect on recall across all five experiments. To confirm this pattern, we conducted a repeated measures multivariate ANOVA with data type (JOLs vs. recall) and anchor as within-subjects variables. To maximize the power to observe a significant interaction, we collapsed all high-anchor and low-anchor items (JOLs and recall) across the five experiments. The results for the medium-anchor items were removed from this analysis because Experiments 1 and 2 included no medium anchors. The ANOVA yielded a significant interaction between anchor and data type, $F(1, 152) = 17.29, p < .001, \eta_p^2 = .10$. Paired t tests revealed a significant anchoring effect on JOLs, difference = 5.64, 95% CI = $[4.02, 7.26]$, Cohen’s $d = 0.56$, but no effect on recall, difference = 0.03, 95% CI = $[-2.50, 2.57]$, Cohen’s $d = 0.002$.

Why did uninformative anchors affect metamemory monitoring? The dual-process and analytic-processing theories for metamemory monitoring have difficulty explaining this effect because the anchor, which was presented after each target word pair in our procedure, could not have affected the fluency with which the word pair was processed, and participants were unlikely to hold a priori or develop online beliefs that

Table 2 Multilevel mediation analysis results of Experiment 4

	<i>b</i>	<i>SE</i>	95% CI
Total effect of anchors on restudy choices	– 0.10	0.05	[– 0.21, 0]
Direct effect of anchors on restudy choices	– 0.02	0.04	[– 0.10, 0.07]
Indirect effect of anchors on restudy choices through JOLs	– 0.09	0.03	[– 0.16, – 0.03]

uninformative anchors (randomly selected by themselves) correlated with successful recall.

Our favored explanation for the effect is based on a combination of the selective-accessibility and cue-utilization theories. These two theories dovetail readily to explain the anchoring effect in metamemory monitoring. Imagine that you have been asked to memorize the word pair *pond–frog* in the context of a low anchor (e.g., 10%). The anchor will induce you (according to the selective-accessibility theory) to consider features of this word pair that might confirm that you will only be about 10% likely to remember it on the later test. You might, for instance, try to generate other strongly related words (e.g., *water*, *lake*) and evaluate how likely they are to come to mind instead of *frog* at test, given the cue *pond* (selective accessibility). When asked to make an absolute judgment (JOL), this feature (that the target has many strong and easily accessed competitors) will induce you to give a lower judgment than would have been the case given a high anchor (cue utilization). Future studies could test this proposal more directly. For example, following Mussweiler and Strack (2000), participants could be asked to judge whether a letter string (e.g., *water*) is a word or nonword (lexical decision) immediately after they have studied a word pair (e.g., *pond–frog*) and answered a comparative (anchor) question. The selective-accessibility hypothesis would be supported if participants were to respond faster to cue-related words, like *water*, in the low- than in the high-anchor condition.

These two theories can also be readily applied to the anchoring effect in restudy choices. Low anchors might induce people to generate strongly related competitors and to appreciate how easy it will be at test to recall competitors as the correct response. This would then serve as a cue when a restudy choice is made (i.e., choosing to restudy the pairs associated with strong competitors).

It is important to note that the anchoring effect in restudy choices does not invalidate the scale distortion theory. A defender of this theory could argue that uninformative anchors first affect the use of the response scale and bias JOLs, and then that participants make their restudy decisions according to their own JOLs by selecting low-JOL items for restudy (Finn, 2008; Yang et al., 2017a). Indeed, in Experiment 4, the multilevel mediation analysis revealed that the anchoring effect in restudy choices was significantly mediated by its effect on JOLs. In summary, the selective-accessibility and

cue-utilization theories combine to provide the best explanation of the effects of uninformative anchors in metamemory monitoring and control, but the scale distortion theory remains a viable alternative account. These theories are not mutually exclusive and might yield overlapping anchoring effects. We emphasize, however, that these theoretical inferences are tentative as the primary goal of the present research was to explore whether anchors affect metamemory monitoring.

In the present and all previous studies exploring the anchoring effect in metamemory monitoring (England & Serra, 2012; Zhao, 2012; Zhao & Linderholm, 2011), participants' attention was explicitly directed to the anchors to ensure that they attended to them. It is unknown whether anchors can bias metamemory monitoring when people do not pay attention to anchors. Wilson, Houston, Etling, and Brekke (1996) proposed that anchoring occurs only when people pay attention to the anchors. However, Critcher and Gilovich (2008) claimed that attention is not a necessary prerequisite for anchoring to occur. They asked participants to view images incidentally containing a high or a low anchor. In a high-anchor picture, for instance, a fictitious linebacker wore a jersey with the number 94, whereas in the low-anchor picture, the number was 54. Then all participants estimated how likely it was that the linebacker would register a sack in a football game. Although the participants were not required to attend to or consider the number on the jersey, they gave higher estimates in the high- than in the low-anchor condition (numerical anchoring). Future research could usefully explore whether incidentally encountered numbers can affect metamemory monitoring. For example, following Critcher and Gilovich, participants could be asked to remember some basketball players while viewing their photos one by one. In each photo, a player wears a jersey with a number ranging from 0 to 100. After viewing each photo, participants could make a JOL to predict how confident they were that they could remember that player later. Such studies could explore whether incidentally encountered numbers can bias metamemory monitoring.

Across all five experiments reported here, we explored the anchoring effect in immediate JOLs (i.e., JOLs made immediately following a study phase). Future research could usefully explore the anchoring effect in delayed JOLs (i.e., JOLs made with a time interval between study and judgment). It is well-known that delayed JOLs are more accurate and less susceptible to biases than are immediate ones (T. O. Nelson

& Dunlosky, 1991; Van Overschelde & Nelson, 2006; Yang et al., 2017a).

Many previous studies have explored factors (such as a priori beliefs) that exist before the encoding phase and can affect metamemory monitoring (Hu et al., 2015; Jia et al., 2015; Mueller & Dunlosky, 2017; Mueller et al., 2016; Undorf & Erdfelder, 2015; Witherby & Tauber, 2017). In addition, other studies have shown that encoding-phase variables (e.g., processing fluency) can affect metamemory monitoring as well (Mueller, Tauber, & Dunlosky, 2013; Susser, Panitz, Buchin, & Mulligan, 2017; Undorf et al., 2017; Yang et al., 2017a). Going beyond these previous studies, the present research has shown that variations that occur *following* the encoding phase (but prior to making a JOL) can also affect metamemory monitoring.

Conclusion

In conclusion, the present research has shown that metamemory judgments, similar to many other forms of metacognitive judgment, are not immune to anchoring effects. Not only informative but also uninformative anchors can affect metamemory monitoring. Moreover, anchors also affect metamemory control, through their effects on metamemory monitoring. The selective-activation and cue-utilization theories in combination provide a plausible theoretical explanation of the anchoring effect in metamemory monitoring.

Author note An award from the China Scholarship Council (CSC), awarded to C.Y., supported this research. All data have been made publicly available via the OSF at <https://osf.io/d73a5/>.

References

- Brewer, N. T., Chapman, G. B., Schwartz, J. A., & Bergus, G. R. (2007). The influence of irrelevant anchors on the judgments and choices of doctors and patients. *Medical Decision Making*, *27*, 203–211. <https://doi.org/10.1177/0272989X06298595>
- Carroll, M., Nelson, T. O., & Kirwan, A. (1997). Tradeoff of semantic relatedness and degree of overlearning: Differential effects on metamemory and on long-term retention. *Acta Psychologica*, *95*, 239–253. [https://doi.org/10.1016/S0001-6918\(96\)00040-6](https://doi.org/10.1016/S0001-6918(96)00040-6)
- Carroll, S. R., Petrusic, W. M., & Leth-Steensen, C. (2009). Anchoring effects in the judgment of confidence: Semantic or numeric priming? *Attention, Perception, & Psychophysics*, *71*, 297–307. <https://doi.org/10.3758/APP.71.2.297>
- Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes*, *79*, 115–153. <https://doi.org/10.1006/OBHD.1999.2841>
- Critcher, C. R., & Gilovich, T. (2008). Incidental environmental anchors. *Journal of Behavioral Decision Making*, *21*, 241–251. <https://doi.org/10.1002/bdm.586>
- Dunlosky, J., & Matvey, G. (2001). Empirical analysis of the intrinsic–extrinsic distinction of judgments of learning (JOLs): Effects of relatedness and serial position on JOLs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1180–1191. <https://doi.org/10.1037/0278-7393.27.5.1180>
- England, B. D., & Serra, M. J. (2012). The contributions of anchoring and past-test performance to the underconfidence-with-practice effect. *Psychonomic Bulletin & Review*, *19*, 715–722. <https://doi.org/10.3758/s13423-012-0237-7>
- Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin*, *32*, 188–200. <https://doi.org/10.1177/0146167205282152>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. <https://doi.org/10.3758/BF03193146>
- Finn, B. (2008). Framing effects on metacognitive monitoring and control. *Memory & Cognition*, *36*, 813–821. <https://doi.org/10.3758/mc.36.4.813>
- Finn, B., & Metcalfe, J. (2007). The role of memory for past test in the underconfidence with practice effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 238–244. <https://doi.org/10.1037/0278-7393.33.1.238>
- Frederick, S. W., & Mochon, D. (2012). A scale distortion theory of anchoring. *Journal of Experimental Psychology: General*, *141*, 124–133. <https://doi.org/10.1037/a0024006>
- Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *Journal of Socio-Economics*, *40*, 35–42. <https://doi.org/10.1016/j.socec.2010.10.008>
- Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology*, *81*, 657–669. <https://doi.org/10.1037/0022-3514.81.4.657>
- Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*, 208–216. <https://doi.org/10.1037/h0022263>
- Hu, X., Li, T., Zheng, J., Su, N., Liu, Z., & Luo, L. (2015). How much do metamemory beliefs contribute to the font-size effect in judgments of learning? *PLoS ONE*, *10*, e0142351. <https://doi.org/10.1371/journal.pone.0142351>
- Hu, X., Liu, Z., Li, T., & Luo, L. (2016). Influence of cue word perceptual information on metamemory accuracy in judgement of learning. *Memory*, *24*, 383–398. <https://doi.org/10.1080/09658211.2015.1009470>
- Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, *21*, 1161–1166. <https://doi.org/10.1177/01461672952111004>
- Jia, X., Li, P., Li, X., Zhang, Y., Cao, W., Cao, L., & Li, W. (2015). The effect of word frequency on judgments of learning: Contributions of beliefs and processing fluency. *Frontiers in Psychology*, *6*. <https://doi.org/10.3389/fpsyg.2015.01995>
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*, 349–370. <https://doi.org/10.1037/0096-3445.126.4.349>
- Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General*, *131*, 147–162. <https://doi.org/10.1037/0096-3445.131.2.147>
- Kornell, N., & Bjork, R. A. (2008a). Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science*, *19*, 585–592. <https://doi.org/10.1111/j.1467-9280.2008.02127.x>

- Kornell, N., & Bjork, R. A. (2008b). Optimising self-regulated study: The benefits—and costs—of dropping flashcards. *Memory*, *16*, 125–136. <https://doi.org/10.1080/09658210701763899>
- Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*, 174–179. <https://doi.org/10.3758/pbr.15.1.174>
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*, 105–125. <https://doi.org/10.1037/1082-989x.7.1.105>
- Mueller, M. L., & Dunlosky, J. (2017). How beliefs can impact judgments of learning: Evaluating analytic processing theory with beliefs about fluency. *Journal of Memory and Language*, *93*, 245–258. <https://doi.org/10.1016/j.jml.2016.10.008>
- Mueller, M. L., Dunlosky, J., & Tauber, S. K. (2016). The effect of identical word pairs on people's metamemory judgments: What are the contributions of processing fluency and beliefs about memory? *The Quarterly Journal of Experimental Psychology*, *69*, 781–799. <https://doi.org/10.1080/17470218.2015.1058404>
- Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning: Does it exemplify fluency effects or reflect people's beliefs about memory? *Journal of Memory and Language*, *70*, 1–12. <https://doi.org/10.1016/j.jml.2013.09.007>
- Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2013). Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review*, *20*, 378–384. <https://doi.org/10.3758/s13423-012-0343-6>
- Mussweiler, T., & Strack, F. (2000). The use of category and exemplar knowledge in the solution of anchoring tasks. *Journal of Personality and Social Psychology*, *78*, 1038–1052. <https://doi.org/10.1037/0022-3514.78.6.1038>
- Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The delayed-JOL effect. *Psychological Science*, *2*, 267–270. <https://doi.org/10.1111/j.1467-9280.1991.tb00147.x>
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida free association, rhyme, and word fragment norms [Database]. Available from <http://w3.usf.edu/FreeAssociation/>
- Newell, B. R., & Shanks, D. R. (2014). Prime numbers: Anchoring and its implications for theories of behavior priming. *Social Cognition*, *32*, 88–108. <https://doi.org/10.1521/soco.2014.32.sup.88>
- Rhodes, M. G. (2016). Judgments of learning: Methods, data, and theory. In J. Dunlosky & S. K. Tauber (Eds.), *The Oxford handbook of metamemory* (pp. 65–80). New York: Oxford University Press.
- Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*, 615–625. <https://doi.org/10.1037/a0013684>
- Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review*, *16*, 550–554. <https://doi.org/10.3758/PBR.16.3.550>
- Scheck, P., Meeter, M., & Nelson, T. O. (2004). Anchoring effects in the absolute accuracy of immediate versus delayed judgments of learning. *Journal of Memory and Language*, *51*, 71–79. <https://doi.org/10.1016/j.jml.2004.03.004>
- Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's z transformation be used? *Journal of Applied Psychology*, *72*, 146–148. <https://doi.org/10.1037/0021-9010.72.1.146>
- Simonson, I., & Drolet, A. (2004). Anchoring effects on consumers' willingness-to-pay and willingness-to-accept. *Journal of Consumer Research*, *31*, 681–690. <https://doi.org/10.1086/425103>
- Slovic, P. (1967). The relative influence of probabilities and payoffs upon perceived risk of a gamble. *Psychonomic Science*, *9*, 223–224. <https://doi.org/10.3758/BF03330840>
- Soderstrom, N. C., & Bjork, R. A. (2014). Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language*, *73*, 99–115. <https://doi.org/10.1016/j.jml.2014.03.003>
- Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, *73*, 437–446. <https://doi.org/10.1037/0022-3514.73.3.437>
- Susser, J. A., Jin, A., & Mulligan, N. W. (2016). Identity priming consistently affects perceptual fluency but only affects metamemory when primes are obvious. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 657–662. <https://doi.org/10.1037/xlm0000189>
- Susser, J. A., Panitz, J., Buchin, Z., & Mulligan, N. W. (2017). The motoric fluency effect on metamemory. *Journal of Memory and Language*, *95*, 116–123. <https://doi.org/10.1016/j.jml.2017.03.002>
- Tauber, S. K., Dunlosky, J., Rawson, K. A., Wahlheim, C. N., & Jacoby, L. L. (2013). Self-regulated learning of a natural category: Do people interleave or block exemplars during study? *Psychonomic Bulletin & Review*, *20*, 356–363. <https://doi.org/10.3758/s13423-012-0319-6>
- Thorsteinson, T. J., Breier, J., Atwell, A., Hamilton, C., & Privette, M. (2008). Anchoring effects on performance judgments. *Organizational Behavior and Human Decision Processes*, *107*, 29–40. <https://doi.org/10.1016/j.obhdp.2008.01.003>
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. <https://doi.org/10.1017/CBO9780511809477.002>
- Undorf, M., & Erdfelder, E. (2015). The relatedness effect on judgments of learning: A closer look at the contribution of processing fluency. *Memory & Cognition*, *43*, 647–658. <https://doi.org/10.3758/s13421-014-0479-x>
- Undorf, M., Zimdahl, M. F., & Bernstein, D. M. (2017). Perceptual fluency contributes to effects of stimulus size on judgments of learning. *Journal of Memory and Language*, *92*, 293–304. <https://doi.org/10.1016/j.jml.2016.07.003>
- Van Overschelde, J. P., & Nelson, T. O. (2006). Delayed judgments of learning cause both a decrease in absolute accuracy (calibration) and an increase in relative accuracy (resolution). *Memory & Cognition*, *34*, 1527–1538. <https://doi.org/10.3758/BF03195916>
- Vuorre, M., & Bolger, N. (2017). *Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience*. Unpublished manuscript. Retrieved from <http://osf.io/s48e2>
- Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, *125*, 387–402. <https://doi.org/10.1037/0096-3445.125.4.387>
- Witherby, A. E., & Tauber, S. K. (2017). The concreteness effect on judgments of learning: Evaluating the contributions of fluency and beliefs. *Memory & Cognition*, *45*, 639–650. <https://doi.org/10.3758/s13421-016-0681-0>
- Yan, V. X., Bjork, E. L., & Bjork, R. A. (2016). On the difficulty of mending metacognitive illusions: A priori theories, fluency effects, and misattributions of the interleaving benefit. *Journal of Experimental Psychology: General*, *145*, 918–933. <https://doi.org/10.1037/xge0000177>
- Yang, C., Huang, S. T., & Shanks, D. R. (2017). *Perceptual fluency affects judgments of learning: The font size effect*. Manuscript submitted for publication.
- Yang, C., Potts, R., & Shanks, D. R. (2017a). The forward testing effect on self-regulated study time allocation and metamemory monitoring. *Journal of Experimental Psychology: Applied*, *23*, 263–277. <https://doi.org/10.1037/xap0000122>
- Yang, C., Potts, R., & Shanks, D. R. (2017b). Metacognitive unawareness of the errorful generation benefit and its effects on self-regulated learning. *Journal of Experimental Psychology: Learning, Memory,*

- and Cognition*, 43, 1073–1092. <https://doi.org/10.1037/xlm0000363>
- Yang, C., & Shanks, D. R. (2017). The forward testing effect: Interim testing enhances inductive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. <https://doi.org/10.1037/xlm0000449>
- Zhao, Q. (2012). Effects of accuracy motivation and anchoring on metacomprehension judgment and accuracy. *Journal of General Psychology*, 139, 155–174. <https://doi.org/10.1080/00221309.2012.680523>
- Zhao, Q., & Linderholm, T. (2011). Anchoring effects on prospective and retrospective metacomprehension judgments as a function of peer performance information. *Metacognition and Learning*, 6, 25–43. <https://doi.org/10.1007/s11409-010-9065-1>