Memory

ISSN: 0965-8211 (Print) 1464-0686 (Online) Journal homepage: https://www.tandfonline.com/loi/pmem20

The influence of visual mental imagery size on metamemory accuracy in judgment of learning

Tongtong Li, Xiao Hu, Jun Zheng, Ningxin Su, Zhaomin Liu & Liang Luo

To cite this article: Tongtong Li, Xiao Hu, Jun Zheng, Ningxin Su, Zhaomin Liu & Liang Luo (2017) The influence of visual mental imagery size on metamemory accuracy in judgment of learning, Memory, 25:2, 244-253, DOI: 10.1080/09658211.2016.1153658

To link to this article: https://doi.org/10.1080/09658211.2016.1153658

Published online: 09 Mar 2016.

MEMORY, 2017
VOL. 25, NO. 2, 244–253
http://dx.doi.org/10.1080/09658211.2016.1153658

Routledge
Taylor & Francis Group

The influence of visual mental imagery size on metamemory accuracy in judgment of learning

Tongtong Li[a], Xiao Hu[a], Jun Zheng[c], Ningxin Su[a], Zhaomin Liu[b] and Liang Luo[c]

[a]State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, People's Republic of China; [b]School of Sociology, China University of Political Science and Law, Beijing, People's Republic of China; [c]Collaborative Innovation Center of Assessment Toward Basic Education Quality, Beijing Normal University, Beijing, People's Republic of China

ABSTRACT

Many studies have found the font size of to-be-remembered words has a significant influence on judgments of learning (JOLs). However, few studies have investigated whether JOLs are affected by the mental imagery size of to-be-remembered words, even when the font sizes themselves are kept identical in study materials. This study investigated whether the visual mental imagery size influences the participants' JOLs and what the underlying mechanisms are. In Experiments 1 and 2, participants learned words with identical font sizes, mentally generated large or small imageries and then made JOLs. We found that JOLs under the large imagery condition were significantly higher than those under the small imagery condition, but actual recall performance exhibited no significant difference. In Experiment 3, participants pressed a button immediately after mental imagery generation and showed that it took significantly longer to generate large imageries than to generate small imageries, and the difference in JOLs between two conditions was no longer significant. In Experiment 4, we used a questionnaire to investigate the contribution of beliefs and found that participants believed large imageries were easier to remember. These findings indicate that imagery size has a significant impact on JOLs, in which beliefs may play a leading role.

ARTICLE HISTORY
Received 30 May 2015
Accepted 7 February 2016

KEYWORDS
Metamemory; judgments of learning; visual mental imagery; beliefs; processing fluency

Judgments of Learning (JOLs) refer to the prospective predictions of future memory performance by a learner (Nelson, 1990). Numerous studies have shown that perceptual information in study materials has a significant influence on JOLs (Alban & Kelley, 2013; Besken & Mulligan, 2013; Rhodes & Castel, 2009; Yue, Castel, & Bjork, 2013). For example, Rhodes and Castel (2008) asked participants to learn words shown in large (48 pt) and small font sizes (18 pt). The results showed that participants tended to believe that their memory performance would be higher with larger font sizes than with smaller font sizes. This font-size effect has been further confirmed by many subsequent studies (Hu, Liu, Li, & Luo, 2015; Li, Xie, Li, & Li, 2014; McDonough & Gallo, 2012; Mueller, Dunlosky, Tauber, & Rhodes, 2014). In these studies, most of the investigators explicitly changed the characteristics of the perceptual information in the study materials. In the present study, we attempt to examine whether JOLs are affected by the mental imagery font size of to-be-remembered words even when the font sizes are identical in the study materials.

Visual mental imagery refers to the image of the object that is kept in a person's mind when the object is no longer in front of him/her (Kosslyn, 1975); it can preserve the external appearance and metric spatial information of the object (Shepard, 1978). Regarding the relationship between mental imagery and JOLs, some studies focused on the influence of interactive imagery on JOLs when it was used as a type of encoding strategy (Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Dunlosky & Nelson, 1994; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Rabinowitz, Ackerman, Craik, & Hinchley, 1982; Robinson, Hertzog, & Dunlosky, 2006; Shaughnessy, 1981). For example, Begg and his colleagues had participants learn noun-noun paired associates by producing interactive imageries or separate imageries and then forming JOLs. They found that, compared to the separate imagery group, the interactive imagery group had better recall performance and higher JOLs. Dunlosky and Nelson (1994) conducted further research and found that the effect of imagery facilitating memory only appeared in the delayed JOL condition, but not in the immediate JOL condition. This result indicated that although interactive imagery is an effective strategy for forming memories, participants are rarely aware of it before having prior knowledge. In other words, when participants are making delayed JOLs they are attempting to

CONTACT Liang Luo luoliang@bnu.edu.cn Collaborative Innovation Center of Assessment Toward Basic Education Quality, Beijing Normal University, Beijing, People's Republic of China, Zhaomin Liu zhaominl@cupl.edu.cn School of Sociology, China University of Political Science and Law, Beijing, People's Republic of China

retrieve the target word from memory, they started realise that they benefited from the interactive imagery strategy.

Most importantly, numerous studies have found that the ease with which an individual can derive an image is negatively correlated with JOLs, such that faster latencies are associated with higher JOLs (Hertzog et al., 2003; Robinson et al., 2006). Based on the above research, although these studies explored how interactive imagery influences JOLs, they regarded interactive imagery as a type of encoding strategy (Kuhlmann & Touron, 2012), and participants took imagery to be a mediator to associate word pairs, which facilitates semantic elaboration and, ultimately, improves memory performance to some extent (Bower, 1970). Therefore, because participants knew that it was more effective after they experienced previous tasks, they believed that the items that were studied by interactive imagery are more memorable and thus should have higher JOLs (Begg et al., 1989). However, mental imagery itself has a large number of properties, such as size, clarity and vividness. Do these properties influence JOLs? To our knowledge, no previous study has examined whether these mental imagery properties influence JOLs.

Although there has been a lack of direct evidence regarding whether the size of mental imagery has an impact on JOLs, the findings concerning the influence of the vividness of mental imagery on individuals' preferences and decision-making have provided indirect evidence on this topic (Chang, 2013; Mandel, Petrova, & Cialdini, 2006). For example, Petrova and Cialdini (2005) used vivid and non-vivid advertising imageries to induce participants to generate mental imagery with varying vividness: after the imagination, the participants were asked to judge their preferences for the products. They found that the more vivid the mental imagery was, the more likely participants were to make positive judgments about the product. In addition, researchers have found that mental imagery can significantly affect individuals' decision-making. For example, Sherman, Cialdini, Schwartzman, and Reynolds (1985) asked participants to imagine the symptoms of different diseases and then to predict the incidence rates of the diseases. They found that when the symptoms of a disease could be imagined more easily, participants believed that the incidence rate of the disease was higher. These studies suggest that the characteristics of mental imagery, such as vividness and clarity, strongly influence people's decision-making and judgments. Decision-making and preference judgments engage multiple psychological processes, including attention-guided encoding, evaluation, retrieval of task-relevant information from memory or external sources, prediction, response, post-decision evaluation of consequences and the resulting updating (Weber & Johnson, 2009). An individual makes the best choice based on his/her evaluation and judgments of the goals and situation (Plous, 1993). Similar to decision-making, the ultimate goals of metamemory monitoring and control are the individual's ability to make adjustments and improvements in subsequent learning processes based on the results of JOLs and to optimise the learning process. Therefore, the processing of JOLs and the decision-making process have much in common (Schwarz, 2004). This led us to ask whether mental imagery size influences JOLs in the same way that clarity and vividness influence decision-making and individuals' preferences.

Furthermore, behavioural studies have shown that mental imagery processing exhibits certain characteristics that are similar to perceptual processing, such as the effect of distance. Kosslyn, Ball, and Reiser (1978) found that after a map was imaged in the mind and participants were asked to scan the point of focus to the target based on the imaged map, the time that participants took increased with the distance between the starting point and the destination, proving the existence of the effect of distance. Broggin, Savazzi, and Marzi (2012) compared the effects of real and mentally generated visual stimuli on simple reaction time (RT) to luminance, contrast, visual motion, and orientation, and found that these variables exert similar effects on visual RT, either when retinally presented or when imagined, providing evidence of some overlap between the structural representation of perception and imagery. In addition, some studies that investigated letter imagery – the visual imagery for a letter – also provided evidence to prove that imagery processing is similar to perceptual processing (Kosslyn et al., 1993; Kosslyn, Thompson, & Alpert, 1997; Podgorny & Shepard, 1978). For instance, Podgorny and Shepard (1978) asked participants from the imagery-with-grid conditions to imagine letters (e.g., T) in an empty grid (two-dimensional 5*5 grid of 25 squares) and make responses to indicate whether a dot fell on or off what was the figural portion of the original grid. The results showed that the difference of reaction time between perceptual-memory condition (a black version of the stimuli was actually presented within the grid) and imagery-with-grid condition was not significant, indicated that the participants benefited from the imagination processing and it was extremely similar to perceptual processing. Neuroimaging studies have also found that imagery and perception have similar neural mechanisms (Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Cichy, Heinzle, & Haynes, 2011; Cui, Jeter, Yang, Montague, & Eagleman, 2007; Kosslyn et al., 1999). These studies, which showed that mental imagery and perception share similar cognitive and neural mechanisms, indicate that font size affects JOLs when it is perceptual; thus, it could also affect JOLs when the size difference is only imagined.

In the current study, across four experiments, we investigate whether the mental imagery generated from different font sizes influences the participants' JOLs and what the underlying mechanisms are. We referred to the research of Podgorny and Shepard (1978) and Kosslyn and his colleagues and asked participants to learn words printed in the same font size and generated large or small font-size imageries based on the provided cues and then made JOLs. Compared to English letters, due to Chinese characters that are pictographic as they portray the object that they represent (e.g., Chee et al., 2000; Tan et al., 2001), Chinese characters have more characteristics related to pictures. The evidence suggested that participants can process the Chinese character similarly to processing pictures, including generating a visual mental imagery of the character itself, which provided a solid foundation for the current study.

We speculate that this paradigm is more beneficial to explore the underlying mechanism of the font-size effect. According to previous studies, the main explanations of the effect are the fluency theory and the beliefs theory (Dunlosky, Mueller, & Tauber, 2014). However, researchers have not reached an agreement regarding which explanation mediates the font-size effect and have yet to differentiate their particular contributions (Mueller et al., 2014). For example, Rhodes and Castel (2008, Experiment 6) manipulated font size (48 pt vs. 18 pt) and font format (standard "picnic" vs. alternating "pIcNiC") to verify the fluency hypothesis, and the results indicated that the font-size effect on JOLs only occurred in the standard format condition but was negligible in the alternating format. Therefore, they argued that the presentation of the alternating format may diminish the fluency of the words, and this is a powerful cue that can eliminate the difference in JOLs between large and small items. Nevertheless, Mueller et al. (2014) had participants decide whether a string of letters is a word or a non-word and recorded their response time. The data indicated no significant difference between response times for large and smaller font-size words. Furthermore, they also performed a survey to explore whether participants have relevant beliefs about how font size influences memory before they participated in the experiments, and the researchers found that participants always believe that larger words are more memorable than smaller words. These evidences suggest that the contribution of processing fluency and beliefs to JOLs do not reached a unanimous conclusion and it is difficult to distinguish the contribution of two factors to the font-size effect.

Why were previous researchers unable to distinguish between processing fluency and the contribution of beliefs to the font-size effects? One predominant reason is that these studies only manipulated the physical size of the study materials explicitly, which led to the same direction of influence of processing fluency and beliefs on JOLs. When words were presented in large font sizes, not only was their processing fluency high but the large font size itself was also in line with people's beliefs. For this reason, in the current study, participants learned words printed in the same font size and then generated large or small font imageries based on the provided cues: this manipulation separates fluency and beliefs. Many studies have confirmed that generating larger imageries require more mental effort and more time. Thus, if mental imagery size matters (with larger imageries producing higher JOLs), then it would provide solid experimental evidence that fluency might be unlikely to account for the influence of variables on JOLs. Moreover, if small size mental imagery produces greater JOLs, then it would confirm the fluency hypothesis that an item that is processed more easily will be judged to be more memorable.

1. Experiment 1

In Experiment 1, we manipulated the mental imagery size. In each trial participants learned words printed in the same font size and then they were instructed to visualise the character itself into large or small size based on the provided cues (large or small 田-shaped grids).[1]

1.1. Methods

1.1.1. Participants

Thirty-one college students (20 females, 11 males) participated in Experiment 1 in return for ¥30. One participant was excluded because of missing memory performance data.

1.1.2. Materials

A set of 46 Chinese 2-character words, for example, 医院 (hospital), were selected from a Chinese database (Cai & Brysbaert, 2010) with a word frequency of between 0.03 and 7.33 per million words. Four words were used for practice, and the remaining 42 were used for the experiment, of which 6 were used as either primary or recency buffer words and were excluded from all of the reported analyses.

The 田-shaped grids had two sizes. The large grid was 11 cm by 11 cm, whereas the small grid was 1.4 cm by 1.4 cm. Two 田-shaped grids in the identical size appeared horizontally on the monitor in parallel.

1.1.3. Procedures

First, participants practiced a mock test procedure. To ensure that participants completely understood the experiment, they were required to verbally describe the process of mental imagery generation (the report was omitted in the experiment).

All of the words were presented in black 90-pt font on a white background. Participants were instructed that they would study words for 2 s, and all words were presented in an identical font size. Each word was immediately followed by a blank screen for 800 ms, after which two 田-shaped grids of identical size randomly appeared and remained on the screen for 10 s. Participants were instructed to visualise the character itself in a large or small size according to the large or small 田-shaped grids, not a visual image of the physical item that the word stimuli represents. For example, when the participants sees the word "大象"(elephant), he or she is instructed to visualise the character "大" and "象" into two 田-shaped grids, not the mental imagery of a large grey creature (see Figure 1). Participants were explicitly

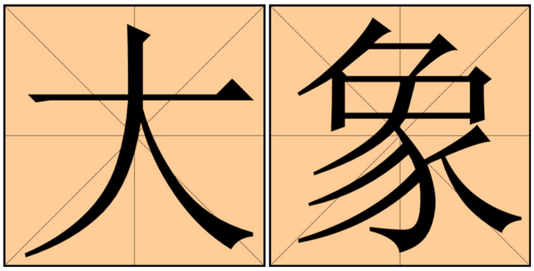

Figure 1. An example of the visualisation of characters in the 田-shaped grids.

told that they did not need to write the word in their mind, which means that they were not supposed to visualise the act of drawing the lines that form the characters; they were just supposed to visualise the whole character already written. After the disappearance of the cue, participants were given 5 s to make a judgment about the likelihood of future recall on a scale from 0 (no recall at all) to 100 (certain to recall). Participants were encouraged to use the entire range of the scale. Immediately following the study list, participants engaged in mathematics exercises for 2 min as a distractor task. Finally, participants were asked to recall as many words as possible and typed the answers into the computer.

Of the 42 trials, half of the cues were presented in a large size and the other half were presented in a small size. The six buffer words were presented in a fixed sequence. The remaining 36 items were presented in a pseudo-random order (with the condition that no more than three items with the same 田-shaped grid size were presented consecutively).

1.2. Results and discussion

The mean JOLs and actual recall performance are presented as percentages in Figure 2. The mean JOLs of participants under the large mental imagery condition ($M = 62.73$, SD = 21.70) were significantly higher than those under the small

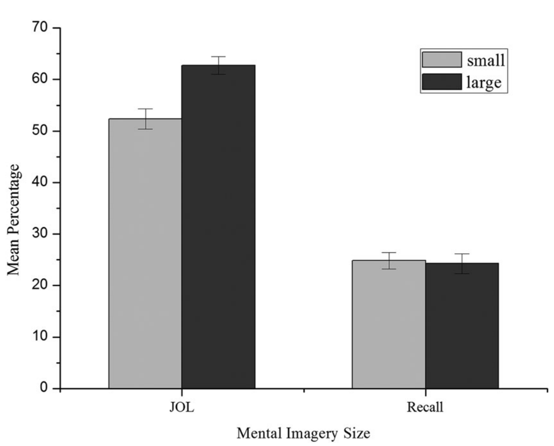

Figure 2. Predicted and recall performance by mental imagery size (presented by the 田-shaped grid) in Experiment 1. Error bars reflect standard errors of the mean in all figures. JOL = Judgment of learning.

Table 1. Mean Kruskal–Goodman gamma correlation for Experiments 1, 2, and 3.

	Experiment 1	Experiment 2	Experiment 3
JOLs-Recall	0.12 (0.39)	−0.009 (0.432)	0.259 (0.23)**
JOLs-Imagery size	0.35 (0.37)**	0.196 (0.439)**	0.128 (0.324)*
Recall-Imagery size	0.15 (0.54)	−0.0787 (0.41)	−0.044 (0.453)
Image time-Imagery size	–	–	0.203 (0.232)**

Note: Values are mean Kruskal–Goodman gamma correlation with SDs in parentheses.
*$p \leq .05$.
**$p \leq .001$ for the difference from zero.

mental imagery condition ($M = 52.35$, SD = 19.06), $t(29) = 4.21$, $p < .001$, Cohen's $d = .51$. Actual recall performance showed no significant differences between the two mental imagery conditions (large mental imagery condition: $M = 24.82\%$, $SD = .17$; small mental imagery condition: $M = 24.26\%$, SD = .21), $t(29) = 0.24$, $p = .81$, Cohen's $d = .04$.

The Kruskal–Goodman gamma correlation was used to investigate the relationship between the mental imagery generation conditions, JOLs, and recall (Nelson, 1984). The results of the gamma correlations from Experiments 1–3 are shown in Table 1. The results show that the mean correlation between JOLs and the size of the mental imagery differed reliably from zero, $t(29) = 5.24$, $p < .001$, Cohen's $d = .47$, indicating that the size of the mental imagery could significantly predict JOLs and that large items could have higher JOLs. Conversely, the mean correlation between recall and the sizes of the mental imagery generated or JOLs were not significantly different from zero ($p > .1$).

Taken together, data from Experiment 1 suggests that participants regarded items imagined in large font sizes as more memorable than items imaged in a small font sizes, indicating that mental imagery size had a strong influence on JOLs, although it was unrelated to subsequent recall. However, in Experiment 1, we used 田-shaped grids of different sizes to assist participants in generating mental imageries of different sizes. Because the two sizes of the 田-shaped grids were significantly different, differences in the perceptual characteristics likely had an impact on the JOLs (Alban & Kelley, 2013; Rhodes & Castel, 2008, 2009). Therefore, we designed Experiment 2 to examine whether participants continue to use mental imagery size as a basis for judgment without the 田-shaped grids.

2. Experiment 2

In Experiment 2, two English letters replaced the 田-shaped grids. Participants were asked to generate mental imageries corresponding in size to the English letter cues (Farah & Kosslyn, 1981). Half of participants were told that the letter A means the large 田-shaped grid and the letter B means the small 田-shaped grid. The remaining half of the participants were told that the letter A means the small 田-shaped grid and the letter B means the large 田-shaped grid. Except for the cue, the design, materials and procedures were entirely identical to

Experiment 1. If the findings from Experiment 1 resulted from visual differences in the size of the 田-shaped grids, they would not be reproducible in Experiment 2.

2.1. Methods

2.1.1. Participants

Thirty-five college students (26 females, 9 males) participated in Experiment 2 in return for ¥30. One participant was excluded due to a misunderstanding of the procedures of Experiment 2.

2.1.2. Design, materials and procedures

The design, materials and procedures were identical to Experiment 1, except that two English letters (A or B), all in a font size of 44, were used as cues instead of the 田-shaped grids. The meanings of the different letters were balanced among the participants. Half of participants were told that the letter A means the large 田-shaped grid and the letter B means the small 田-shaped grid. The remaining half of the participant were told that the letter A means the small 田-shaped grid and the letter B means the large 田-shaped grid. In addition, we also trained participants to ensure they memorised the meanings of the letters and the size of the imageries that they were to generate.

2.2. Results and discussion

Similar to Experiment 1, the mean JOLs of participants under the large mental imagery condition ($M = 67.77$, SD = 19.22) were significantly higher than those under the small mental imagery condition ($M = 60.58$, SD = 23.14), $t(33) = 3.55$, $p = .001$, Cohen's $d = .34$. The actual recall performance was not significantly different between the two imagery conditions (large mental imagery condition: $M = 21.24\%$, SD = .14; small mental imagery condition: $M = 20.75\%$, SD = .12), $t(33) = .23$, $p = .82$, Cohen's $d = .04$ (see Figure 3).

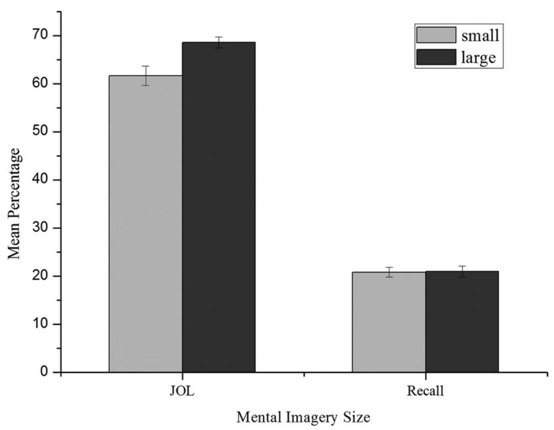

Figure 3. Predicted and recall performance by mental imagery size (presented by letter A or B) in Experiment 2. JOL = Judgment of learning.

The mean correlation between JOLs and the imagery size differed reliably from zero, $t(33) = 2.96$, $p < .05$, Cohen's $d = .29$. However, the mean correlation between recall and the imagery size or JOLs was not significantly different from zero ($t < 1$) (see Table 1).

The results in Experiment 2 led to a similar conclusion as those of Experiment 1 in that imagery size influences JOLs, which excluded the influence of visual differences in the physical characteristics of the 田-shaped grids on the results of Experiment 1. Thus, Experiments 1 and 2 indicated that the JOLs for words with large mental imageries were significantly higher than those for words with small imageries. In the next set of experiments, we investigated how the mental imagery size influences JOLs. Previous research emphasised that various cues can influence JOLs via experience-based or theory-based processes. Experience-based processes reflect the subjective experience (e.g., fluency) in the study phase, and theory-based processes reflect beliefs about how the cues are related to memory (Koriat, Bjork, Sheffer, & Bar, 2004). In Experiments 3 and 4, we separately examined the influence of imagery generation fluency and beliefs on JOLs.

3. Experiment 3

In Experiment 3, we investigated whether the mental imagery size influences JOLs through imagery generation fluency. Previous studies have revealed that it takes more time to image large pictures compared to small pictures (Farah & Kosslyn, 1981; Kosslyn, 1978; Kosslyn et al., 1978). However, it is unclear whether generating large-size Chinese characters also requires more time. In addition, in both Experiments 1 and 2, the time allowed for participants to generate an image was fixed (10 s), implying that we did not know whether there was any difference in the time needed to generate large vs. small imageries. Thus, in Experiment 3, we recorded the time spent to generate the image. If generating large imageries takes less time than generating small imageries, the higher image generation fluency may lead to higher JOLs for words with large imageries. On the other hand, if generating large imageries takes more time, as in previous studies, then the fluency difference would be inconsequent with the JOL difference and may not contribute to JOLs.

3.1. Methods

3.1.1. Participants

Thirty college students (26 females, 4 males) participated in Experiment 3 in return for ¥30. Three participants[2] took over 10 s to generate imagery in approximately 40% of the trials; thus, their data were excluded.

3.1.2. Design, materials and procedures

The design, materials and procedures were consistent with those of Experiment 1, except that participants were asked to press the "ENTER" key immediately after generating the mental image. To maintain consistency between the experiments, the display time of the cue (田-shaped grids) was kept unchanged (10 s). If participants pressed the "ENTER" key within 10 s, the cue remained on the screen until the end of the 10 s. If participants did not press the "ENTER" key within 10 s, the cue disappeared. The imagery generation time was measured from the moment of the appearance of the cue to the moment the participant pressed the "ENTER" key; if the participant did not press the "ENTER" key within 10 s, the generation time was recorded as 10 s.

3.2. Results and discussion

We found that the generation time under the large imagery condition was significantly longer ($M = 5456.74$ ms, SD = 338.23) than under the small imagery condition ($M = 5012.47$ ms, SD = 308.49), $t(27) = 3.497$, $p < .001$, Cohen's $d = .29$.

The mean JOLs under the two imagery conditions exhibited no significant difference (large mental imagery condition: $M = 59.38$, SD = 17.24; small mental imagery condition: $M = 57.32$, SD = 17.04), $t(26) = .99$, $p = .330$, Cohen's $d = .11$. Similarly, actual recall performance showed no significant difference between the two imagery conditions (large mental imagery condition: $M =$ 25.51%, SD = 17.04; small mental imagery condition: $M =$ 25.10%, SD = 13.20), $t(26) = .14$, $p = .89$, Cohen's $d = .02$ (see Figure 4).

The mean gamma correlation between recall and the size of the imagery was not significantly different from zero, while that between the size of the imagery and JOLs was almost significantly different from zero, $t(26) =$ 2.06, $p = .050$, Cohen's $d = .17$ However, the mean gamma correlation between recall and JOLs was significantly different from zero, $t(26) = 5.87$, $p < .001$, Cohen's $d = .36$. In addition, the mean gamma correlation between the size of the imagery and the imagery generation time was significantly different from zero, $t(26) = 4.55$, $p < .001$, Cohen's $d = .30$ (see Table 1).

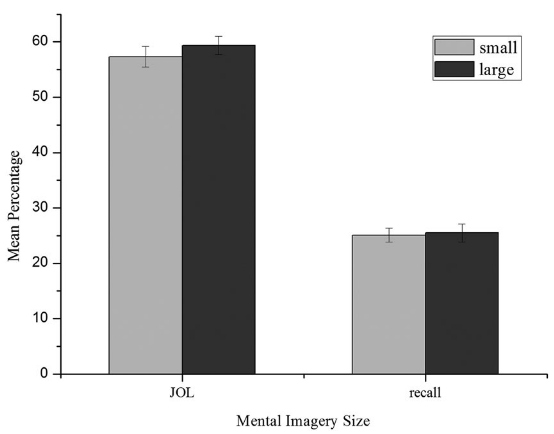

Figure 4. Predicted and recall performance by mental imagery size (presented by the 田-shaped grid) in Experiment 3. JOL = Judgment of learning.

The results of Experiment 3 confirmed that the time required by participants to generate a large image was longer than that required to generate a small image, which is consistent with previous findings (Farah & Kosslyn, 1981; Kosslyn, 1978; Kosslyn et al., 1978). In addition, there was no longer a significant difference between JOLs under the large and small imagery conditions, which will be discussed in detail in the general discussion. The results of Experiment 3 indicate that processing fluency might not be an effective cue in Experiments 1 and 2. What, then, was the cue that enabled the participants to predict better recall performance regarding items with large imageries? Previous studies show that beliefs may play an important role in making JOLs (Koriat et al., 2004; Mueller et al., 2014). Therefore, we designed Experiment 4 to investigate the role of beliefs.

4. Experiment 4

In Experiment 4, we used a questionnaire to examine whether participants believed that words with large imagery sizes are easier to remember. The questionnaire described Experiment 1 in detail, and then, participants were asked to estimate recall for different conditions without actually having experienced the actual experiment. Because participants are not presented with any stimuli, there is no influence of direct experience on the predictions. If participants showed a significant difference in predicted performance for words that generated different mental imageries with different font sizes, the results were considered to indicate that people's beliefs have a considerable impact on JOLs (Li et al., 2014; Mueller, Tauber, & Dunlosky, 2013; Mueller et al., 2014).

4.1. Methods

4.1.1. Participants

Two-hundred participants (92 females, 115 males) participated in the survey in return for ¥5. One participant gave answers that exceeded the total number of words when predicting recall performance; thus, the data from this participant were excluded.

4.1.2. Materials and procedures

The experiment was conducted in the form of an online questionnaire. In the questionnaire, the details of Experiment 1 were described, and examples of 田-shaped grids of different sizes were presented. Participants were asked first to read the description of the experiment and to separately estimate the recall performance of items with large and small imagery sizes. The order of asking for the estimates was counterbalanced across participants.

4.2. Results and discussion

A paired-sample *T*-test of the predictions showed that predicted performance under the large imagery condition ($M = 69.71\%$, SD = 23.40) was significantly higher than that under the small imagery condition ($M = 63.65\%$, SD = 25.11), $t(198) = 6.67$, $p < .001$, Cohen's $d = .25$.

The prediction results in Experiment 4 were similar to those in Experiment 1 and Experiment 2. Participants in Experiment 4 did not actually participate in the experiment but only predicted their future performance based on the description of the experiment. Therefore, participants were unable to perceive the fluency of mental imagery processing, and thus the differential fluency of processing could not have influenced their JOLs. This result indicates that participants might hold the belief that words with large imagery sizes are easier to remember and that these beliefs had an influence on the prediction of future recall.

5. General discussion

In this study, we examined the effect of mental imagery size on learners' JOLs. In Experiments 1 and 2, participants studied words with identical font sizes and then, based on cues, mentally generated large or small imageries and made JOLs. The results showed that JOLs under the large imagery condition were significantly higher than those under the small imagery condition, but actual recall did not exhibit any significant difference, indicating that the mental imagery of different font sizes led to a metamemory illusion. In Experiment 3, participants were asked to press a button immediately after imagery generation, and the results showed that the generation of large imageries required significantly more time than the generation of small imageries. In addition, JOLs under the two imagery conditions were no longer significantly different. In Experiment 4, we adopted a questionnaire to examine the role of beliefs and found that even if participants did not take part in actual experiments, based only on the mere description of the experiment, they believed that words with a large imagery size were easier to remember.

Previous studies on metamemory have established that various cues exert their influence on JOLs mainly through experience-based or theory-based processes (Koriat et al., 2004). The theory-based processes refer to naive theories or metacognitive beliefs held by participants on how various cues affect recall (Mueller et al., 2013, 2014), whereas the experience-based processes reflect the subjective experience directly acquired by participants in the learning process (Koriat, Nussinson, Bless, & Shaked, 2008). Processing fluency is an important experience-based cue (Bjork, Dunlosky, & Kornell, 2013). Imagery studies have found that, in relation to the generation of small imagery, the generation of large imagery requires more time (Farah & Kosslyn, 1981; Kosslyn, 1978; Kosslyn et al., 1978). The data analysis results from Experiment 3 also confirm that the generation of large font imagery requires more time than the generation of small font imagery. This indicates that the generation of large imagery demands more mental effort, which leads us to believe that processing words with a large imagery size is less fluent than processing words with a small imagery size. If mental imagery exerts its effect on JOLs through processing fluency, JOLs under the small imagery condition should be greater than those under the large imagery condition. However, the results from Experiment 1 and Experiment 2 contradict this assumption and do not support the above hypothesis. Thus, we conclude that the processing fluency of imagery generation has little, if any, impact on the JOLs in Experiment 1 and Experiment 2.

What led to the significantly higher JOLs under the large imagery condition in Experiment 1 and Experiment 2? Many studies have shown that people generally hold the belief that words with a large font size are easier to remember than those with a small font size (Li et al., 2014; Mueller et al., 2014). These studies have posited that the effect of beliefs on JOLs might be implicit and subtle. For example, Alban and Kelley (2013) found that individuals thought that heavy items were easier to remember than lighter items, and they speculated that this finding was most likely due to individuals' stereotypes concerning weight (heavy objects are usually positive, while light objects are negative). Thus, they reversed the individuals' beliefs (e.g., upgrades in computers signify that the change from heavy to light represents progress and technological development) by priming and found that the JOLs were also reversed. Mueller et al. (2013, 2014) asked participants to make predictions concerning their future recall performance before the study, when they had not had any contact with the study materials such that it was impossible for the individuals to perceive factors such as the difficulty of the study materials. Nevertheless, participants believed that larger font sizes would yield better recall performance. We speculate that in Experiment 1 and Experiment 2, although the difference in font size only existed in mental imagery, participants likely continued to use their beliefs that words in large font sizes are easier to remember, which may have had a significant influence on JOLs through theory-based processes. In Experiment 4, this hypothesis was validated through the questionnaire, which provided evidence that most participants believed that words with large imagery are easier to remember, indicating that this belief led to the significantly higher JOLs under the large imagery condition. However, although the result from Experiment 4 provide evidence that people's beliefs may play a large role on JOLs, the mechanism is indirect, and further research is needed to establish how beliefs about imagery size mediate the effect on JOLs.

Why was the gamma correlation between JOLs and recall not significant in Experiment 1 and Experiment 2 but was significant in Experiment 3? Before answering this question, we should think about another question first-Why, then, did the significant difference between JOLs under the large imagery condition and small imagery condition disappear in Experiment 3 when participants were asked to press a key immediately after imagery generation? Comparing Experiment 1 and Experiment 3, except for the request that participants press the button after imagery generation, all of the experimental operations were identical. The reason for the results might be that when asked to press the button after generating mental imageries, participants realised the difference between the time to generate the large and small imageries, and the important experience-based cue of fluency started to exert an influence on JOLs. In Experiment 1, when participants did not realise the difference in the image generation time, it was mainly their beliefs that had a strong impact on JOLs. However, in Experiment 3, participants began to use the generation time as an effective cue. That is, processing fluency and beliefs simultaneously affected JOLs, albeit in opposite directions: generating small imageries took less time than generating large imageries, whereas words with a large imagery size were believed to be easier to remember. The directions of the effect of fluency and beliefs were quite contrary, leading to the disappearance of the significant difference between the JOLs under the large imagery and small imagery conditions in Experiment 3. These results indicated that the role of processing fluency and its degree might be restricted in some conditions to some extent (e.g., whether participants realise the difference of time required to generate imageries). Only when certain conditions were met did processing fluency become active and have an effect on JOLs. The gamma correlation in Experiment 3 also confirmed this speculation, which is that the relative accuracy improved significantly when participants made JOLs based on beliefs and fluency at the same time.

According to a new theory in the area of metacognition – the analytic processing theory (Mueller, Dunlosky, & Tauber, 2015) – when people explicitly to make a JOL, they adopt an analytic problem-solving approach to reduce their uncertainty about future memory performance. In the Experiments 1 and 2, participants might have a belief about why the size of imagery influences memory and they used the belief about the cue (the size of imagery) as they made JOLs. They might believe that the large imagery is huge and it will be memorised more effectively. However, in the Experiment 3, the operation – in which participants were asked to press a button after generating the imagery – might lead them to search for more cues to reduce uncertainty about which words they will remember, and the cue is imagination time which is a more effective cue to JOLs. In addition, perhaps, it might be that the action of pressing the button itself makes participants realise that something happened that we are not aware of, which leads participants to improve their metacognitive ability. Therefore, this open question needs more attention in the future studies.

Furthermore, we found that there was a different magnitude of the JOLs made in Experiments 1 and 2. To explore whether the difference of JOLs had an impact on results of the Experiment 1 and Experiment 2, we conducted a post-hoc 2 (imagery conditions: large imagery or small imagery) × 2(cue type: the 田-shaped grids in Experiment 1 or the letter in Experiment 2) ANOVA, and the result showed that the main effect of imagery conditions (F (1,126) = 5.40, p = .02, $\eta_p^2 = .04$) and cue type (F (1,126) = 4.16, p = .04, $\eta_p^2 = .03$) were significant, but the interaction effect was not ($p > .6$). It is important to note that the interaction effect of imagery conditions and cue types was not significant, which that means the cue type would not contaminate the influence of imagery condition and has no other effect on the font-size effect on JOLs, which was most important to the present study.

Given that the only change between Experiment 1 and 2 was the cue type (aside from the participants), then, why was there a significant difference in mean JOL magnitude across Experiments 1 and 2? We put forward three possible explanations as follows. First, the manipulation in Experiment 2 may have led to participants visualising a pictorial representation of the word the characters represented rather than visualising the characters themselves, and this is what led to this difference in JOLs magnitude between the experiments. Concerning this issue, we explored whether word imageability has an impact on JOLs.[3] The analysis results showed that in Experiment 1 and 2, only the main effect of imagery condition was significant; the main effect of imageability and the interaction were not significant (see footnote 3 for details), indicating that JOLs were not influenced by words' imageability. From this analysis it appears this explanation may not be valid. Second, the different cues used in the two experiments induced different levels of difficulty with regard to visualisation. Compared to Experiment 1(the 田-shaped grids), participants in Experiment 2(the letters) could control their imagination, the size of characters they imagined might have been slightly different , and visualisation with letters might be easier and participants might have had a stronger feeling of control, which resulted in higher JOLs in Experiment 2. To investigate this hypothesis, we considered that, if participants in Experiment 2 felt the task was easier because of the feeling of control, their characters' visualisation sizes might be more diverse than those in Experiment 1. Furthermore, the diversity of JOLs in Experiment 2 was also larger than the diversity in Experiment 1. Therefore, we explored the variability of JOLs of two experiments by computing the coefficient of variation (Allison, 1978)[4] (Exp 1: $CV_{large} = .35$, $CV_{small} = .36$; Exp 2: $CV_{large} = .28$, $CV_{small} = .38$) and the CVs magnitude did not differ across the two experiments, which indicated that this hypothesis might not be tenable. In fact, we have trained the participants in visualisation manipulation and reduced this possibility. Of course, there was also a possibility that the influence of difficulty was a systematic error; consequently, the JOLs increased in both of the imagery conditions of Experiment 2. Third, an individual's characteristics. It is because JOL is a type of subjective judgement (Nelson, 1990), so JOLs are more influenced by individual's characteristics. Therefore, we guess that the difference of JOLs between two experiments might be due to the participants' judgement preference. In addition, the JOLs difference among the different experiments also appeared in the previous studies. For example, in the first study about the font-size effect on JOLs – Rhodes and Castel (2008) – in Experiment 1 ($M_{large} = 60.81$, $M_{small} = 48.63$) and the time 1 condition of Experiment 2 ($M_{large} \approx 52$, $M_{small} \approx 42$), although the methodologies of the two conditions were identical, the JOLs in the two experiments were numerically different.[5] Of course, this issue should be explored further in future studies.

In this study, through the mental imagery task, the acting directions of the two factors on JOLs were successively separated, providing a new avenue of inquiry for future studies. Meanwhile, the current study identified a new cue that has an effect on JOLs, and we have called this new effect the imagery-size effect. We refer to beliefs that may play a leading role in JOLs and the contribution of processing fluency that may be conditional. Furthermore, the current results, especially the results of Experiment 3, suggest that both processing fluency and beliefs could influence JOLs, but their interactions require further research.

Notes

1. The 田-shaped grid is an aid for writing Chinese characters. In our experiments, this grid helps learners to imagine to-be-remembered words in a large or small font size.
2. There were three participants who took longer than 10 s for imagery generation in more than 40% of the trials (max: 83.33%, min: 41.67%) in Experiment 3. Because the data for such a high percentage of the trials were missed, there were concerns that the characters were actually not completely generated. Therefore, the data for these three participants were excluded. The average percentage of missed trials among the remaining 27 participants was 6.39% (max: 27.78%, min: 0%). The data were analysed using two methods. In the first method, all of the trial data were analysed, whereas in the second method, the trials with no reaction for which the imagery generation time was recorded as 10 s were all excluded, and only the data for the remaining trials were analysed. Because the results of both methods were similar, only the analytical results based on the data using the first method are presented here.
3. To explore this issue, we conducted a questionnaire about the imageability of the words used in the present study. Imageability is defined as words' capacity to arouse mental imageries of things or event (Altarriba, Bauer, & Benvenuto, 1999; Cortese & Fugett, 2004). One hundred participants rated the words on how difficult it is to form a mental image in a 7-points scale ("1" means very difficult to arouse a mental image; "7" means very quickly and easily to arouse a mental image). According to the rating results, we equally divided the learning materials (36 words) into two groups, the high imageability group ($M = 6.14$, SD = .65) and the low imageability group ($M = 4.77$, SD = .21), and the imageability estimation between the two groups was significant ($t(34) = 8.48$, $p < .001$, Cohen's $d = 2.38$). The 2(imageability groups: high and low) × 2(imagery conditions: large and small) post-hoc ANOVA results was as followed. The main effect of imagery condition was significant (Exp 1: $M_{small} = 52.75$, $M_{large} = 62.81$, $F(1,116) = 6.67$, $p = .011$, $\eta_p^2 = .05$; Exp 2: $M_{small} = 61.69$, $M_{large} = 68.44$, $F(1,136) = 3.21$, $p = .08$, $\eta_p^2 = .02$), and the main effect of imageability (Exp 1: $M_{low} = 55.3$, $M_{high} = 60.26$, $F(1,116) = 1.62$, $p = .21$, $\eta_p^2 = .01$; Exp 2: $M_{low} = 62.27$, $M_{high} = 67.86$, $F(1,136) = 2.21$, $p = .14$, $\eta_p^2 = .02$) and the interaction was not significant ($p > .7$).
4. The coefficient of variation (CV) is a measure of dispersion of data relative to the mean, and defined as the standard deviation of a random variable in ratio to the expectation of the random variable (Allison, 1978).
5. In Rhodes and Castel's study (2008), they did not report the magnitude of JOLs in the large condition and small condition in time 1 and time 2, respectively, so the value of JOLs under the time 1 condition was estimated from the figure. Regardless of the time, the JOLs of Experiment 2 in the large condition was 42.67, and was 36.28 for the small condition, which was far less than the JOLs magnitude in Experiment 1 in both the large and small condition. Furthermore, due to the lack of experimental data, we did not analyse the data or perform a statistical test.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by the National Natural Science Foundation of China grant number (31100817, 31000506); and Beijing Higher Education Young Elite Teacher Project and the project for young people who are devoted to the researches of humanities and social sciences (grant number 10YJCXLX032) of the Ministry of Education of the People's of Republic of China; and the Training Program Foundation for the Excellent Youth Scholars of China University of Political Science and Law.

References

Alban, M. W., & Kelley, C. M. (2013). Embodiment meets metamemory: Weight as a cue for metacognitive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1628–1634.

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, *23*(15), 1427–1431.

Allison, P. D. (1978). Measures of inequality. *American Sociological Review*, *43*(6), 865–880.

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 578–602.

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*(5), 610–632.

Besken, M., & Mulligan, N. W. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition*, *41*(6), 897–903.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444.

Bower, G. H. (1970). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 529–533.

Broggin, E., Savazzi, S., & Marzi, C. A. (2012). Similar effects of visual perception and imagery on simple reaction time. *The Quarterly Journal of Experimental Psychology*, *65*(1), 151–164.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS One*, *5*(6), e10729.
Chang, C. (2013). Imagery fluency and narrative advertising effects. *Journal of Advertising*, *42*(1), 54–68.
Chee, M. W., Weekes, B., Lee, K. M., Soon, C. S., Schreiber, A., Hoon, J. J., et al. (2000). Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: Evidence from fMRI. *Neuroimage*, *12*(4), 392–403.
Cichy, R. M., Heinzle, J., & Haynes, J. D. (2011). Imagery and perception share cortical representations of content and location. *Cerebral Cortex*, *22*, 372–380.
Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 384–387.
Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research*, *47*(4), 474–478.
Dunlosky, J., Mueller, M., & Tauber, S. K. (2014). The contribution of processing fluency (and beliefs) to people's judgments of learning. In D. S. Lindsay, C. M. Kelley, A. P. Yonelinas, & H. L. Roediger III (Eds.), *Remembering: Attributions, processes, and control in human memory: Papers in honour of Larry L. Jacoby* (pp. 46–64). New York: Psychology Press.
Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language*, *33*(4), 545–565.
Farah, M. J., & Kosslyn, S. M. (1981). Structure and strategy in image generation. *Cognitive Science*, *5*(4), 371–383.
Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 22–34.
Hu, X., Liu, Z., Li, T., & Luo, L. (2015). Influence of cue word perceptual information on metamemory accuracy in judgment of learning. *Memory* (in press). doi: 10.1080/09658211.2015.1009470
Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, *133*(4), 643–656.
Koriat, A., Nussinson, R., Bless, H., & Shaked, N. (2008). Information-based and experience-based metacognitive judgments: Evidence from subjective confidence. In J. Dunlosky & R. A. Bjork (Eds.), *A handbook of memory and metamemory* (pp. 117–135). Mahwah, NJ: Erlbaum.
Kosslyn, S. M. (1975). Information representation in visual images. *Cognitive Psychology*, *7*(3), 341–370.
Kosslyn, S. M. (1978). Measuring the visual angle of the mind's eye. *Cognitive Psychology*, *10*(3), 356–389.
Kosslyn, S. M., Alpert, N. M., Thompson, W. K., Maljkovic, V., Weise, S. B., Chabris, C. F., et al. (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, *5*(3), 263–287.
Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(1), 47–60.
Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Ganis, G., … Alpert, N.M. (1999). The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science*, *284*(5411), 167–170.
Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *Neuroimage*, *6*(4), 320–334.
Kuhlmann, B. G., & Touron, D. R. (2012). Mediator-based encoding strategies in source monitoring in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1352–1364.
Li, F., Xie, R., Li, X., & Li, W. (2014). The influence of perceptual information on control processes involved in self-regulated learning: Evidence from item selection. *Psychonomic Bulletin & Review*, *22*(4), 1007–1013.
Mandel, N., Petrova, P. K., & Cialdini, R. B. (2006). Images of success and the preference for luxury brands. *Journal of Consumer Psychology*, *16*(1), 57–69.
McDonough, I. M., & Gallo, D. A. (2012). Illusory expectations can affect retrieval-monitoring accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(2), 391–404.
Mueller, M. L., Dunlosky, J., & Tauber, S. K. (2015). 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*. doi: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.
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*(2), 378–384.
Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*(1), 109–133.
Nelson, T. O. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation*, *26*, 125–173.
Petrova, P. K., & Cialdini, R. B. (2005). Fluency of consumption imagery and the backfire effects of imagery appeals. *Journal of Consumer Research*, *32*(3), 442–452.
Plous, S. (1993). *The psychology of judgment and decision making* (pp. 15–21). New York, NY: Mcgraw-Hill Book Company.
Podgorny, P., & Shepard, R. N. (1978), Functional representations common to visual perception and imagination. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(1), 21–35.
Rabinowitz, J. C., Ackerman, B. P., Craik, F. I., & Hinchley, J. L. (1982). Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology*, *37*(6), 688–695.
Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*(4), 615–625.
Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review*, *16*(3), 550–554.
Robinson, A.E., Hertzog, C., & Dunlosky, J. (2006). Aging, encoding fluency, and metacognitive monitoring. *Aging, Neuropsychology, and Cognition*, *13*, 458–478.
Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, *14*(4), 332–348.
Shaughnessy, J. J. (1981). Memory monitoring accuracy and modification of rehearsal strategies. *Journal of Verbal Learning and Verbal Behavior*, *20*(2), 216–230.
Shepard, R. N. (1978). The mental image. *American Psychologist*, *33*(2), 125–137.
Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease the mediating effect of ease of imagery. *Personality and Social Psychology Bulletin*, *11*(1), 118–127.
Tan, L. H., Liu, H. L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J. H. (2001). The neural system underlying Chinese logograph reading. *Neuroimage*, *13*(5), 836–846.
Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, *60*, 53–85.
Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is – and is not – a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition*, *41*(2), 229–241.