Memory

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/pmem20

Does the reactivity effect of judgments of learning transfer to learning of new information?

Baike Li, Wenbo Zhao, Aike Shi, Yongen Zhong, Xiao Hu, Meng Liu, Liang Luo & Chunliang Yang

MEMORY
https://doi.org/10.1080/09658211.2023.2208792

Does the reactivity effect of judgments of learning transfer to learning of new information?

Baike Li [a], Wenbo Zhao[b], Aike Shi[a], Yongen Zhong[a], Xiao Hu[c,e], Meng Liu[d], Liang Luo[a,c] and Chunliang Yang[a,e]

[a]Institute of Developmental Psychology, Faculty of Psychology, Beijing Normal University, Beijing, People's Republic of China; [b]Collaborative Innovation Center of Assessment for Basic Education Quality, Beijing Normal University, Beijing, People's Republic of China; [c]Faculty of Psychology, Beijing Normal University, Beijing, People's Republic of China; [d]School of Psychology, South China Normal University, Guangzhou, People's Republic of China; [e]Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education, Beijing Normal University, Beijing, People's Republic of China

ABSTRACT

Making judgments of learning (JOLs) can reactively change memory, a phenomenon termed the *reactivity effect*. The current study was designed to explore whether the reactivity effect transfers to subsequent learning of new information. Participants studied two blocks of words (Experiment 1) or related word pairs (Experiments 2 & 3). In Block 1, participants in the experimental (JOL) group made a JOL while studying each item, whereas the control (no-JOL) group did not make item-by-item JOLs. Then both groups studied Block 2, in which they did not make JOLs, and finally, they took a test on Blocks 1 and 2. Across Experiments 1 −3, the results showed superior Block 1 test performance in the JOL than in the no-JOL group, demonstrating a positive reactivity effect. Critically, there was minimal difference in Block 2 test performance between the two groups, implying little transfer of the positive reactivity effect to subsequent learning of new information. Furthermore, Experiment 3 demonstrated that the reactivity effect still failed to transfer even when participants explicitly appreciated the benefits of making JOLs. Educational implications are discussed.

ARTICLE HISTORY
Received 29 March 2022
Accepted 24 April 2023

KEYWORDS
Judgments of learning; reactivity effect; transfer; learning of new information

Judgments of learning (JOLs; metacognitive judgments about the likelihood of remembering an item in a later memory test) are of vital importance for human learning and memory because learners typically regulate their study activities (e.g., decisions about when, what, and how to study) according to their JOLs. A large number of studies have been conducted to figure out to what extent JOLs are accurate in reflecting actual memory status, to investigate mechanisms underlying JOL formation, and to explore the relations between metamemory monitoring and control (for reviews, Rhodes, 2016; Rhodes & Tauber, 2011; Yang et al., 2021). Numerous studies have found that learners are prone to allocate more time to studying items perceived as less-well studied than to those perceived as well studied (Dunlosky & Hertzog, 1997; Dunlosky & Thiede, 2004; Nelson & Narens, 1994; Verkoeijen et al., 2005; Yang et al., 2017). These findings reflect that metamemory monitoring (e.g., JOLs) can affect memory in an indirect way through its influences on metamemory control (Finn, 2008; Metcalfe & Finn, 2013; Rhodes, 2016; Rhodes & Castel, 2009).

An emerging body of recent studies found that making JOLs can also affect memory in a direct way. To be specific, recent studies found that instructing learners to make item-by-item JOLs can reactively alter memory itself, a phenomenon referred to as the *memory reactivity effect* (for a review, see Double et al., 2018; Double & Birney, 2019). Even though many studies have been conducted to explore the reactive influences of making JOLs on memory for studied information, to our knowledge, no research has been conducted to investigate whether the reactivity effect can transfer to subsequent learning of new information. The current study aims to fill this gap. Exploring the transferability of the reactivity effect to learning new information may provide practical implications guiding educational practice.

Below we briefly summarise empirical findings of the reactivity effect, then introduce two theories proposed to

CONTACT Chunliang Yang chunliang.yang@bnu.edu.cn Institute of Developmental Psychology, Faculty of Psychology, Beijing Normal University, 19 Xinjiekou Wai Street, Beijing 100875, People's Republic of China; Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education, Beijing Normal University, 19 Xinjiekou Wai Street, Beijing 100875, People's Republic of China

explain the effect, and finally provide an overview of the current study.

Reactivity effects of making JOLs

It has been shown that many forms of metacognitive judgments can reactively change the very entity being judged (for a review, see Double & Birney, 2019). A concrete example of reactivity is the effect of making JOLs on memory (Double et al., 2018; Janes et al., 2018; Li et al., 2021; Mitchum et al., 2016; Myers et al., 2020; Rivers et al., 2021; Soderstrom et al., 2015; Witherby & Tauber, 2017; Zhao et al., 2022). For instance, in Soderstrom et al.'s (2015) Experiment 1, participants were randomly divided to two groups (JOLs *vs*. no-JOL) and instructed to study a list of word pairs, half of which were strongly related (e.g., *blunt-sharp*) and the other half were weakly related (e.g., *boxer-terrible*). The total exposure time for each pair was 8 s for both the JOL and no-JOL groups. For the JOL group, at the last 4 sec for each pair, participants were asked to make a JOL to predict the likelihood that they would remember the pair in a later memory test. By contrast, the no-JOL group did not need to make JOLs. In a later cued-recall test, the JOL group recalled statistically more strongly related pairs and numerically more weakly related pairs than the no-JOL group, demonstrating a reactivity effect (for related findings, Janes et al., 2018; Mitchum et al., 2016; Myers et al., 2020; Tekin & Roediger, 2020; Witherby & Tauber, 2017).

Results from previous research demonstrated that making JOLs enhances memory for related word pairs (Janes et al., 2018; Soderstrom et al., 2015), word lists (Li et al., 2021; Zhao et al., 2022), and pictures (Shi et al., 2022), but has minimal influences on memory for text passages (Ariel et al., 2021) and unrelated word pairs (Janes et al., 2018; Rivers et al., 2021). It has also been observed that making JOL reactively facilitates retention for children (Zhao et al., 2022) and young adults (Janes et al., 2018; Mitchum et al., 2016; Rivers et al., 2021), but the reactivity effect was not observed in older adults (Tauber & Witherby, 2019).

Putative mechanisms underlying positive reactivity

As discussed above, many previous studies have documented that making JOLs can reactively enhance memory for related pairs and word lists (for a review, see Double et al., 2018). Several theories have been proposed to explain the positive reactivity effect of JOLs, including the *enhanced-engagement theory*, and the *cue-strengthening theory*.

The enhanced-engagement theory assumes that positive reactivity is derived from enhanced engagement induced by the requirement of making JOLs (Tauber & Witherby, 2019; Tekin & Roediger, 2020; Zhao et al., 2022). Specifically, participants' attention gradually wanes and their mind wandering systematically increases across a pro-longed learning task (Seli et al., 2016). The requirement of making item-by-item JOLs should reduce (or even prevent) attention waning and enhance learning engagement. That is, participants have to engage in the ongoing learning task in order to make an appropriate JOL for each item, and the enhanced learning engagement in turn produces superior memory gains (for related discussion, see Tauber & Witherby, 2019). Supporting evidence for this theory comes from Shi et al. (2022), which showed that making JOLs significantly reduced mind wandering during the encoding phase, and the reduced mind wandering statistically mediated the positive reactivity effect on memory for images.

Another available explanation is the cue-strengthening theory (Soderstrom et al., 2015), which was proposed to explain the reactivity effect on learning of word pairs. Soderstrom et al. (2015) observed that making JOLs improved cued recall of strongly related word pairs, but had little influence on cued recall of unrelated word pairs. These researchers assumed that participants had to search for "diagnostic" cues to make a reasonable JOL for each pair (Dunlosky & Matvey, 2021; Koriat, 1997). The activated cues, induced by the requirement of making JOLs, in turn strengthened the relatedness between the cue and the target for related word words, producing a positive reactivity effect. By contrast, because there is no pre-existing relatedness between the cue and the target for unrelated word pairs, making JOLs, therefore, fail to benefit recall of unrelated word pairs, leading to no reactivity effect (e.g., Soderstrom et al., 2015).

Transfer of positive reactivity

Previous studies demonstrated that some effective study strategies, such as retrieval practice, can not only improve memory for studied materials (e.g., Roediger & Karpicke, 2006; Yang et al., 2021) but also facilitate subsequent learning of new information (e.g., Kliegl & Bäuml, 2021; Yang et al., 2018). Additionally, in the domain of metacognition, Bonder and Gopher (2019) found that reporting trial-by-trial confidence ratings (i.e., confidence ratings about response correctness) reactively enhanced decision accuracy, and the positive reactivity effect of confidence ratings on decision accuracy successfully transferred to a subsequent task, in which there was no need to report response confidence. Furthermore, Birney et al. (2017) demonstrated that, compared with a control condition in which performance was largely constant across the entire process of the task, the prior experience of making confidence rating impeded subsequent performance across a reasoning task. That is, making confidence ratings reactively alter task trajectories. These findings imply that the reactivity effect of retrospective metacognitive judgments (i.e., confidence ratings) is transferable to a subsequent task.

Although previous studies found that making JOLs can reactively facilitate memory for studied related word pairs and word lists (for a review, see Double et al., 2018), it has not been explored whether these positive reactivity effects transfer to subsequent learning of new information. The aim of the current study is to fill this gap. We are also interested in whether learners would actively adopt the strategy of making JOLs to learn new information in self-regulated learning settings.

It is reasonable to expect that positive reactivity is transferable. Previous studies have consistently found that when participants realise that a given learning strategy can improve learning performance and they have adopted this strategy in a prior learning session, they will continue using this learning strategy in a subsequent learning task (Sun et al., 2022; Yan et al., 2016).[1] For instance, it has been well documented that interleaved presentation of category exemplars is more effective for inductive learning than massed presentation, a phenomenon known as the interleaving effect (Carpenter et al., 2012; Feng et al., 2019; Zhao et al., 2015). Sun et al. (2022) found that when participants realised the benefits of interleaved learning (by comparison with blocked learning), they actively adopted this strategy to learn new categories in a new learning task. Zhao et al. (2022) recently demonstrated that learners do metacognitively appreciate the benefits of making JOLs for learning and memory. Hence, these findings point to the inference that positive reactivity may be transferable to new learning. For instance, even when there is no explicit requirement of making JOLs in a sequent learning task, participants may covertly make JOLs to facilitate their learning because they know that making JOLs is beneficial (Zhao et al., 2022).

It has to be noted that there are also reasons to expect limited transfer of positive reactivity. As discussed above, both the enhanced-engagement and cue-strengthening theories assume that positive reactivity is a task-specific phenomenon (i.e., positive reactivity is induced by the specific task requirement of making JOLs), and when there is no requirement of making JOLs in a subsequent new learning task, positive reactivity would disappear, leading to no transfer. For instance, according to the enhanced-engagement theory, participants have to sustain their attention on the ongoing learning task in order to make reasonable item-by-item JOLs, and enhanced-engagement theory is induced by the specific task requirement of making JOLs. In the same way, the cue-strengthening theory claims that the cues activated to inform JOLs strengthen cue-target relations for related word pairs, in turn leading to positive reactivity. Clearly, both theories propose that positive reactivity is caused by the specific requirement of making JOLs. When there is no requirement of making JOLs in a subsequent learning task, positive reactivity would disappear. Overall, neither the enhanced-engagement nor the cue-strengthening theories provide strong evidence in favour of transfer of positive reactivity to subsequent learning of new information.

In brief, it is difficult to make a clear prediction regarding whether positive reactivity is transferable to learning of new information, and this question has not been explored by far. The current study aims to fill this gap by examining whether making item-by-item JOLs in a prior learning session can promote subsequent learning of new information.

Overview of the current study

To explore the transferability of positive reactivity, the current study asked two (JOL vs. no-JOL) groups of participants to study two blocks of unrelated words (Experiment 1) or related word pairs (Experiment 2). The JOL group made a JOL for each study item in Block 1. By contrast, the no-JOL group did not make JOLs in Block 1. Then both group studied Block 2, during which both groups did not make JOLs. To foreshadow, Experiments 1 and 2 consistently observed a positive reactivity effect in Block 1, but there was little difference in test performance in Block 2, indicating limited transfer of positive reactivity to learning of new information.

A possible explanation for the limited transfer findings observed in Experiments 1 and 2 was that participants in the JOL group did not realise the benefits of making JOLs before studying Block 2. Experiment 3 was conducted to test this explanation. Another aim of Experiment 3 was to explore a potential intervention to promote the transfer of positive reactivity. To achieve these two aims, Experiment 3 asked two (JOL vs. no-JOL) groups of participants to complete two learning tasks. In Task 1, the JOL and no-JOL groups studied two blocks of related word pairs, with the JOL group making JOLs for one block but not for the other, and the no-JOL group not making JOLs for any block. Task 1 was implemented to let the JOL group experience positive reactivity first hand, in turn improving their metacognitive appreciation of the benefits of making JOLs. Then, both groups completed Task 2, in which they studied new pairs without the requirement of making JOLs. We expect that firsthand experience of positive reactivity in Task 1 would encourage the JOL group to covertly make JOLs in Task 2, producing superior learning performance in Task 2 in the JOL than in the no-JOL group (Sun et al., 2022; Yan et al., 2016).

Experiment 1

Prior research consistently found that making JOLs reactively enhances the recognition of word lists (Li et al., 2021; Shi et al., 2022; Zhao et al., 2022). Experiment 1 employed word lists as study materials to explore whether the positive reactivity effect on learning of a prior word list transfers to subsequent learning of a new word list.

Method

Participants

A pilot study, with six participants in each group, was conducted to determine the required sample size, and the experimental procedure in the pilot study was identical to that in the formal experiment. The pilot results showed that the effect size of the interaction was Cohen's $d = 0.713$. A power analysis, conducted via G*Power (Faul et al., 2007), showed that about 32 participants were required in each group to observe a significant interaction at 0.80 power. Finally, 67 participants (*M* age = 20.731, *SD* = 2.890; 66 females) were recruited from the Beijing Normal University (BNU) participant pool, with 34 randomly allocated into the JOL group and 33 into the no-JOL group.

All participants signed an agreement to participate, were tested individually in a sound-proofed cubicle, and received financial remuneration. The protocol was approved by the Institutional Review Board of BNU Faculty of Psychology.

Materials

The stimuli were 352 two-character Chinese words extracted from the Chinese word database developed by Cai and Brysbaert (2010). The word frequency of these words ranged from 0.03 to 19.44 per million. Thirty-two words were used for practice and the other 320 words were used in the formal experiment. In the formal experiment, 160 words were studied during the learning task and served as "old" items in the forced-choice recognition test, with the other 160 words as "new" items.

To prevent any item-selection effects, for each participant, the 160 words were randomly divided into two lists, with 80 words in each list. The two lists were randomly assigned to Blocks 1 and 2. In addition, the presentation sequence of words in each block was randomly decided by the computer for each participant. All stimuli were presented via the Matlab *Psychtoolbox* package (Kleiner et al., 2007).

Design and procedure

The experiment involved a 2 (group: JOL *vs.* no-JOL) × 2 (block: Block 1 *vs.* Block 2) mixed design. Group was a between-subject variable, and block was a within-subject variable (see Figure 1).

Before the formal experiment, participants completed a practice task to familiarise the experimental procedure (see Appendix A for experimental instructions). The procedure of the practice task was the same as that of the main experiment (see below for details).

In the formal experiment, participants studied two blocks of words, with 80 words in each block. Participants in the JOL group were informed that they would need to make memory predictions for words in Block 1 but did not need to make such predictions in Block 2. Critically, they were asked to remember as many words as they could in both blocks because all words would be finally tested, regardless of whether they needed to make predictions or not. By contrast, participants in the no-JOL group were instructed to remember as many words in both blocks as they could in preparation for a subsequent memory test, and they did not receive instructions about making JOLs.

For the no-JOL group, participants studied the words one-by-one and block-by-block. In Block 1, the 80 words were presented one-by-one in random order. Before the presentation of each word, a cross sign appeared at the centre of the screen for 0.5 s to mark the inter-stimulus interval. Immediately following, a word appeared on the screen for 6 s in total.[2] Then, the next trial started. This cycle repeated until the end of the block, with a new word studied in each cycle. After studying Block 1, participants in the no-JOL group were instructed to study Block 2. The procedure of Block 2 was the same as that for Block 1, except that participants studied 80 new words in Block 2.

The procedure for the JOL group's Block 1 was similar to that for the no-JOL group's Block 1, except that participants needed to make item-by-item JOLs while studying each word. Specifically, in Block 1, each word was first presented for 3 s, following which the word remained on screen for another 3 s with a slider presented below it (see Figure 1). Participants were instructed to predict the likelihood that they would remember the word in a later memory test on a slider ranging from 0 (*Sure I will not remember it*) to 100 (*Sure I will remember it*). The scale was presented for 3 s, and participants made their JOLs by dragging and clicking the scale pointer. If they successfully made a JOL within the 3 s time-window, the word remained on screen for the left duration of the 3 s to ensure that the total exposure duration for each word was 6 s. If they did not successfully make a JOL during the required time-window, a message box appeared onscreen to remind them to carefully make predictions for the following words during the required time-window. Participants clicked the mouse to remove the message box and trigger the next trial. The procedure for the JOL group's Block 2 was identical to that for the no-JOL group's Block 2. That is, the 80 words were presented one-by-one in a random order, for 6 s each, and participants did not need to make item-by-item JOLs.

After participants studied both blocks, they solved math problems (e.g., *16 + 45 = ___?*) for 5 min, which served as a distractor task. After that, they took a forced-choice recognition test. Specifically, the 160 studied and 160 new words were randomly paired to form 160 pairs, with each pair consisting of an "old" and a "new" word. The pairs were presented one-by-one in random order. Before presenting each pair, a cross sign was presented for 0.5 s. Next, the two words were randomly allocated to the left and the right side of the screen. Participants

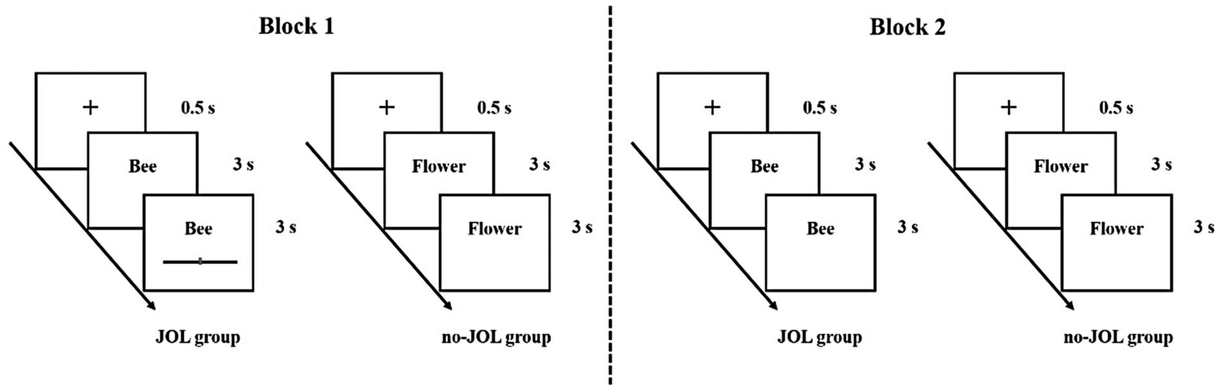

Figure 1. The experimental flowchart for Experiment 1.

were instructed to indicate which word was "old". When a recognition choice was made, the next test trial started automatically. There was no time pressure and no feedback in the forced-choice recognition test.

Results

The primary research interest was to explore whether the reactivity effect on learning of Block 1 can transfer to subsequent learning of Block 2, and hence test performance results are presented below. Item-by-item JOLs were not the focus of the current study and hence are reported in Appendix B.

Recognition performance in each condition is depicted in Figure 2. A Bayesian mixed analysis of variance (ANOVA) was conducted via JASP Version 0.15.0.0 (https://jasp-stats.org), with all parameters set as default (Goss-Sampson, 2019; van Doorn et al., 2021). Bayes Factors for the alternative over the null hypothesis (BF_{10}) can be interpreted as relative evidence in favour of the alternative hypothesis (i.e., evidence supporting the existence of an effect over the absence of the effect).

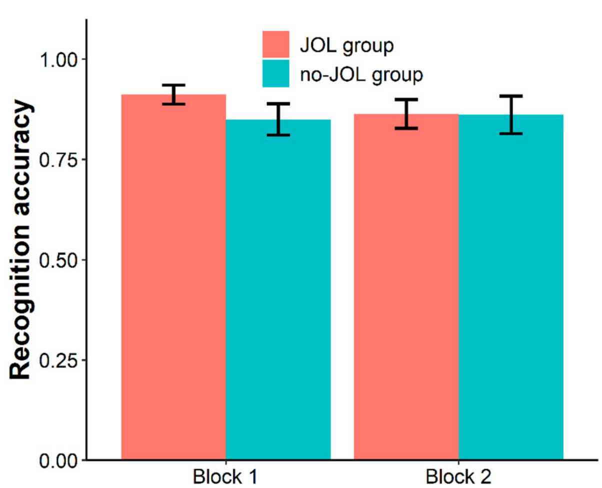

Figure 2. Recognition accuracy as a function of group and block in Experiment 1. Error bars represent 95% CI.

The Bayesian mixed ANOVA, with block as the within-subjects variable, group as the between-subjects variable, and recognition performance as the dependent variable, showed no main effect of block, $F(1, 65) = 3.634$, $p = .061$, $\eta_p^2 = .053$, $BF_{10} = 0.838$, and no main effect of group, $F(1, 65) = 1.650$, $p = .203$, $\eta_p^2 = .025$, $BF_{10} = 0.690$. Of critical interest, there was a significant interaction between block and group, $F(1, 65) = 9.592$, $p = .003$, $\eta_p^2 = .129$, $BF_{10} = 12.500$.

A pre-planned Bayesian independent *t*-test showed a significant difference in Block 1 recognition performance between the JOL ($M = 0.911$, $SD = 0.070$) and the no-JOL group ($M = 0.850$, $SD = 0.115$), difference = 0.062 [0.016, 0.108], $t(65) = 2.666$, $p = .010$, $d = 0.652$, $BF_{10} = 4.753$, reflecting a positive reactivity effect on learning of word lists. By contrast, there was little difference in Block 2 recognition performance between the JOL ($M = 0.864$, $SD = 0.107$) and the no-JOL group ($M = 0.861$, $SD = 0.137$), difference = 0.003 [−0.057, 0.063], $t(65) = 0.087$, $p = .931$, $d = 0.021$, $BF_{10} = 0.251$, implying that the positive reactivity effect failed to transfer to learning of new words.

A Bayesian paired *t*-test showed that, in the JOL group, performance for Block 1 words ($M = 0.911$, $SD = 0.070$) was significantly better than that for Block 2 words ($M = 0.864$, SD = 0.107), difference = 0.048 [0.018, 0.078], $t(33) = 3.223$, $p = .003$, $d = 0.553$, $BF_{10} = 12.726$, re-confirming the positive reactivity effect on memory for word lists (Li et al., 2021; Zechmeister & Shaughnessy, 1980; Zhao et al., 2022). Twenty-one participants showed a positive reactivity effect, 11 showed the converse pattern, and the other two were ties. In contrast, in the no-JOL group, there was no statistically detectable difference in recognition performance between Blocks 1 ($M = 0.850$, $SD = 0.115$) and 2 ($M = 0.861$, $SD = 0.137$), difference = 0.012 [−0.013, 0.036], $t(32) = 0.952$, $p = .348$, $d = 0.173$, $BF_{10} = 0.283$.

Discussion

Akin to prior research, Experiment 1 found that making concurrent JOLs significantly enhanced memory for word lists (Li et al., 2021; Zhao et al., 2022). More importantly, Experiment 1 observed minimal difference in recognition performance in Block 2, indicating that the positive reactivity effect does not transfer to learning of a new word list.

Experiment 2

It is premature to make a firm conclusion based on results from a single experiment. Hence, Experiment 2 was conducted to conceptually replicate the main findings of Experiment 1. Prior research consistently found that making JOLs reactively promotes cued recall of related word pairs (Janes et al., 2018; Li et al., 2021; Rivers et al., 2021; Soderstrom et al., 2015; Witherby & Tauber, 2017). Experiment 2 hence employed related word pairs as study stimuli to further test whether the positive reactivity effect is transferable.

Method

Participants

A pilot study, with six participants in each group, was conducted to determine the required sample size, and the experimental procedure in the pilot study was identical to that in the formal experiment. The pilot results showed that the effect size for the interaction between group (JOL *vs.* no-JOL) and block (Block 1 *vs.* Block 2) was Cohen's $d = 0.655$. A power analysis, conducted via G*Power (Faul et al., 2007), showed that 38 participants in each group were required to observe a significant interaction at 0.80 power. Finally, 76 participants were recruited from Beijing Normal University (BNU), with a mean age of 20.908 ($SD = 1.955$) years, and 71 females. They were randomly allocated to each group, with 38 in each group.

All participants signed agreements to participate, were tested individually in a sound-proofed cubicle, and received financial remuneration. The protocol was approved by the Institutional Review Board of BNU Faculty of Psychology.

Materials

The stimuli were 100 semantically related Chinese word pairs (e.g., *mouse-keyboard*) selected from Hu et al. (2016). Hu and colleagues asked participants to rate the semantic relatedness of the word pairs on a scale ranging from 1 (*completely unrelated*) to 4 (*strongly related*). The average of relatedness rating for the selected word pairs was 3.416 ($SD = 0.262$). Eighty pairs were used in the formal experiment, with the other 20 used for practice.

To prevent any item-selection effects, for each participant, the 80 pairs were randomly divided into two lists, with 40 pairs in each list. For each participant, the two lists were randomly assigned to Blocks 1 and 2. In addition, the present sequence of pairs in each block was randomly decided by computer for each participant. All stimuli were presented via the Matlab *Psychtoolbox* package (Kleiner et al., 2007).

Design and procedure

The experiment involved a 2 (group: JOL *vs.* no-JOL) × 2 (block: Block 1 *vs.* Block 2) mixed design. Group was a between-subjects variable, and block was a within-subjects variable (see Figure 3).

The procedure in Experiment 2 was similar to that in Experiment 1. Specifically, in the JOL group, participants were instructed to make item-by-item JOLs in Block 1, and they did not need to do that in Block 2. They were encouraged to remember as many word pairs as they could in both blocks, because all word pairs would be finally tested. In the no-JOL group, participants were informed to remember as many word pairs as possible in both blocks, and they did not receive instructions about making memory predictions.

For the no-JOL group's Blocks 1 and 2, each word pair was presented on the screen for 8 s in total for participants to study. For the JOL group's Block 1, each word pair was firstly presented on the screen for 4 s, following which the word pair remained on the screen for another 4 s with a slider presented below it. Participants were instructed to predict the likelihood they would remember the word pair in a later memory test on a slider ranging from 0 (*Sure I will not remember it*) to 100 (*Sure I will remember it*) by dragging and clicking the slider to make a JOL during the 4 s time window. For the JOL group's Block 2, each word pair was presented for 8 s, and participants did not need to make JOLs.

After participants studied both blocks, they solved math problems (e.g., *16 + 45 = ___?*) for 5 min, which served as a distractor task. Then both groups completed a cued recall test on all word pairs. Specifically, the 80 cue words were presented one-by-one in a random order, and participants were required to recall the corresponding targets. There was no time pressure and no feedback in the cued recall test.

Results

Cued recall performance is depicted in Figure 4. A Bayesian mixed ANOVA, with block as the within-subject variable, with group as the between-subject variable, and recall performance as the dependent variable, was conducted. The results revealed a main effect of block, $F(1, 74) = 23.227$, $p < .001$, $\eta_p^2 = .239$, $BF_{10} = 1122.995$. There was no main effect of group, $F(1, 74) = 1.493$, $p = .226$, $\eta_p^2 = .020$, $BF_{10} = 0.657$. Of critical interest, the interaction between block and group was significant, $F(1, 68) = 6.900$, $p = .010$, $\eta_p^2 = .085$, $BF_{10} = 4.202$ (see Figure 4).

A pre-planned Bayesian independent *t*-test showed a significant difference in Block 1 recall performance between the JOL ($M = 0.859$, $SD = 0.111$) and the no-JOL

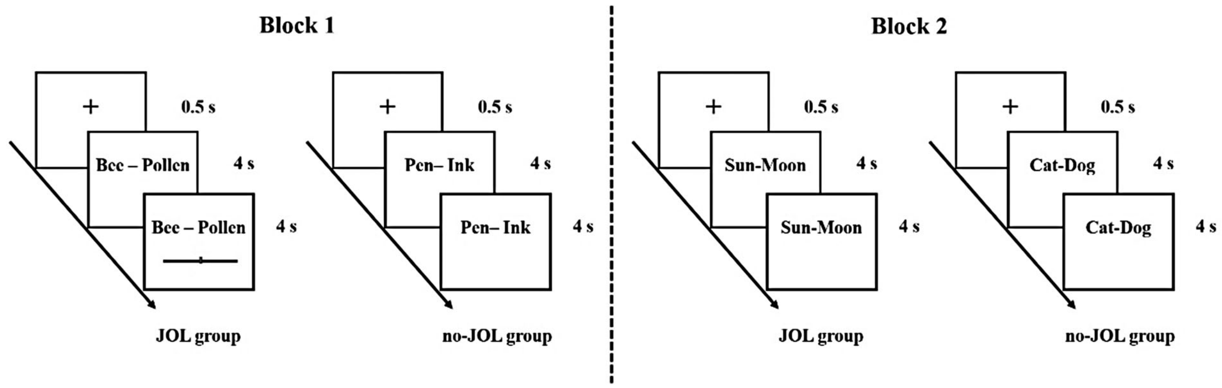

Figure 3. The experimental flowchart for Experiment 2.

group ($M = 0.779$, $SD = 0.199$), difference = 0.080 [0.006, 0.153], $t(74) = 2.149$, $p = .037$, $d = 0.493$, $BF_{10} = 1.697$, replicating the classic positive reactivity effect of making JOLs on memory for related word pairs (Double et al., 2018). By contrast, there was minimal difference in Block 2 recall performance between the JOL ($M = 0.767$, $SD = 0.180$) and the no-JOL group ($M = 0.752$, $SD = 0.203$), difference = 0.015 [−0.073, 0.103], $t(74) = 0.344$, $p = .732$, $d = 0.079$, $BF_{10} = 0.250$, reflecting limited transfer of positive reactivity toward learning of new related word pairs.

A pre-planned Bayesian paired *t*-test showed that, in the JOL group, word pairs in Block 1 ($M = 0.859$, $SD = 0.111$) were recalled better than those in Block 2 ($M = 0.767$, $SD = 0.180$), difference = 0.091 [0.050, 0.133], $t(37) = 4.848$, $p < .001$, $d = 0.786$, $BF_{10} = 959.598$, re-confirming the positive reactivity effect on memory for related word pairs (Janes et al., 2018; Rivers et al., 2021; Zhao et al., 2022). Twenty-seven participants showed a positive reactivity effect, seven showed the converse pattern, and the other four were ties. By contrast, in the no-JOL group, there was minimal difference in test performance between Blocks 1 ($M = 0.779$, $SD = 0.199$) and 2 ($M = 0.752$, $SD = 0.203$), difference = 0.027 [−0.005, 0.059], $t(37) = 1.718$, $p = .094$, $d = 0.279$, $BF_{10} = 0.664$.

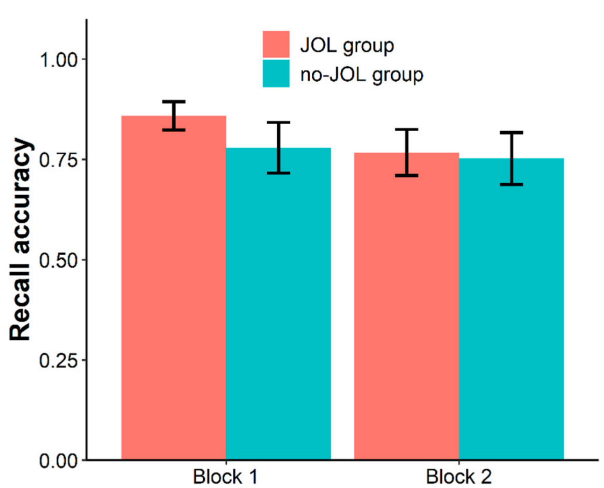

Figure 4. Recall accuracy as a function of group and block in Experiment 2. Error bars represent 95% CI.

Discussion

Akin to prior research, Experiment 2 found that making concurrent JOLs significantly enhanced memory for related word pairs (Janes et al., 2018; Li et al., 2021; Rivers et al., 2021; Soderstrom et al., 2015; Witherby & Tauber, 2017). Consistent with Experiment 1, Experiment 2 observed Bayesian evidence supporting no transfer of positive reactivity to learning of new related pairs.

Experiment 3

Experiments 1 and 2 consistently demonstrated limited transfer of positive reactivity to learning of new word lists and related word pairs. A possible explanation is that participants in the JOL group might lack metacognitive awareness about the benefits of making JOLs. If they did not know that making JOLs facilitated their learning, they would be reluctant to use this strategy (i.e., covertly making JOLs) when learning new materials, leading to minimal transfer of positive reactivity. We term this explanation as the *lack of awareness* explanation.

To test this explanation, Experiment 3 had the JOL group to personally experience the benefits of making JOLs before they studied new information. Specifically, in Experiment 3's Task 1, the JOL group studied related pairs in a JOL and a no-JOL block, and then completed a cued-recall test. Task 1 was included to let the JOL group get firsthand experience of positive reactivity and enhance their metacognitive appreciation of the benefits of making JOLs. After completing Task 1, the JOL and no-JOL groups studied new related pairs in Task 2. Prior research did show that firsthand experience of the benefits of a given learning strategy can enhance metacognitive awareness and promote the application of the strategy in a subsequent learning task (Sun et al., 2022). Accordingly, we expect that firsthand experience of positive reactivity in Task 1 would encourage the JOL group to covertly make JOLs in Task 2, leading to superior recall performance in the JOL than in the no-JOL group (i.e., successful transfer of positive reactivity).

Method

Participants

A pilot study, with nine participants in each group, was conducted to determine the required sample size, and the experimental procedure in the pilot study was identical to that in the formal experiment. The pilot results showed that the effect size for the interaction between group (JOL *vs.* no-JOL) and block (Block 1 *vs.* Block 2) was Cohen's $d = 0.840$. A power analysis, conducted via G*Power (Faul et al., 2007), showed that 24 participants in each group were required to observe a significant interaction at 0.80 power. Finally, 72 participants (*M* age = 22.305, *SD* = 2.307; 63 females) were recruited from \BNU participant pool, with 36 randomly allocated to each group. All participants signed agreements to participate, were tested individually in a sound-proofed cubicle, and received financial remuneration. The protocol was approved by the Institutional Review Board of BNU Faculty of Psychology.

Materials

The stimuli were 140 semantically related Chinese word pairs, selected from Hu et al. (2016). The average relatedness rating for the selected word pairs was 3.300 (*SD* = 0.279). One hundred and twenty pairs were used in the formal experiment, with the other 20 used for practice.

Experiment 3 contained two learning tasks. To prevent any item-selection effects, the 120 pairs were randomly divided into two lists, with 60 pairs in each list. The two lists were randomly assigned to Tasks 1 and 2. In Task 1, the word pairs were randomly assigned to Blocks 1 and 2, with 30 pairs in each block. In addition, the present sequence of pairs in each block, the block sequence, and the list sequence was randomly decided by computer for each participant. All stimuli were presented via the Matlab *Psychtoolbox* package (Kleiner et al., 2007).

Design and procedure

Task 1 involved a 2 (group: JOL *vs.* no-JOL) × 2 (block: Block 1 *vs.* Block 2) × 2 (task: Task 1 *vs.* Task 2) mixed design. Group was a between-subject variable, and block and task were within-subject variables.

The procedure of Task 1 was same as in Experiment 1, except that participants studied 30 related pairs in each block. That is, in Block 1, the JOL group studied 30 related pairs and made item-by-item JOLs, whereas the no-JOL group studied those word pairs without making JOLs. Next, both groups studied another 30 related pairs in Block 2, in which they did not make JOLs. Then, they attended to a 5-min distractor task, and completed a cued recall test on the 60 studied word pairs.

After the completion of Task 1, each participant in the JOL group was shown his/her own memory performance for JOL and no-JOL pairs. We provided participants their own test performance with the aim to enhance their awareness about the benefits of making JOLs. After viewing their own test performance, they were asked to report which strategies helped them learn better. They made their choices by choosing one from two options: (1) Making JOLs, and (2) Not making JOLs. This question was implemented to measure participants' awareness of the positive reactivity effect.

Different from the JOL group, participants in the no-JOL group were provided a learning scenario in that some students are preparing for a course exam, with some materials studied through massed learning (i.e., studying the materials for 3 h in a single day) and others studied through spaced learning (i.e., studying the materials for one hour in each of three successive days). They were asked to report which study strategy helps them study better. They made their choices by choosing one of two options: (1) Spaced, and (2) Massed. This question is a filler question, implemented to make the task procedure between the JOL and no-JOL groups as similar as possible. Note that the no-JOL group had no experience of making JOLs in Task 1, which made it impossible to measure their awareness of the reactivity effect.

After answering the awareness question, both groups engaged in the second learning task. In Task 2, both groups studied 60 new related word pairs (8 sec each) without making JOLs, completed a 5-min distractor task, and then took a cued recall test on them. The second learning task was administered to investigate whether the positive reactivity effect transfers to learning of new related word pairs when participants explicitly knew that making JOLs was beneficial for learning.

Results

Test performance in Task 1

Recall performance in each condition is depicted in Figure 5. A mixed Bayesian ANOVA, with block as the within-subjects variable, group as the between-subjects variable, and recall performance in Task 1 as the dependent variable, showed no main effect of block, $F(1, 70) = 4.819$, $p = .031$, $\eta_p^2 = .064$, $BF_{10} = 1.409$, and no main effect of group, $F(1, 70) = 2.696$, $p = .105$, $\eta_p^2 = .037$, $BF_{10} = 1.053$. Of critical interest, there was a significant interaction between block and group, $F(1, 70) = 7.966$, $p = .006$, $\eta_p^2 = .102$, $BF_{10} = 6.525$.

A pre-planned Bayesian independent *t*-test showed a significant difference in Block 1 recall performance between the JOL ($M = 0.806$, $SD = 0.155$) and the no-JOL group ($M = 0.689$, $SD = 0.233$), difference = 0.117 [0.024, 0.210], $t(70) = 2.499$, $p = .015$, $d = 0.589$, $BF_{10} = 3.345$, reflecting a positive reactivity effect on learning of

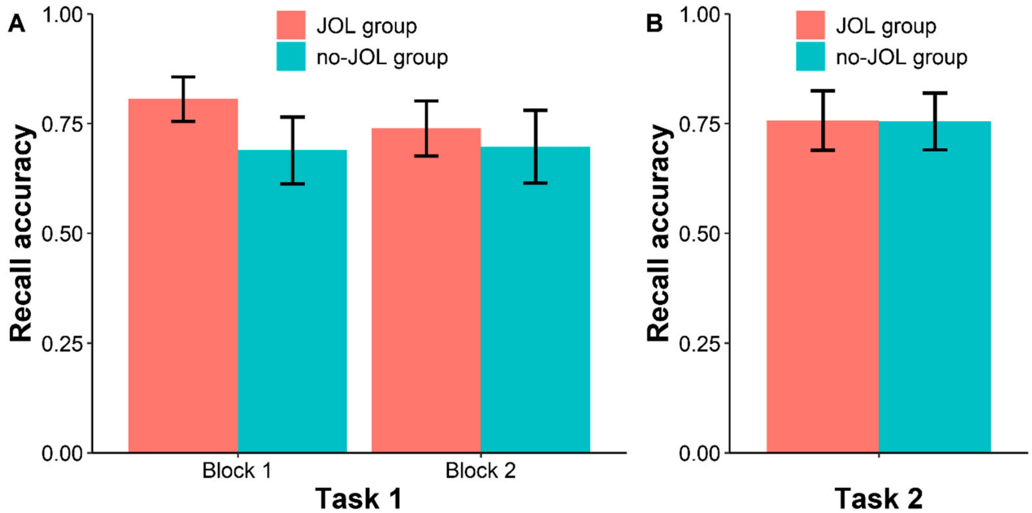

Figure 5. Panel A: Recall accuracy as a function of group and block in Experiment 3's Task 1. Panel B: Recall accuracy as a function of group in Experiment 3's Task 2. Error bars represent 95% CI.

related pairs. By contrast, there was little difference in Block 2 recall performance between the JOL ($M = 0.739$, $SD = 0.193$) and the no-JOL group ($M = 0.697$, $SD = 0.254$), difference = 0.042 [−0.064, 0.148], $t(65) = 0.784$, $p = .436$, $d = 0.185$, $BF_{10} = 0.317$, implying that the positive reactivity effect failed to transfer to learning of new related pairs in Task 1's Block 2 and replicating the main findings of Experiments 1 and 2.

A Bayesian paired *t*-test showed that, in the JOL group, recall performance for Block 1 word pairs ($M = 0.806$, $SD = 0.155$) was significantly better than that for Block 2 word pairs ($M = 0.739$, $SD = 0.193$), difference = 0.067 [0.026, 0.107], $t(35) = 3.369$, $p = .002$, $d = 0.562$, $BF_{10} = 18.305$, replicating the positive reactivity effect on learning of related pairs. Twenty-one participants showed a positive reactivity effect, nine showed the converse pattern, and the other six were ties. In contrast, in the no-JOL group, there was no statistically detectable difference in recall performance between Block 1 ($M = 0.689$, $SD = 0.233$) and 2 ($M = 0.697$, $SD = 0.254$), difference = −0.008 [−0.044, 0.028], $t(35) = -0.470$, $p = .641$, $d = -0.078$, $BF_{10} = 0.198$.

Overall, the above results successfully replicated the main findings of Experiments 1 and 2 by showing that the positive reactivity effect did not transfer to subsequent learning of new information.

Metacognitive awareness

For the JOL group, according to participants' responses to the awareness question about reactivity, they were classified into two categories: (1) JOL > no-JOL (i.e., participants who believed that making JOLs helped them study better) and (2) JOL < no-JOL (i.e., participants who believed that not making JOLs helped them study better). The results showed that 66.7% (24 out of 36) of participants believed JOL > no-JOL, which was significantly greater than the proportion (33.3%) of participants believing JOL < no-JOL, $\chi^2(1) = 4$, $p = .046$ (for related findings, see Zhao et al., 2022). These results reflect that, before the initiation of Task 2, most of participants in the JOL group did believe that making JOLs is beneficial for learning of related word pairs.[3]

For the no-JOL group, based on participants' responses, they were classified into two categories: (1) Spaced > Massed and (2) Spaced < Massed. The results showed that 52.8% (19 out of 36) of participants believed Spaced > Massed, which was statistically different from the proportion (47.2%) of participants who believed Spaced < Massed, $\chi^2(1) = 0.111$, $p = .739$.

Recall performance in Task 2

A pre-planned Bayesian independent *t*-test showed a little difference in recall performance between the JOL ($M = 0.757$, $SD = 0.208$) and the no-JOL group ($M = 0.755$, $SD = 0.198$), difference = 0.002 [−0.093, 0.098], $t(70) = 0.048$, $p = .962$, $d = 0.011$, $BF_{10} = 0.243$, implying that the positive reactivity effect failed to transfer to learning of new related pairs in Task 2 even when participants in the JOL group explicitly knew that making JOL enhanced their learning.

Before the initiation of Task 2, 12 out of 36 participants in the JOL group believed JOL < no-JOL. To be more careful, we conducted another Bayesian independent *t*-test, in which the results from these 12 participants were excluded, leaving final data from 24 participants in the JOL group who believed JOL > no-JOL. After excluding these 12 participants, there was still no statistically detectable difference in recall performance in Task 2 between the JOL ($M = 0.696$, $SD = 0.187$) and the no-JOL group ($M = 0.755$, $SD = 0.198$), difference = −0.059 [−0.161, 0.043], $t(58) = -1.152$, $p = .254$, $d = -0.304$, $BF_{10} = 0.464$. These results reconfirm limited transferability of the reactivity effect to the learning of new information.

More importantly, these results suggest that limited transferability of the reactivity effect observed in Experiment 2 should not be attributed to a lack of awareness about the benefits of making JOLs.

Discussion

Experiment 3 replicated the main findings of Experiment 2 by showing that making concurrent JOLs induced a positive reactivity effect on learning of related word pairs, and the positive reactivity effect failed to transfer to learning of new related pairs in Task 1's Block 2. More importantly, even when participants in the JOL group explicitly acknowledged the beneficial effects of making JOLs, the positive reactivity effect still failed to transfer to learning related pairs in Task 2.

General discussion

Although many recent studies have consistently demonstrated that making concurrent JOLs can reactively change memory itself (Li et al., 2021; Zhao et al., 2022), no research has been conducted to explore whether the reactivity effect is transferable to subsequent learning of new information. The current study conducted three experiments to explore whether the positive reactivity effects on memory for word lists and related word pairs are transferable.

In the current study, word lists (Experiment 1) or related word pairs (Experiments 2 and 3's Task 1) were studied in two blocks. In Block 1, participants in the JOL group were instructed to make concurrent JOLs, whereas those in the no-JOL group did not make JOLs. In Block 2, both groups did not make JOLs. The principal findings documented in the current study were that making JOLs significantly enhanced memory for word lists (Experiment 1) and related word pairs (Experiments 2 and 3's Task 1) in Block 1, but the positive reactivity effect failed to transfer to learning of Block 2, in which there was no requirement of making concurrent JOLs. To our knowledge, the current study is the first to demonstrate that making concurrent JOLs fails to benefit subsequent learning of new word lists (Experiment 1) and related word pairs (Experiments 2 and 3).

The limited transferability of the positive reactivity effect on learning of word lists observed in Experiment 1 can be explained by the enhanced-engagement theory. The enhanced-engagement theory asserts that to make appropriate JOLs, participants have to focus their attention on the ongoing learning task, and the enhanced engagement in turn boosts learning outcomes (Li et al., 2021; Yang et al., 2015; Zechmeister & Shaughnessy, 1980; Zhao et al., 2022). To be specific, the enhanced-engagement theory assumes that the positive reactivity effect results from the requirement of making JOLs. Without such a requirement, the positive reactivity effect diminishes or disappears. Indeed, Experiment 1 demonstrated that the positive reactivity effects failed to transfer to learning of new words in Block 2.

The limited transferability of the positive reactivity effect documented in Experiments 2 and 3 (Task 1) is consistent with the cue-strengthening theory (Rivers et al., 2021; Soderstrom et al., 2015). Participants had to search for "diagnostic" cues (e.g., the level of relatedness between the cue and the target word) to inform JOL formation (Koriat, 1997; Mueller et al., 2013). The activated cues, induced by the requirement of making JOLs, in turn enhanced the relatedness between the cue and the target for related pairs, producing superior retention. When making JOLs was not required in Block 2, the positive reactivity therefore disappeared. The cue-strengthening theory and enhanced engagement theory are not mutually exclusive, and those two theories jointly assume that positive reactivity is a task-specific phenomenon. Once the requirement of making JOLs is removed, the positive reactivity effect would suspend.

Another possible explanation of the limited transfer findings observed in Experiments 1 and 2 is that participants in the JOL group might lack metacognitive appreciation of the beneficial effect of making JOLs on memory, leading them not to adopt making JOLs as an effective learning strategy in the subsequent learning task. Experiment 3 was conducted to test this lack of awareness explanation and to explore if enhancing metacognitive awareness through firsthand experience can improve the transfer of positive reactivity. The results showed that, even after participants obtained firsthand experience of positive reactivity and explicitly realised the benefits of making JOLs, positive reactivity still failed to transfer to learning of related pairs in Task 2. Such a finding is inconsistent with the lack of awareness explanation.

It is somewhat striking that enhanced metacognitive awareness of the benefits of making JOLs did not promote transfer of the reactivity effect. It is possible that participants in the JOL group still did not make covert JOLs in Task 2, even though they clearly realised the benefits of making JOLs. For instance, it is well-known that, besides metacognitive awareness, there are many other factors affecting study strategy usage, such as perceived mental effort. Learners are frequently reluctant to adopt mental taxing strategies and prefer to use less effortful ones during self-regulated learning (Kirk-Johnson et al., 2019; Macaluso et al., 2022; McDaniel & Einstein, 2020). As discussed above, participants needed to search for appropriate cues to make a reasonable JOL for each item, and the cue-search process required them to spend extra mental effort. In Experiment 3, even though participants in the JOL group did realise that making JOLs is helpful for learning (Zhao et al., 2022), they might not continue covertly making JOLs when studying new pairs in Task 2 because the JOL-making process was mentally taxing.

Another possible explanation is that enhanced metacognitive awareness might have successfully stimulated participants to make covert JOLs in Task 2, but making covert JOLs is less effective in enhancing learning than making overt JOLs. However, this possibility should be carefully adopted as evidence has shown that whether a memory strategy (e.g., retrieval practice) was implemented covertly or overtly led to an equivalent mnemonic effect (Putnam & Roediger, 2013; Smith et al., 2013), even though the comparison between covert and overt JOL has not yet been experimentally justified. We expect to directly test it in the future.

Overall, the present findings run counter to the lack of awareness explanation. The mechanisms underlying the limited transfer of positive reactivity remain largely unknown. Future research is encouraged to further explore why the positive reactivity effect of making JOLs does not transfer to learning of new information.

The documented findings bear some educational implications. Class curriculum is generally too full to spare time for making item-by-item JOLs, and it is unusual for a teacher to ask students to make a JOL after teaching each knowledge point in real educational settings. Hence, it is important to explore whether positive reactivity is transferable to learning of new information (i.e., whether making JOLs for some items in a prior learning session can promote subsequent learning of new items). The answer from the current study is negative. Specifically, the current study observed that once the requirement of making JOLs disappears, the positive reactivity effect will suspend. Although making JOLs does not enhance learning of new information, it can reactively enhance memory for studied word lists and related word pairs (Li et al., 2021; Shi et al., 2022; Zhao et al., 2022) and the positive reactivity effect is long-lasting (i.e., at least 24 h, see Witherby & Tauber, 2017). Teachers should bear in mind that the positive reactivity effect does not transfer. When the class curriculum is too full and there is no sufficient time for students to make JOLs for all study materials, instructors should consider asking students to make JOLs only for key knowledge points.

Conclusion

Making concurrent JOLs induces a positive reactivity effect on learning of related word pairs and word lists, but the positive reactivity effect does not transfer to subsequent learning of new information when making JOLs is no longer required. Enhancing metacognitive awareness does not promote transfer of positive reactivity. Lack of metacognitive awareness tends to be not responsible for limited transfer of positive reactivity.

Notes

1. It should be noted that, besides metacognitive awareness, other factors (e.g., perceived mental effort) also affect study strategy selection. See the General Discussion for detailed discussion.
2. To avoid a ceiling effect in recognition performance, the presentation duration for each word was set to 6 s (rather than 8 s), which was identical to that of Li et al. (2021).
3. A point-biserial correlation analysis showed a positive correlation between the magnitude of the reactivity effect (i.e., the difference in recall performance between Task 1's Blocks 1 and 2) and metacognitive awareness (with JOL > no-JOL coded as 1 and JOL < no-JOL as 0), $r_{pb} = .436$, $p = .008$, indicating that participants who demonstrated a stronger reactivity effect were more likely to appreciate the benefits of making JOLs.

Author note

The data contained in this project are publicly available at Open Science Framework (https://osf.io/7hjyz/).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by National Natural Science Foundation of China: [grant no 32000742; 32171045]; the Research Program Funds of the Collaborative Innovation Center of Assessment toward Basic Education Quality at Beijing Normal University: [grant no 2022-01-132-BZK01].

ORCID

Baike Li http://orcid.org/0000-0002-9382-9613

References

Ariel, R., Karpicke, J. D., Witherby, A. E., & Tauber, S. K. (2021). Do judgments of learning directly enhance learning of educational materials? *Educational Psychology Review*, *33*, 693–712. https://doi.org/10.1007/s10648-020-09556-8.

Birney, D. P., Beckmann, J. F., Beckmann, N., & Double, K. S. (2017). Beyond the intellect: Complexity and learning trajectories in Raven's Progressive Matrices depend on self-regulatory processes and conative dispositions. *Intelligence*, *61*, 63–77. https://doi.org/10.1016/j.intell.2017.01.005

Bonder, T., & Gopher, D. (2019). The effect of confidence rating on a primary visual task. *Frontiers in Psychology*, *10*, 2674. https://doi.org/10.3389/fpsyg.2019.02674

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, *5*, e10729. https://doi.org/10.1371/journal.pone.0010729.

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review*, *24*, 369–378. https://doi.org/10.1007/s10648-012-9205-z.

Double, K. S., & Birney, D. P. (2019). Reactivity to measures of metacognition. *Frontiers in Psychology*, *10*, 27–55. https://doi.org/10.3389/fpsyg.2019.02755

Double, K. S., Birney, D. P., & Walker, S. A. (2018). A meta-analysis and systematic review of reactivity to judgements of learning. *Memory (Hove, England)*, *26*, 741–750. https://doi.org/10.1080/09658211.2017.1404111.

Dunlosky, J., & Hertzog, C. (1997). Older and younger adults use a functionally identical algorithm to select items for restudy during multitrial learning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *52*, 178–186. https://doi.org/10.1093/geronb/52B.4.P178

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*, 1180–191. https://doi.org/10.1037/0278-7393.27.5.1180

Dunlosky, J., & Thiede, K. W. (2004). Causes and constraints of the shift-to-easier-materials effect in the control of study. *Memory & Cognition*, *32*, 779–788. https://doi.org/10.3758/BF03195868.

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.

Feng, K., Zhao, X., Liu, J., Cai, Y., Ye, Z., Chen, C., & Xue, G. (2019). Spaced learning enhances episodic memory by increasing neural pattern similarity across repetitions. *The Journal of Neuroscience*, *39*, 5351. https://doi.org/10.1523/JNEUROSCI.2741-18.2019.

Finn, B. (2008). Framing effects on metacognitive monitoring and control. *Memory & Cognition*, *36*, 813–821. https://doi.org/10.3758/MC.36.4.813.

Goss-Sampson, M. (2019). *Statistical analysis in JASP: A guide for students*: JASP.

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.

Janes, J. L., Rivers, M. L., & Dunlosky, J. (2018). The influence of making judgments of learning on memory performance: Positive, negative, or both? *Psychonomic Bulletin & Review*, *25*, 2356–2364. https://doi.org/10.3758/s13423-018-1463-4.

Kirk-Johnson, A., Galla, B. M., & Fraundorf, S. H. (2019). Perceiving effort as poor learning: The misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice. *Cognitive Psychology*, *115*, 101237. https://doi.org/10.1016/j.cogpsych.2019.101237

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception 36 ECVP Abstract Supplement*.

Kliegl, O., & Bäuml, K.-H. T. (2021). When retrieval practice promotes new learning – The critical role of study material. *Journal of Memory and Language*, *120*, 104253. https://doi.org/10.1016/j.jml.2021.104253

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.

Li, B., Zhao, W., Zheng, J., Hu, X., Su, N., Fan, T., Yin, Y., Liu, M., Yang, C., & Luo, L. (2021). Soliciting judgments of forgetting reactively enhances memory as well as making judgments of learning: Empirical and meta-analytic tests. *Memory & Cognition*, *50*, 1061–1077. https://doi.org/10.3758/s13421-021-01258-y.

Macaluso, J. A., Beuford, R. R., & Fraundorf, S. H. (2022). Familiar strategies feel fluent: The role of study strategy familiarity in the misinterpreted-effort model of self-regulated learning. *Journal of Intelligence*, *10*, 83. https://doi.org/10.3390/jintelligence10040083.

McDaniel, M. A., & Einstein, G. O. (2020). Training learning strategies to promote self-regulation and transfer: The knowledge, belief, commitment, and planning framework. *Perspectives on Psychological Science*, *15*, 1363–1381. https://doi.org/10.1177/1745691620920723

Metcalfe, J., & Finn, B. (2013). Metacognition and control of study choice in children. *Metacognition and Learning*, *8*, 19–46. https://doi.org/10.1007/s11409-013-9094-7.

Mitchum, A. L., Kelley, C. M., & Fox, M. C. (2016). When asking the question changes the ultimate answer: Metamemory judgments change memory. *Journal of Experimental Psychology: General*, *145*, 200–219. https://doi.org/10.1037/a0039923.

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.

Myers, S. J., Rhodes, M. G., & Hausman, H. E. (2020). Judgments of learning (JOLs) selectively improve memory depending on the type of test. *Memory & Cognition*, *48*, 745–758. https://doi.org/10.3758/s13421-020-01025-5.

Nelson, T., & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe, & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). The MIT Press.

Putnam, A. L., & Roediger, H. L. (2013). Does response mode affect amount recalled or the magnitude of the testing effect? *Memory & Cognition*, *41*, 36–48. https://doi.org/10.3758/s13421-012-0245-x.

Rhodes, M. G. (2016). Judgments of learning: Methods, data, and theory. In John Dunlosky & S. K. Tauber (Eds.), *The Oxford handbook of metamemory* (pp. 65–80). New York, NY: Oxford University Press.

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.

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin*, *137*, 131–148. https://doi.org/10.1037/a0021705.

Rivers, M. L., Janes, J. L., & Dunlosky, J. (2021). Investigating memory reactivity with a within-participant manipulation of judgments of learning: Support for the cue-strengthening hypothesis. *Memory*, *29*, 1342–1353. https://doi.org/10.1080/09658211.2021.1985143.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning:Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x.

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-wandering with and without intention. *Trends in Cognitive Sciences*, *20*, 605–617. https://doi.org/10.1016/j.tics.2016.05.010.

Shi, A., Xu, C., Zhao, W., Shanks, D. R., Hu, X., Luo, L., & Yang, C. (2022). Judgments of learning reactively facilitate visual memory by enhancing learning engagement. *Psychonomic Bulletin & Review*, https://doi.org/10.3758/s13423-022-02174-1

Smith, M. A., Roediger, H. L., & Karpicke, J. D. (2013). Covert retrieval practice benefits retention as much as overt retrieval practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1712–1725. https://doi.org/10.1037/a0033569.

Soderstrom, N. C., Clark, C. T., Halamish, V., & Bjork, E. L. (2015). Judgments of learning as memory modifiers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 553–558. https://doi.org/10.1037/a0038388.

Sun, Y., Shi, A., Zhao, W., Yang, Y., Li, B., Hu, X., Shanks, D. R., Yang, C., & Luo, L. (2022). Long-lasting effects of an instructional intervention on interleaving preference in inductive learning and transfer. *Educational Psychology Review*, *34*, 1679–1707. https://doi.org/10.1007/s10648-022-09666-5.

Tauber, S. K., & Witherby, A. E. (2019). Do judgments of learning modify older adults' actual learning? *Psychology and Aging*, *34*, 836–847. https://doi.org/10.1037/pag0000376.

Tekin, E., & Roediger, H. L. (2020). Reactivity of judgments of learning in a levels-of-processing paradigm. *Zeitschrift für Psychologie*, *228*, 278–290. https://doi.org/10.1027/2151-2604/a000425.

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., & Hinne, M. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*, 813–826. https://doi.org/10.3758/s13423-020-01798-5.

Verkoeijen, P. P. J. L., Rikers, R. M. J. P., & Schmidt, H. G. (2005). The effects of prior knowledge on study-time allocation and free recall: Investigating the discrepancy reduction model. *The Journal of Psychology*, *139*, 67–79. https://doi.org/10.3200/JRLP.139.1.67-79.

Witherby, A. E., & Tauber, S. K. (2017). The influence of judgments of learning on long-term learning and short-term performance. *Journal of Applied Research in Memory and Cognition*, *6*, 496–503. https://doi.org/10.1016/j.jarmac.2017.08.004.

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. https://doi.org/10.1037/xge0000177.

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*, 399–435. https://doi.org/10.1037/bul0000309.

Yang, C., Potts, R., & Shanks, D. R. (2017). 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., Potts, R., & Shanks, D. R. (2018). Enhancing learning and retrieval of new information: A review of the forward testing effect. *npj Science of Learning*, *3*, 8. https://doi.org/10.1038/s41539-018-0024-y.

Yang, C., Yu, R., Hu, X., Luo, L., Huang, T. S. T., & Shanks, D. R. (2021). How to assess the contributions of processing fluency and beliefs to the formation of judgments of learning: Methods and pitfalls. *Metacognition and Learning*, *16*, 319–343. https://doi.org/10.1007/s11409-020-09254-4.

Yang, H., Cai, Y., Liu, Q., Zhao, X., Wang, Q., Chen, C., & Xue, G. (2015). Differential neural correlates underlie judgment of learning and subsequent memory performance. *Frontiers in Psychology*, *6*, 1699. https://doi.org/10.3389/fpsyg.2015.01699

Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society*, *15*, 41–44. https://doi.org/10.3758/BF03329756.

Zhao, W., Li, B., Shanks, D. R., Zhao, W., Zheng, J., Hu, X., Su, N., Fan, T., Yin, Y., Luo, L., & Yang, C. (2022). When judging what you know changes what you really know: Soliciting metamemory judgments reactively enhances children's learning. *Child Development*, *0*, 1–13. https://doi.org/10.1111/cdev.13689

Zhao, X., Wang, C., Liu, Q., Xiao, X., Jiang, T., Chen, C., & Xue, G. (2015). Neural mechanisms of the spacing effect in episodic memory: A parallel EEG and fMRI study. *Cortex*, *69*, 76–92. https://doi.org/10.1016/j.cortex.2015.04.002

Appendices

Appendix A: Experimental instructions for the JOL group

In the formal experiment, you will study 2 blocks of Chinese words. The study time for each pair is 6 s. In the first block, after each word appears for 3 s, a scale slider, ranging from 0 to 100, will appear below the word. Please predict the likelihood that you would remember the word in a later memory test through dragging and clicking the slider (0 = *Sure I will not remember it*; 100 = *Sure I will remember it*). The scale will be presented for 3 s, and please make your prediction during 3 s. After studying all words in the first block, you will start to study a new block of words, and the study procedure is identical to that in the first block, except that you will not need to make memory predictions when studying each word in the second block. Please try to remember as many words as you could regardless of whether you need to make memory predictions or not, because all words will be finally tested.

If you fully understand the instructions, please click the mouse to start the practice task. If not, please consult our experimenter.

Appendix B: JOL results

Experiment 1

For the JOL group, participants successfully provided item-by-item JOLs to 98.4% (SD = 1.5%) of words in Block 1. The average of JOLs was 63.860 (*SD* = 12.401). The averaged *G* across participants was 0.106 (*SD* = 0.464), 95% CI [−0.056, 0.268], which is not significantly different from 0, $t(33) = 1.329$, $p = .193$, Cohen's $d = 0.228$, $BF_{10} = 0.411$.

Experiment 2

For the JOL group, participants successfully provided item-by-item JOLs to 98.7% (*SD* = 1.5%) of word pairs in Block 1. The average of JOLs was 71.405 (*SD* = 10.218). For each participant, a Gamma (*G*) correlation was calculated to measure relative accuracy of JOLs. The averaged *G* across participants was 0.206 (*SD* = 0.370), 95% CI [0.085, 0.328], which is significantly greater than 0, $t(37) = 3.439$, $p = .001$, Cohen's $d = 0.558$, $BF_{10} = 22.168$.

Experiment 3

For the JOL group, participants successfully provided item-by-item JOLs to 98.7% (*SD* = 2.4%) of word pairs in Block 1. The average of JOLs was 70.038 (*SD* = 11.187). For each participant, a Gamma (*G*) correlation was calculated to measure relative accuracy of JOLs. Data from three participants were removed from Gamma calculation because their data contain too many identical values in JOL. The averaged *G* across participants was 0.219 (*SD* = 0.337), 95% CI [0.100, 0.339], which is significantly greater than 0, $t(32) = 3.742$, $p < .001$, Cohen's $d = 0.651$, $BF_{10} = 43.269$.