Metacognitive illusions can be reduced by monitoring recollection during study

Nicholas C. Soderstrom\textsuperscript{a} & Matthew G. Rhodes\textsuperscript{b}
\textsuperscript{a} Department of Psychology, University of California, Los Angeles, CA, USA
\textsuperscript{b} Department of Psychology, Colorado State University, Fort Collins, CO, USA
Published online: 17 Sep 2013.
Metacognitive illusions can be reduced by monitoring recollection during study

Nicholas C. Soderstrom\(^1\) and Matthew G. Rhodes\(^2\)

\(^1\)Department of Psychology, University of California, Los Angeles, CA, USA
\(^2\)Department of Psychology, Colorado State University, Fort Collins, CO, USA

Prior work has shown that judgments of learning (JOLs) are prone to an auditory metacognitive illusion such that loud words are given higher predictions than quiet words despite no differences in recall as a function of auditory intensity. The current study investigated whether judgments of remembering and knowing (JORks)—judgments that focus participants on whether or not recollective details will be remembered—are less susceptible to such an illusion. In Experiment 1, participants studied single words, making item-by-item JOLs or JORks immediately after study. Indeed, although increased volume elevated judgement magnitude for both JOLs and JORks, the effect was significantly attenuated when JORks were elicited. Experiment 2 replicated this finding and additionally demonstrated that participants making JORks were less likely than participants making JOLs to choose to restudy quiet words relative to loud words. Taken together, these results suggest that JORks are impacted less—in terms of both metacognitive monitoring and control—by irrelevant perceptual information than JOLs. More generally, these data support the contention that metacognitive illusions can be attenuated by simply changing the way metacognitive judgments are solicited, an important finding given that subjective experiences guide self-regulated learning.

**Keywords:** Metamemory; Judgments of learning; Recollection; Metacognitive illusions.

The importance of making valid subjective judgments is underscored by the fact that people act on their subjective experiences. A student, for example, realising that material for an upcoming exam has not been mastered, might reread the material. The degree to which such self-regulated learning is beneficial, however, depends on the accuracy of the preceding subjective assessments. This relationship between *monitoring* one’s cognitive processes, and cognitive and behavioural *control*, has been of focal interest in metacognitive research (see Bjork, Dunlosky, & Kornell, 2013).

One commonly elicited prospective metacognitive judgement is the judgement of learning (JOL), in which participants are asked to assess the likelihood of remembering an item on a later test. JOLs likely reflect inferences based on the cues available to the rememberer and thus their accuracy depends on the degree to which these cues are diagnostic of future performance (Koriat, 1997). Indeed, JOLs made immediately after studying an item are often weakly-to-moderately accurate, suggesting that diagnostic cues are, to some extent, taken into account when making JOLs. However, there are many cases in which JOLs are unrelated—or even negatively related—to actual memory performance (e.g., Benjamin, Bjork, & Schwartz, 1998; Rhodes & Castel, 2008, 2009; Soderstrom & McCabe, 2011). Of greatest relevance to the current study, Rhodes and Castel...
(2009) reported a metacognitive illusion such that participants gave higher JOLs for words presented in a loud volume relative to those presented in quiet volume, whereas actual recall performance did not differ based on auditory intensity (see also Foster & Sahakyan, 2012). Furthermore, participants chose to restudy quiet items more than loud items, despite equivalent memory performance. Thus, Rhodes and Castel (2009) demonstrated that people mistakenly believed that volume would impact future memory performance and acted on this false belief. Finding that faulty monitoring drove control judgments is of key practical and theoretical importance as it provides empirical support that monitoring has a causal influence on learning (cf. Metcalfe & Finn, 2008).

Given that JOLs usually show moderate predictive accuracy and are prone to metacognitive illusions, much research seeks to identify methods of improving the accuracy of immediate prospective metacognitive judgments.1 McCabe and Soderstrom (2011) identified one such method by asking participants to base their memory predictions on the episodic memory experiences of remembering and knowing (Tulving, 1985). Briefly, remembering is defined as recollecting details of an event (e.g. seeing someone and being able to recall the person’s name and where you know them from), whereas knowing is characterised by a sense of familiarity in the absence of recollective details (e.g. being confident that you know the person but without memory for the person’s name or any other details).2 These subjective states are typically assessed retrospectively (for a review, see Gardiner, 2002); however, McCabe and Soderstrom developed prospective remember-know judgments, called judgments of remembering and knowing (JORKs). Thus, compared to JOLs that ask people to assess the likelihood that studied material will be remembered in the future, JORKs ask people to monitor whether self-generated details associated with the studied material will be remembered in the future (see also Soderstrom, McCabe, & Rhodes, 2012).

McCabe and Soderstrom (2011) observed that JORKs generally led to higher levels of predictive accuracy than JOLs. For example, participants in their first experiment studied single words, making either item-by-item JOLs (‘WILL Remember <1—2—3> WON’T Remember’) or JORKs (‘Will you [1] Recollect, [2] Know, or [3] Forget?’) after each item.3 Comparing these metacognitive judgments with actual memory performance on a recognition test showed that JORKs predicted memory better than JOLs. Consequently, the authors concluded that ‘...JORKs can be considered judgments that focus participants’ attention on information that is more closely associated with target retrievability than are immediate JOLs’ (McCabe & Soderstrom, 2011, p. 616). That is, JORKs seem to be based more on diagnostic cues than JOLs.

If JORKs are indeed based more on diagnostic information than JOLs, then the inverse should also be true: JORKs should be relatively less susceptible to the influence of irrelevant information. We examined this possibility in Experiment 1 by having participants make either item-by-item JOLs or JORKs after hearing loud and quiet words, a manipulation itself not expected to impact actual memory performance. We predicted that JOLs would be highly prone to this metacognitive illusion such that participants would mistakenly regard loud words as more memorable than quiet words (replicating Rhodes & Castel, 2009). In contrast, we predicted that JORKs would be less susceptible, or even immune, to this auditory manipulation. Experiment 2 added an element of metacognitive control by soliciting restudy choices after making judgments. If participants making JOLs are highly susceptible to the auditory illusion then they should choose to more frequently restudy quiet items that are mistakenly regarded as less memorable than loud items. If JORKs are less susceptible to the auditory illusion, then restudy choices should differ less (if at all) based on volume.

---

1 Imposing a delay between studying and making JOLs for items has been shown to dramatically increase the relative accuracy of JOLs, a finding termed the delayed-JOL effect (Nelson & Dunlosky, 1991; for a review, see Rhodes & Tauber, 2011). However, of current interest are immediate judgments.

2 A long-standing debate concerns whether episodic memory is based on a single process of memory strength or on dual processes composed of recollection (often equated with remembering) and familiarity (often equated with knowing; for a review, see Yonelinas, 2002). We are simply interested in the subjective experiences of remembering and knowing, and make no contention whether these states accurately map onto the processes of recollection and familiarity, respectively.

3 JOLs are typically made on a 0%–100% scale; however, given that JORKs were made on a 3-point scale (Recollect, Know, or Forget?), JOLs also needed to be for the two judgments to be directly comparable. Also, for JORKs the term ‘recollect’ was used instead of ‘remember’ in order to avoid potential confusion among participants regarding the differences between the subjective states of remembering and knowing.
EXPERIMENT 1

Method

Participants
Forty undergraduates from Colorado State University (M age = 20.85 years) participated for course credit. There were 20 participants in each of the two between-subjects conditions (JOLs and JORKs). Participants were tested individually.

Materials
Studied items were taken from Rhodes and Castel (2009) and consisted of 42 single nouns recorded in a male voice. Items were edited such that loud items were approximately three times as loud as quiet (conversational) items (see Rhodes and Castel for more details regarding these items). Half of the items were presented in a loud volume, whereas the other half were presented in a quiet volume, and each item was presented often equally at each volume level. Of the 42 total items, 6 served as buffer items (3 at the beginning of the list and 3 at the end of the list), which were presented often equally in either a loud or quiet volume and were excluded from analyses.

Procedure
Before the study phase began, participants listened to four words presented one-at-a-time through a set of speakers located next to the computer monitor. This phase ensured that participants could clearly identify the presented items; adjustments in volume level were made by the experimenter when appropriate. Participants were then informed that words would be heard one-at-a-time at different volumes and that they would make memory predictions after each item in anticipation of a later test. For JOLs, participants assessed the likelihood, on a scale from 1 (very likely) to 3 (not likely at all), that the item would be recalled on a later test. For JORKs, participants were asked to assess whether recollective details associated with the item would be remembered. Like JOLs, JORKs were made on a 3-point scale; however, these judgments were made based on the following categories: 1 = Recollect, 2 = Know and 3 = Forget. Participants in the JORK condition were read instructions explaining that their predictions would be based on the experiences of remembering and knowing, with these judgments explained in detail (based on McCabe & Geraci, 2009; see also McCabe & Soderstrom, 2011). JOLs and JORKs were fixed-paced at 4 sec and were made by pressing the 1–3 keys on the keyboard. Participants were encouraged to use the entire scale when making their judgments. Studied items were presented in a fixed–random order at a 4-sec rate, with the condition that no more than three items of the same volume were presented consecutively. After all words had been given a judgement, participants engaged in a 4-min filler task that consisted of writing down as many states in the United States as possible. Participants were then given 4 min to freely recall, in any order, the words previously presented in the study phase using a blank sheet of paper provided by the experimenter.

Results and discussion
To determine whether JOLs and JORKs were differentially influenced by volume, we conducted a 2 (Judgement Type: JOL vs. JORK) × 2 (Volume: loud vs. quiet) mixed-model analysis of variances (ANOVA) on mean judgement magnitude (see Figure 1a: Note that for purposes of the Figure, JOLs and JORKs were reversed scaled such that a higher value reflects a higher judgement, e.g. a ‘3’ reflects a high-confidence JOL and
a ‘recollect’ JORK. The opposite was true when participants made their judgments during the experiment.) Regardless of Judgement Type (i.e. JOLs or JORKs), loud words were given higher predictions than quiet words, $F(1,38) = 72.52$, $MSE = 5.15, \eta^2 = .66$. However, the interaction between Judgement Type and Volume was reliable, $F(1,38) = 5.62, MSE = .40, \eta^2 = .13$. For loud words, judgement magnitude was lower for JORKs than JOLs, $t(38) = 2.13, p < .05, d = .67$, whereas no difference was found for quiet words ($p > .05$). Consequently, the difference in the magnitude of mean predictions for loud and quiet words was smaller for JORKs compared to JOLs ($d = .65$ vs $d = .66$, respectively), indicating that JORKs were influenced less by this perceptual cue. Consistent with previous research (Foster and Sahakyan, 2012; Rhodes and Castel, 2009), and shown in Figure 1b, volume had no effect on final recall for either JOLs or JORKs ($F's < 1$).

Goodman–Kruskal gamma correlations (see Nelson, 1984) were also calculated to examine the extent to which JOLs and JORKs were informed by volume. Gamma correlations will be positive to the extent that higher judgments were given for loud words relative to quiet words. Indeed, the mean gamma correlations between JOLs and volume ($G = .68, SE = .04$) and JORKs and volume ($G = .31, SE = .10$) were significantly greater than zero, $t(19) = 16.49, p < .001$ and $t(19) = 3.06, p < .01$, respectively. Critically, however, the mean JORK-volume gamma was significantly weaker than the mean JOL-volume gamma, $t(38) = 3.45, p = .001$ and $d = 1.08$, providing convergent evidence that JORKs were less influenced by this irrelevant cue than JOLs.$^4$

One potential explanation for the reduction in the magnitude of JORKs compared to JOLs as a function of volume is that participants making JORKs may have simply been reluctant to use the ‘remember’ prediction, whereas participants making JOLs may have been less reluctant to use the highest confidence (very likely) prediction. By this account, differences between JOLs and JORKs emerged because of the frequency with which participants were willing to give a particular response, regardless of whether the item was loud or quiet.

To address this issue, we analysed response distributions for JOLs and JORKs as a function of volume (see Table 1). We first conducted a 2 (Judgement Type: JOL vs. JORK) × 2 (Volume: loud vs. quiet) × 3 (Judgement Frequency: 1/Remember, 2/Know, 3/Forget) mixed-model ANOVA. (Note that because the frequency of responses sums to 1 for both JOLs and JORKs, main effects of Judgement Type could not be calculated; therefore, our focus was on the main effects of Volume and Judgement Frequency, and the various interactions.) The only significant effects included an interaction between Volume and Judgement Frequency, $F(2,76) = 27.27, MSE = 1.32, \eta^2 = .42$, and the three-way interaction between Judgement Type, Volume and Judgement Frequency, $F(2,76) = 4.81, MSE = .23, \eta^2 = .11$.

We unpacked this three-way interaction by conducting separate 2 (Judgement Type: JOL vs. JORK) × 3 (Judgement Frequency: 1/Remember, 2/Know, 3/Forget) mixed-model ANOVAs for each volume level. For loud items, a main effect of Judgement Frequency was found, $F(2,76) = 26.15, MSE = 1.09$ and $\eta^2 = .41$, indicating that the frequency of judgments was not equally distributed across the three response categories. Similar to loud items, for quiet items there was a main effect of Judgement Frequency, $F(2,76) = 7.67, MSE = .39, \eta^2 = .17$; however, this was qualified by an interaction between Judgement Type and Judgement Frequency, $F(2,76) = 4.26, MSE = .22, \eta^2 = .10$. For the proportion of quiet items given the highest response of 1/Remember, JORKs showed an advantage (JOLs = .16; JORKs = .30), $t(38) = 2.54, p < .05, d = .81$, whereas for the proportion of loud items given a medium response of 2/Know, JOLs showed an advantage (JOLs = .43; JORKs = .27), $t(38) = 2.94, p < .05, d = .93$. No difference was found for the lowest response of 3/Forget ($p > .05$). Thus, participants making JOLs and JORKs differed in terms of how their judgments—particularly, judgments of 1/Remember and 2/Know for quiet items—were distributed across items rather than the overall frequency in which these judgments were provided. For completeness, Table 1 also reports the probability of correct recall given each level of prediction for both JOLs and JORKs. A 2 (Judgement Type: JOL vs. JORK) × 2 (Volume: loud vs. quiet) × 3 (Recall Probability: 1/Remember, 2/Know and 3/Forget) mixed-model ANOVA yielded no interactions with Judgement Type and thus no follow-up analyses were performed.
EXPERIMENT 2

Experiment 1 showed that, although participants making JOLs or JORKs erroneously predicted loud words to be more memorable than quiet words, this metacognitive illusion was significantly attenuated for JORKs. Experiment 2 sought to replicate and extend this finding to a measure of metacognitive control. Specifically, after making each JOL or JORK, participants indicated whether they wanted the opportunity to restudy that item before the test. Following the idea that monitoring affects control (Nelson & Narens, 1990), we expected that participants would seek to restudy items regarded as least memorable. Thus, we predicted that those making JORKs, as result of being less susceptible to the auditory illusion, would be less likely than participants making JOLs to choose to restudy quiet words relative to loud words (cf. Dunlosky and Thiede, 1998). That is, to the extent that metacognitive monitoring informs subsequent control decisions, monitoring differences between JOLs and JORKs should also be reflected in participants’ restudy choices.

Method

Participants

Sixty undergraduates from Colorado State University (M_{age} = 19.53 years) participated for course credit. There were 30 participants in each of the two between-subjects conditions (JOLs and JORKs). Participants were tested individually.

Materials and procedure

The materials, design and procedure were identical to Experiment 1 with one exception: Immediately after making each JOL or JORK, participants made a restudy choice indicating whether they would like to restudy the item prior to the test. Participants were given 4 sec to press either the ‘1’ (Yes) or ‘2’ (No) key to indicate their choice. Restudy choices were not honoured at any point during the experiment.

Results and discussion

Similar to Experiment 1, we first conducted a 2 (Judgement Type; JOL vs. JORK) × 2 (Volume; loud vs. quiet) mixed-model ANOVA on mean judgement magnitude (see Figure 2a). Loud words were given higher judgments than quiet words, F(1,58) = 31.43, MSE = 3.01, η^2 = .35. A marginal main effect of Judgement Type was found, F(1,58) = 3.07, MSE = .42, η^2 = .05, p = .09. However, the interaction between Judgement Type and Volume was reliable, F(1,58) = 4.79, MSE = .46, η^2 = .08. For loud words, judgement magnitude was lower for JORKs than JOLs, t(58) = 2.93, p < .01, d = .76, whereas no difference was found for quiet words.
Thus, similar to Experiment 1, the gap between mean predictions for loud and quiet words was smaller for JORKs compared to JOLs (.20 vs. .44, respectively). This finding was also supported by judgement-volume gamma correlations. These correlations reliably differed from zero for both JOLs ($G = .48; \text{SE} = .08$) and JORKs ($G = .22; \text{SE} = .07$), $t(29) = 5.89$, $p < .001$ and $t(29) = 2.95$, $p < .01$, respectively; however, the JORK-volume gamma correlation was reliably weaker than the JOL-volume gamma correlation, $t(58) = 2.38$, $p < .05$, $d = .62$.

Recall performance is presented in Figure 2b. The only significant effect was that, unexpectedly, loud words were recalled more than quiet words, $F(1,58) = 5.51$, MSE = .04, $\eta^2 = .09$. This main effect of Volume is inconsistent with both the results from Experiment 1 and previous research (Foster and Sahakyan, 2012; Rhodes and Castel, 2009), and thus this finding should be interpreted cautiously. Indeed, when gamma correlations between volume and recall were calculated—which measure the extent to which volume is related to recall—neither JOLs ($G = .08; \text{SE} = .09$) nor JORKs ($G = .11; \text{SE} = .08$) yielded gammas that differed reliably from zero ($p's > .05$).

Restudy choices were analysed by conducting a 2 (Judgement Type: JOL vs. JORK) × 2 (Volume: loud vs. quiet) mixed-model ANOVA (see Figure 3). A marginal main effect of Judgement Type was found, $F(1,58) = 4.00$, MSE = .41, $\eta^2 = .09$, $p = .05$, as was a reliable main effect of Volume, $F(1,58) = 30.77$, MSE = .67, $\eta^2 = .35$. However, these effects were qualified by a reliable Judgement Type by Volume interaction, $F(1,58) = 9.04$, MSE = .20, $\eta^2 = .14$. The proportion of loud words chosen for restudy did not differ between those making JOLs and JORKs ($p > .05$), but differed substantially for quiet words. Specifically, the percentage of quiet items chosen for restudy was 20% points lower for participants making JORKs than JOLs, $t(58) = 3.02$, $p < .01$, $d = .78$. Consequently, this led to a smaller difference in volume-driven restudy choices in the JORK compared to the JOL condition (.07 vs. .23, respectively). Consistent with this finding, gamma correlations between volume and restudy decisions were weaker for JORKs ($G = -.19; \text{SE} = .09$) than JOLs ($G = -.48; \text{SE} = .07$), $t(54) = 2.60$, $p < .05$, $d = .70$. Thus, not only were JORKs less sensitive to volume, but subsequent restudy choices were also less sensitive to this irrelevant cue.5

Similar to Experiment 1, we analysed response distributions for JOLs and JORKs as a function of

---

5The observed differences in restudy choices between JOL and JORK participants provide evidence against the argument that volume was an equally powerful cue in both conditions, but that high-confidence predictions were simply distributed differently across JOLs and JORKs. If this was the case, both conditions should have chosen to restudy a similar proportion of loud vs. quiet items.
volume by conducting a 2 (Judgement Type: JOL vs. JORK) × 2 (Volume: loud vs. quiet) × 3 (Judgement Frequency: 1/Remember, 2/Know, 3/Forget) mixed-model ANOVA. These data are presented in Table 1. The significant effects included a main effect of Judgement Frequency, $F(2,116) = 15.34$, MSE = .93, $\eta^2 = .21$, an interaction between Volume and Judgement Frequency, $F(2,116) = 17.28$, MSE = .77, $\eta^2 = .23$, and the three-way interaction between Judgement Type, Volume and Judgement Frequency, $F(2,116) = 6.02$, MSE = .27, $\eta^2 = .09$. To unpack the three-way interaction, separate 2 (Judgement Type: JOL vs. JORK) × 3 (Judgement Frequency: 1/Remember, 2/Know, 3/Forget) mixed-model ANOVAs were conducted for each volume level. For loud items, a main effect of Judgement Frequency was found, $F(2,116) = 23.65$, MSE = 1.14, $\eta^2 = .29$, but this was qualified by an interaction between Judgement Type and Judgement Frequency, $F(2,116) = 4.63$, MSE = .22, $\eta^2 = .07$. For the proportion of loud items given the highest response of 3/Forget, there was a JORK advantage (JOLs = .49; JORKs = .23), $t(58) = 2.42$, $p < .05, d = 62$; however, for the proportion of loud items given the lowest response of 3/Forget, there was a JORK advantage (JOLs = .12; JORKs = .23), $t(58) = 2.90$, $p < .05, d = 76$. No difference was found for the medium response of 2/Know. Turning to the quiet items, the main effect of Judgement Frequency was reliable, $F(2,116) = 9.83$, MSE = .60, $\eta^2 = .15$; however, this was qualified by an interaction between Judgement Type and Judgement Frequency, $F(2,116) = 4.22$, MSE = .24, $\eta^2 = .07$. For the proportion of quiet items given the medium response of 2/Know, JOLs showed an advantage (JOLs = .51; JORKs = .37), $t(58) = 3.12$, $p < .05, d = .87$. No differences were found for responses of 1/Remember or 3/Forget ($p$’s > .05). Thus, similar to Experiment 1, no differences were shown in the overall frequency of judgments of 1/Remember, 2/Know or 3/Forget; rather, those making JOLs and JORKs differed in terms of how these judgments were distributed across loud and quiet items. Finally, Table 1 also reports conditional recall on each level of prediction for both JOLs and JORKs. Similar to Experiment 1, a 2 (Judgement Type: JOL vs. JORK) × 2 (Volume: loud vs. quiet) × 3 (Recall Frequency: 1/Remember, 2/Know, 3/Forget) mixed-model ANOVA yielded no interactions with Judgement Type and thus no follow-up analyses were performed.

**GENERAL DISCUSSION**

Results from the current study support the notion that metacognitive illusions can be reduced by having participants monitor recollection, rather than confidence, during study. Following Rhodes and Castel (2009), participants in Experiment 1 heard loud and quiet words, making item-by-item JOLs or JORKs. Although JOLs and JORKs were impacted by this irrelevant auditory cue, JORKs were influenced significantly less than JOLs. Experiment 2 replicated and extended this finding to a measure of metacognitive control showing that participants making JORKs were less likely than participants making JOLs to choose to restudy quiet words relative to loud words. As measured by the mean weighted effect sizes collapsed across experiments, soliciting JORKs led to medium-to-large decreases in judgement magnitude for loud items ($d = .723$) and produced weaker gamma correlations between volume and judgement magnitude ($d = .796$) relative to JOLs. Thus, the change in predictions produced by JORKs was far from trivial. Moreover, participants making JORKs were not simply reluctant to use the ‘remember’ prediction. In fact, averaging across experiments, ‘remember’ predictions were given just as frequently as the highest magnitude JOL (approximately 34% of the time). Thus, the difference was not in how often these predictions were given, but rather when they were given.

That JORKs were less susceptible to this auditory-based metacognitive illusion is consistent with the notion that JORKs are based, to a larger extent than JOLs, on diagnostic (i.e. relevant) information (McCabe and Soderstrom, 2011). We showed that metacognitive monitoring can be safeguarded (albeit not entirely so) from fleeting irrelevant information by asking participants to base their memory predictions on whether self-generated details associated with the studied material would be remembered (JORKs), as opposed to simply assessing confidence in the future memorability of an item (JOLs). Such a finding is important given that people act on their subjective experiences. If one’s metacognitive monitoring is poor or faulty, so too will be that person’s subsequent study decisions. Experiment 2 clearly showed that not only were monitoring processes for JORKs less impacted by auditory intensity, but that so were restudy choices. This suggests that encouraging people to monitor
recollection during study might lead to more effective, efficient self-regulated learning.

Admittedly, arguing that JORKs reflect more diagnostic information than JOLs does not indicate the specific bases for these judgments. Indeed, actual remembering and knowing were not measured in the current study, and thus, one can only infer such information from the emergent data and how participants were instructed to make these judgments. Future research might delve into this issue by employing think-aloud protocols in which participants are asked to continuously verbalise their thoughts during certain phases of the experiment (McCabe, Geraci, Bowman, Sensenig, and Rhodes, 2011). Such a methodology might reveal important differences (and similarities) regarding the information that is used by participants when making JOLs and JORKs. Relatedly, future research might also determine why auditory intensity is used to inform memory predictions. One possibility put forth by Rhodes and Castel (2008) is that volume-driven JOLs might be the result of loud items being perceived as more fluent, or easier to process, than quiet items. Alternatively, the auditory illusion might reflect a general theory that loud words are more memorable than quiet words. Regardless of its origin, our data indicate that such irrelevant information exerts less influence when participants predict whether self-generated details associated with items will be remembered later (JORKs) as compared to simply monitoring confidence (JOLs).

In summary, the current study demonstrated that metacognitive illusions can be attenuated by simply changing the way metacognitive judgments are solicited. Specifically, asking participants to monitor recollection during study (JORKs; McCabe and Soderstrom, 2011) reduced an auditory metacognitive illusion previously shown to impact JOLs (Rhodes and Castel, 2009). Furthermore, the relative immunity of JORKs to this illusion aligned with subsequent restudy choices, which highlights the dependent relationship between metacognitive monitoring and control. These data have substantial implications in various domains—particularly educational settings—as minimising metacognitive errors should optimise self-regulated learning.

REFERENCES


Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information:


