

Distributed Metacognition: Increased Bias and Deficits in Metacognitive Sensitivity When Retrieving Information From the Internet

Timothy L. Dunn^{1, 2}, Connor Gaspar³, Daev McLean³, Derek J. Koehler³, and Evan F. Risko³

¹ Warfighter Performance Department, Naval Health Research Center, San Diego, California, United States

² Leidos Inc., San Diego, California, United States

³ Department of Psychology, University of Waterloo



Our metacognitive ability to monitor and evaluate our cognitive performance is central to efficient and adaptive behaviors. Research investigating this ability has focused largely on tasks that rely exclusively on internal processes (e.g., memory). However, our day-to-day cognitive activities often consist of the mixes of internal and external processes. In the present investigation, we expand research on metacognition to this distributed domain. We examined participants' ability to accurately monitor their performance in a knowledge retrieval task when they were required to rely on only their internal knowledge and when required to rely on both internal knowledge and utilizing the internet. One hundred and ninety-four participants completed an online study consisting of answering general knowledge questions. Individuals were also randomly assigned to provide accuracy judgments either prospectively or retrospectively. Results revealed metacognitive bias (i.e., overconfidence) increased when using the internet and when making retrospective judgments. Metacognitive sensitivity was also worse when using the internet, especially when individuals made prospective judgments about what their performance would be. Furthermore, metacognitive bias was positively related across the internal knowledge and internet conditions. These results provide the beginnings of an understanding of metacognition and behavior in distributed cognitive contexts involving the internet.

Keywords: metacognition, distributed cognition, metacognitive accuracy, transactive systems, internet search

Supplemental materials: <https://doi.org/10.1037/tmb0000039.supp>

Metacognitive processes provide an important tool in predicting our cognitive performance either when situated in a task or after-the-fact, and allow individuals to adapt behaviors accordingly (Ackerman & Thompson, 2017; Fleming et al., 2012; Metcalfe & Finn, 2008; Thiede et al., 2003). Although there is a long and rich history of research on metacognition, much of it has focused on cognitive acts that rely largely on internal processes (e.g., recalling items studied from a list then estimating the number correct). However, individuals frequently integrate “external” aids into

cognitive acts (Clark, 1997, 2008; Gray et al., 2006; Heersmink, 2013; Hutchins, 1995, 2010; Michaelian & Sutton, 2013; Risko & Gilbert, 2016; Sterelny, 2010; Sutton, 2010; Wilson, 2002). For instance, individuals off-load a tremendous amount of information retrieval onto the internet with various consequences (Ferguson et al., 2015; Fisher et al., 2015; Risko & Gilbert, 2016; Sparrow et al., 2011; Ward, 2013). How then does thinking in this more distributed manner influence our metacognitive abilities to predict and monitor our cognitive performance? In the present investigation, we

Action Editor: Danielle S. McNamara was the action editor for this article.

ORCID iDs: Timothy L. Dunn  <https://orcid.org/0000-0002-6085-7677>; Connor Gaspar  <https://orcid.org/0000-0002-6648-2925>; Daev McLean  <https://orcid.org/0000-0003-4345-5083>.

A preprint version of this article is available at <https://timothydunn.co>

The data analyzed and reported in the submitted article have not been used in any prior instances.

Neither the Department of the Navy nor any other component of the Department of Defense has approved, endorsed, or authorized this product.

This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) and funding from the Canada Research Chairs program to Evan F. Risko.

Conflicts of Interest: The authors report no real or potential conflicts of interest related to this article.

Open Science Disclosures:

 The data are available at <https://doi.org/10.17605/OSF.IO/K6JPW>

Open Access License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC-BY-NC-ND). This license permits copying and redistributing the work in any medium or format for noncommercial use provided the original authors and source are credited and a link to the license is included in attribution. No derivative works are permitted under this license.

Disclaimer: Interactive content is included in the online version of this article.

Contact Information: Correspondence concerning this article should be addressed to Timothy L. Dunn, Warfighter Performance Department, Naval Health Research Center, 140 Sylvester Rd., San Diego, CA 92152, United States. Email: timothy.dunn12.ctr@mail.mil

examine this question by having individuals predict and assess their cognitive performance in a general knowledge fact retrieval task when relying on internal processes or relying on both internal processes and the internet.

Metacognition and Cognitive Offloading

Metacognition has come to play a central role in research on cognitive offloading—the trading off of internal processing through external means (e.g., Boldt & Gilbert, 2019; Dunn & Risko, 2016; Gilbert, 2015; Hu et al., 2019; Risko & Gilbert, 2016; Weis & Wiese, 2019). The fusing of these lines of research specifically looks to examine how individuals think about their thinking when including the use of the body, environment, and/or technology, and how metacognitive feelings and judgments influence the decision to off-load cognitive processing (Dunn & Risko, 2016; Gilbert, 2015; Risko & Dunn, 2015). For example, Gilbert (2015) demonstrated that participants' confidence in their memory predicted their likelihood of offloading, where individuals with low confidence were more likely to off-load memory demands when given the opportunity. This persisted even after controlling for actual memory performance. Furthermore, engaging in cognitive offloading can influence our metacognitions. A contemporary example is that searching for information on the internet influences individuals' subsequent judgments about their own knowledge (Fisher et al., 2015). In a similar vein, Ferguson et al. (2015) demonstrated that individuals were more likely to report not knowing the answer to a question when they had the internet available. This suggests that metacognitive errors in evaluating our performance may arise from thinking in more distributed contexts. Such results point to the theoretical and practical importance of better understanding our metacognitions in these more distributed contexts.

Metacognitive Judgments in More- or Less-Distributed Cognitive Tasks

Metacognitive judgments are commonly viewed as a kind of cue-based inference (Dunlosky et al., 2014; Koriat, 1997; Mueller et al., 2016). From a cue-utilization perspective, a metacognitive judgment consists of the application, implicitly or explicitly, of some rule/heuristic over available cues. According to the analytic processing theory of judgments of learning (JOL; Dunlosky et al., 2014; Mueller et al., 2016) individuals search for variability across studied items that can be plausibly related to memory (i.e., cues). As an example, studied items may consist of either words or nonwords, where the cue is lexicality which varies across the items. Individuals may then apply the heuristic that words are easier to remember than nonwords over the lexicality cue to infer what their memory performance will be. Metacognitive judgments are often further classified into theory- or mnemonic-based judgments. The former relies on an explicit inference based on preexisting beliefs. The latter relies on the online experience of internal cues and is used more automatically (e.g., Koriat, 1997; Koriat & Bjork, 2006).

Situating cognitive offloading within a cue-utilization framework, the choice between performing a cognitive task internally or integrating an external aid can be seen as differing along two dimensions: (1) the types of available cues and (2) the nature of individuals' beliefs applied to those cues. With respect to (1),

integrating an external aid could be viewed as reducing the contribution of mnemonic (internal) cues to a metacognitive judgment. For example, when retrieving information from an external store such as the internet, the feeling of fluency which is typically part-and-parcel of internal retrieval (Kelley & Lindsay, 1993) could be less salient, ignored, or noninformative (e.g., there is a lack of directed internal search). Given such internal cues are often reliable predictors of performance (Koriat & Adiv, 2012; Michaelian, 2012; Proust, 2008; Reber & Unkelbach, 2010), this predicts a potential reduction in metacognitive accuracy when retrieving from an external source. That is, distributed metacognitive judgments may lose a beneficial source of information when accurate judgments are critical.

Judgment error, due to the lack of mnemonic cues, may be compounded by potential beliefs about the reliability of external aids. It is likely that individuals' preexisting beliefs about external aids, such as the internet, hold them as being highly reliable and associated with more reliable outcomes relative to if that functions were performed internally. This likely has metacognitive consequences, such as a potential insensitivity to errors and reduced metacognitive accuracy, and/or ignoring internal cues. Distributed prospective judgments (i.e., those made to predict upcoming performance) may be especially affected given their heavy reliance on prior rather than direct experience (Siedlecka et al., 2016). Nonetheless, it is important to also note that the use of external aids could be associated with cues not available internally (e.g., speed of search result return, cues available in a search results page; Risko et al., 2016; Stone & Storm, 2021). If these cues are reliable (or more reliable than internal ones), then they could compensate for a lack of internal cues or may even improve metacognitive performance relative to relying only on internal processes. The idea that internal cues can be misleading in forming our metacognitive judgments is well known (e.g., Benjamin et al., 1998).

Metacognitive Ability in More- or Less-Distributed Cognitive Tasks

If a given individual tends to be more metacognitively accurate in less-distributed contexts, then do they also tend to be more metacognitively accurate in more-distributed contexts? This question bears directly on ongoing debates about the extent to which a domain general metacognitive skill exists (Kelemen et al., 2000; Maki et al., 2005; Mazancieux et al., 2020; Mengelkamp & Bannert, 2010). Kelemen et al. (2000) found significant correlations in bias scores (i.e., extent of over or underconfidence), but not *G* or discrimination scores (i.e., ability to discriminate correct from incorrect responses within a given individual), within tasks across time and across different tasks (for a similar result, see Mengelkamp & Bannert, 2010). Mazancieux et al. (2020) recently reported that the tendency to report high confidence in one task is correlated with the tendency to report high confidence in another task. It is worth noting that the tasks used in this work relied completely on internal processes. The present investigation looks to provide a novel contrast to this work as the general task is the same (i.e., knowledge retrieval), but differs in the resources that participants are allowed to utilize while completing them (i.e., internal only vs. a mix of internal and external).

The Present Investigation

The present investigation examined participants' actual accuracy and their metacognitive judgments of their accuracy in a knowledge retrieval task. Participants answered fact-based general knowledge questions relying solely on internal processes or using internal processes and the internet to retrieve the answer (e.g., Ferguson et al., 2015; Risko et al., 2016). Risko et al. (2016) provided evidence that participants can make accurate judgments about their own speed of retrieval of unknown answers to general knowledge questions using the internet. This suggests that individuals have some ability to predict their performance in retrieving information from the internet—what the authors called a “feeling-of-findability.” In the present study, participants either made a prospective judgment about their general knowledge performance before completing the trial (but after being given information about the upcoming trial) or a retrospective judgment of their performance following completion of the trial. Critically, the use of both types of judgments provides an opportunity to examine the extent to which experience performing the cognitive act (i.e., retrieving the answers to the questions) contributes to metacognitive performance. For example, retrospective judgments might be expected to be superior provided the direct experience of performing the task (Fleming et al., 2016).

Two lists were used across the internal (i.e., internal knowledge alone) and external (i.e., internal knowledge and the internet) conditions based on careful item selection using pilot testing and previous research (Ferguson et al., 2015; Risko et al., 2016). The items were selected to yield similar average accuracy across conditions in a pilot study online ($N = 850$). In other words, we attempted to make the internal and external conditions equivalently “easy” based on the proportion of correct answers derived from the pilot study ($M = 88%$ and $M = 89%$ for the internal list and external list, respectively). We also selected items in the external condition that would be difficult for most people to answer when relying on their own knowledge. For example, *What is the name of the rubber roller on a typewriter (platen)*, and *What is the name of the mountain range that separates Asia from Europe (Ural)*; see the Supplemental Appendix for full item lists). Moreover, we asked participants whether they could have answered the question without the use of the internet following each trial. This was done in an attempt to limit the contribution of internal knowledge. When comparing an internal condition to an external condition there is typically, and possibly always, an asymmetry in the sense that the external condition *does not preclude* an internal contribution (i.e., internal + external sources). Though the internal condition *does preclude* the contribution of the external artifact (i.e., internal alone). In the present study, this could complicate the interpretation of the results, for example, when an individual knows the answer to a question in the external condition without the use of the internet. In this case, their metacognitive judgment may not reflect a judgment about their performance together with the internet. By equating item difficulty across conditions and removing items for which individuals reported having known the answer prior to searching the internet we can reduce, though not completely eliminate, this issue.

Method

Participants

One hundred and ninety-four participants were recruited through Amazon's Mechanical Turk to take part in the study in exchange for

financial compensation. All individuals were U.S. citizens and 18 years of age or older. Individuals were randomly assigned to either a prospective or retrospective judgment group. The study was approved by the University of Waterloo's Office of Research Ethics.

Design

A 2 (Knowledge Location: Internal vs. External) \times 2 (Judgment Type: Prospective vs. Retrospective) mixed design was utilized. The internal knowledge condition consisted of individuals attempting to use their memory alone to answer the general knowledge questions. The external knowledge condition consisted of individuals using their memory and the internet to search for answers. The knowledge location factor was manipulated within-subjects and the order of condition completion was counterbalanced across participants. Judgment type was manipulated between subjects and was assigned randomly to create two balanced groups. For each cell of the 2 \times 2 design both actual (i.e., accuracy on the general knowledge questions) and judged accuracy (i.e., individuals' metacognitive estimates of their performance) were considered.

Stimuli and Apparatus

Participants were presented with two lists of 30 general knowledge questions selected from Tauber et al.'s (2013) updated list of normed general knowledge questions (see Supplemental Materials). The task was presented through Qualtrics online survey software.

Procedure

Following instructions, participants performed two blocks of trials across both knowledge location conditions. On each trial, participants were first presented with a general knowledge question. Participants in the prospective judgment condition were then asked to provide an estimate of how likely they were to provide the correct answer using a slider with end points of 0 and 100 in increments of 1. This screen did not appear for participants in the retrospective condition. In the internal location condition, a screen was then presented with an empty box where participants were to type their answer. In the external location condition, participants were presented with the general knowledge question and instructed to “Click next when you are ready to begin your search.” On the next screen the question remained displayed and participants were instructed to “Please search the Internet now. Click ‘NEXT’ when you have found the answer.” Finally, they were presented with the box to input their answer.

In the prospective judgment condition, participants in the internal location condition moved on to the next trial and participants in the external condition were asked if they knew the answer to the question before they searched the internet for it following the entry of their estimate. In the retrospective judgment condition, participants in both the internal and external conditions entered their estimate after entering their answer and before the question pertaining to prior knowledge in the external condition. At the end of the experiment participants were asked to provide an estimate as to how many of the questions that they believed that they had answered correctly in each block out of 30. Because we only analyze questions in the external condition that participants did not know before looking up the answer, we do not report an analysis of these questions as their estimates would refer to all questions. The entire

experiment took approximately 30 min to complete. All data and code for the experiment are openly available via the Open Science Framework (Dunn et al., 2021).

Results

Metacognitive *bias* and *sensitivity* are both critical factors to consider when assessing metacognition, though they are often conflated (Fleming & Lau, 2014). Our results first focus on metacognitive *bias* (or Type 2 bias) which highlights over/underconfidence or the calibration of judgments to performance, and is indexed by actual performance and judged performance for each knowledge condition and judgment type. The second portion of results focuses on the relations between judgments and accuracy which is referred to as metacognitive *sensitivity* (or metacognitive accuracy). We measure sensitivity here with Gamma correlations and discrimination scores.

Data analysis and visualization was completed in the R programming language (R Core Team, 2016) with the assistance of add-on packages (*boot*, Canty & Ripley, 2016; *effsize*, Torchiano, 2016; *ez*, Lawrence, 2015; *ggplot2*, Wickham, 2009; *hmisc*, Harrell, 2016; *psych*, Revelle, 2016; *reshape2*, Wickham, 2007). Five trials were removed due to individuals not following instructions. We removed 36.4% of trials because participants said they knew the answer prior to searching in the external condition. Including these trials in analyses of the external condition would, in theory, contaminate metacognitive judgments with solely internal ones. That is, individuals would likely need to *only* rely on metacognitions related to internal processing when generating confidence ratings for these questions. Moreover, an analysis including these items yielded the same overall pattern of results. For specific statistics related to this analysis please see Footnote 1.¹ Greenhouse–Geisser corrections (Greenhouse & Geisser, 1959) are used to adjust significance levels when Sphericity assumptions are violated. Welch’s *t* tests are used where applicable. Effect sizes are reported using generalized eta squared (η_G^2 ; Bakeman, 2005) and/or Cohen’s *d* (Cohen, 1988). As noted above, half of the participants started the experiment in the internal condition and the other half in the external condition. When we included “Order” as a between-subject factor in the reported analyses we found no main effect or any interactions with the variable.

Metacognitive Bias: Actual and Judged Accuracy

A 2 (Measure: Absolute Actual vs. Absolute Judged) \times 2 (Memory Location: Internal vs. External) \times 2 (Judgment Type: Prospective vs. Retrospective) mixed-design analysis of variance (ANOVA) was performed assessing the complete crossing of all factors. Results demonstrated a significant main effect of memory location, $F(1, 192) = 33.47, p < .001, \eta_G^2 = .04$, where accuracy (both actual and judged) was higher in the internal condition where individuals used their internal knowledge relative to when they also used the internet. There was also a significant three-way interaction between measure, location, and judgment type, $F(1, 192) = 7.2, p = .008, \eta_G^2 = .005$. Comparison of the left and right panels in Figure 1 demonstrates the three-way interaction where actual and judged accuracy for each judgment type noticeably differed across the location types. We break down these patterns as a function of the

internal and external conditions separately in the subsequent section. No other model terms were significant, F ’s $< 3.11, p$ ’s $> .07$.

To probe the three-way interaction further we conducted two 2 (Measure: Absolute Actual vs. Absolute Judged) \times 2 (Judgment Type: Prospective vs. Retrospective) ANOVAs, one for each location. In the internal condition, there was a main effect of measure, $F(1, 192) = 15.5, p < .001, \eta_G^2 = .02$, such that participants’ actual accuracy was higher than their judged accuracy across judgment types (i.e., underconfidence). There was no interaction between measure and judgment type, $F(1, 192) = 1.96, p = .16, \eta_G^2 = .002$. Thus, overall individuals were underconfident in their memory regardless of whether they provided their judgments prospectively or retrospectively (see the left panel of Figure 1). In the external condition, there was no main effect of measure, $F(1, 192) = .005, p = .82, \eta_G^2 = .0001$, but there was a significant interaction between measure and judgment type, $F(1, 192) = 4.33, p = .039, \eta_G^2 = .009$. In the prospective condition, participants were numerically underconfident but not significantly so, $F(1, 96) = 1.66, p = .20, \eta_G^2 = .008$. They were significantly overconfident in the retrospective condition, $F(1, 96) = 4.44, p = .038, \eta_G^2 = .013$. In contrast to the internal condition, the patterns of actual and judged accuracy varied as a function of when individuals provided their judgments in the external condition. Similar to both judgment types in the internal condition, individuals were qualitatively underconfident in their performance when incorporating the internet and providing judgments prospectively. However, individuals became overconfident when providing judgments after utilizing the internet (see the right panel of Figure 1).

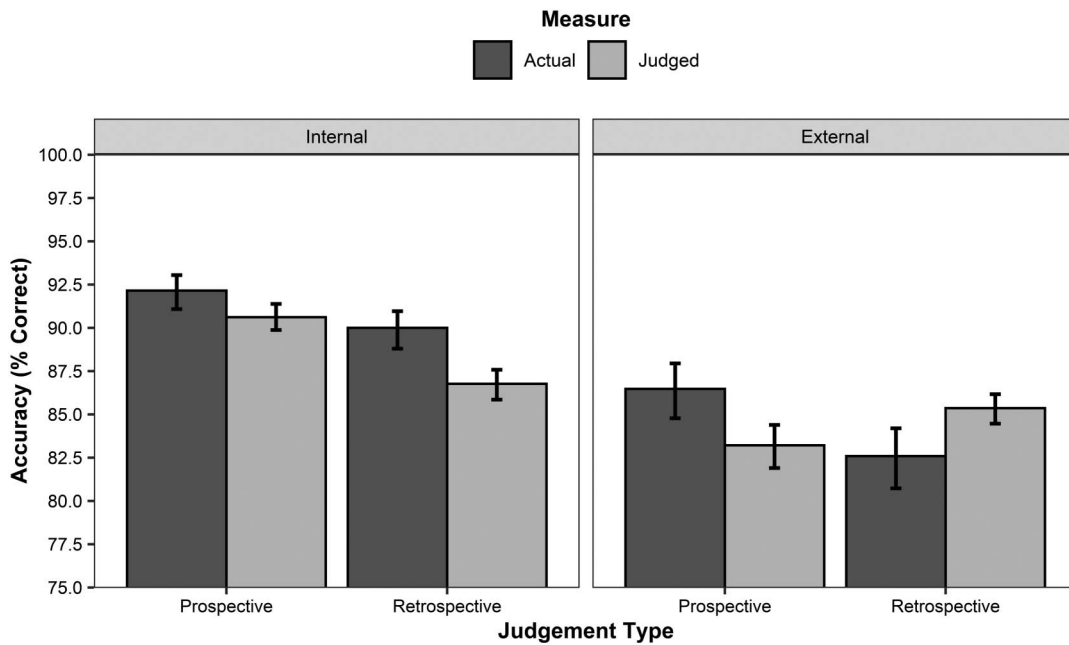
Metacognitive Sensitivity

Gamma

We first assessed metacognitive accuracy by computing Gamma coefficients (Nelson, 1984). Gamma coefficients can range from -1 (a perfect negative association) to $+1$ (a perfect positive association), with 0 representing an absence of association (i.e., ratings at random). Overall, positive Gamma’s indicate better agreement between judgments and performance (e.g., one is actually correct when confident). Here, we present average Gamma correlations as a function of conditions. Given a lack of variability in accuracy (e.g., always correct) or confidence ratings (e.g., an individual never changed their rating over trials), 27.3% ($n = 53$) of participants had no Gamma coefficient and were removed from the following analyses ($n = 35$ for the prospective condition, $n = 18$ for the retrospective condition). As a final note, we additionally observed 24 occurrences where participants achieved a perfectly inverse Gamma of -1 , and 45 occurrences where there was perfectly

¹ In the main analysis of the external condition trials on which individuals reported that they would have known the answer to the question without searching the internet were removed. We conducted a reanalysis of the external condition with those trials included. The results were qualitatively the same with the following exception. With respect to judged accuracy unlike the main analysis, there was no main effect of location, $F(1, 192) = .08, p = .78, \eta_G^2 < .001$. There was an interaction between location and judgment type, $F(1, 192) = 6.69, p = .01, \eta_G^2 = .01$, such that confidence was not significantly different across locations when given prospectively, $F(1, 96) = 2.83, p = .095, \eta_G^2 = .01$, but confidence was higher in the external condition when given retrospectively, $F(1, 96) = 4.8, p = .03, \eta_G^2 = .01$.

Figure 1
Actual and Judged Accuracy as a Function of Location and Judgment Type



Note. Error bars are 95% bias corrected and accelerated (BCa) confidence intervals.

positive Gamma of +1. For the purpose of main analysis below we included these participants. When these “perfect” correlations were excluded in a secondary analyses the results were not qualitatively different.²

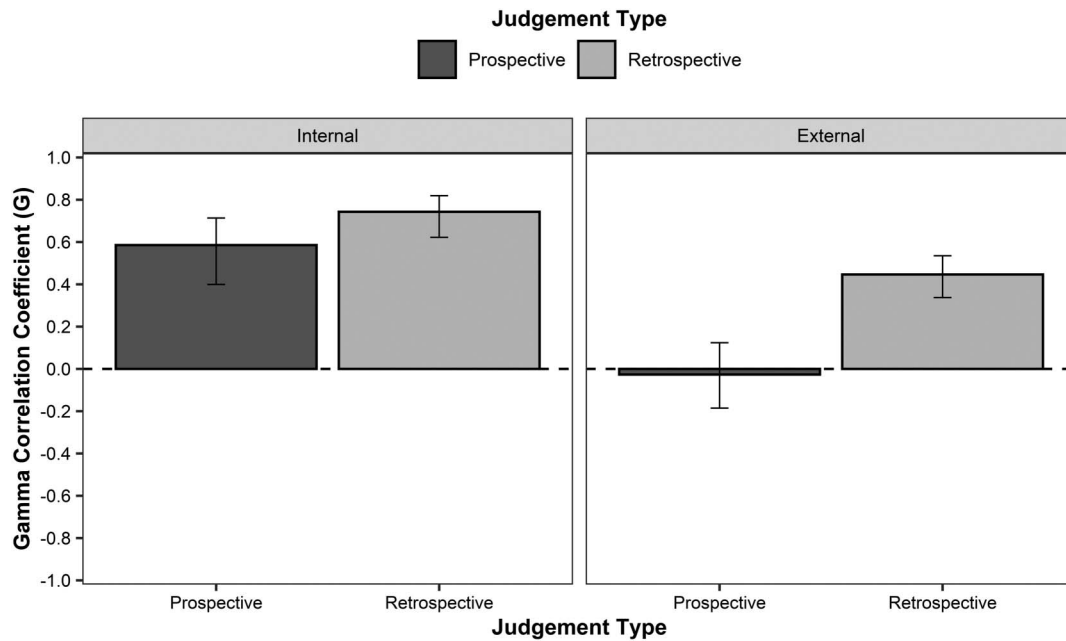
A 2 (Knowledge Location: Internal vs. External) \times 2 (Judgment Type: Prospective vs. Retrospective) mixed ANOVA demonstrated significant main effects of knowledge location, $F(1, 139) = 49.84$, $p < .001$, $\eta_G^2 = .15$, and judgment type, $F(1, 139) = 24.72$, $p < .001$, $\eta_G^2 = .08$, such that metacognitive accuracy was overall higher in the internal relative to external location and retrospective relative to prospective conditions. There was also a significant two-way interaction between location and judgment type, $F(1, 139) = 6.46$, $p = .012$, $\eta_G^2 = .02$. For the internal location, metacognitive accuracy in the prospective condition, $M = .59$, $SD = .62$, did not differ from the retrospective condition, $M = .74$, $SD = .42$, Welch’s $t(102.9) = 1.71$, $p = .09$, $d = .29$, d 95% CI [-0.05, .63]. However, for the external location, metacognitive accuracy was much lower in the prospective condition, $M = -.03$, $SD = .63$, relative to the retrospective condition, $M = .45$, $SD = .43$, Welch’s $t(103.3) = 5.06$, $p < .001$, $d = .86$, d 95% CI [.51, 1.21]. Taken together, individuals produced better agreement between their judgments and performance in the internal condition as a whole, and when judging their performance retrospectively when using the internet. Interestingly, this association was zero (i.e., random ratings) when individuals were asked to judge their performance prior to integrating the internet. Recall that in the external prospective condition individuals were somewhat underconfident, though this pattern was not significant (cf., retrospective judgments in the external condition where individuals were significantly overconfident, but also demonstrated better metacognitive accuracy; see Figure 2).

Discrimination

We next assessed metacognitive sensitivity through an individual’s ability to discriminate correct from incorrect answers (i.e., knowing when one is either right or wrong). Accuracy judgments were analyzed as a function of actual accuracy within and across

² In the metacognitive accuracy section we chose to include those participants who produced perfect Gamma correlation coefficients (either +1 or -1). Although not statistically improbable, the likelihood a priori of any individual being perfectly metacognitively accurate in a real-world setting is very unlikely. Specifically, the perfect Gamma coefficients we observed here are likely a function of accuracy overall being relatively high in the experiment. Thus, we further elected to reanalyze Gamma correlations that excluded the following: (a) those who produced gamma coefficients of -1 and (b) those who produced gamma coefficients of -1 and/or +1. In the former scenario, the sample size was reduced by 14.8% ($N = 120$) relative to the main analysis. However, the pattern of results is similar. Participants were more metacognitively accurate retrospectively than prospectively in the external condition, $t(118) = 3.49$, $p = .001$, $d = .62$, d 95% CI [.24, 1.00], but not in the internal condition, $t(118) = .88$, $p = .38$, $d = .16$, d 95% CI [-0.21, .53], $F(1, 118) = 6.37$, $p = .012$, $\eta_G^2 = .025$, judgment type by location interaction. In the latter scenario, where the sample size was reduced by 43.9% ($N = 79$), the originally observed main effects are conserved, but the interaction is not significant. Participants in the external condition were more metacognitively aware retrospectively than prospectively, $t(77) = 2.27$, $p = .025$, $d = .49$, d 95% CI [.028, .954], but not in the internal condition, $t(77) = .68$, $p = .496$, $d = .15$, d 95% CI [-0.30, .61], $F(1, 77) = 2.08$, $p = .154$, $\eta_G^2 = .013$, judgment type by location interaction. Critically, in each case the overall pattern of results stays the same: Individuals’ metacognitive accuracy was worse prospectively relative to retrospectively in the external condition, with no difference across judgment type in the internal condition. Furthermore, it is likely that the loss of the significant between-subjects interaction effect in the second scenario is a result of a loss of power, not from a change in the pattern of results due to exclusions.

Figure 2
Average Gamma Correlations as a Function of Location and Judgment Type



Note. Error bars are 95% bias corrected and accelerated (BCa) confidence intervals. More positive coefficients represent better metacognitive accuracy.

locations, and between judgment types. Forty-three of the 194 participants performed every trial correctly in either the internal (16.0%) or external condition (8.8%), and therefore had no estimates of accuracy when incorrect. These participants were removed for the purposes of this analysis.

A 2 (Actual Accuracy: Correct vs. Incorrect) \times 2 (Knowledge Location: Internal vs. External) \times 2 (Judgment Type: Prospective vs. Retrospective) ANOVA was performed to compare the ability to discriminate across the two memory locations. There was a significant two-way interaction between actual accuracy and memory location, $F(1, 149) = 113.44, p < .001, \eta_G^2 = .096$. There was no three-way interaction between actual accuracy, memory location, and judgment type, $F(1, 149) = 1.91, p = .17, \eta_G^2 = .002$ (see Figure 3). The ability of individuals' to discriminate between right and wrong answers was thus overall worse in the external condition. As is apparent in Figure 3, judged accuracy when incorrect drew closer to judged accuracy when correct integrating the internet. This is in contrast to the better discrimination as demonstrated in the internal condition.

Correlations Across Internal and External Memory Locations

Bias (i.e., participants' average judged accuracy minus their actual accuracy) and discrimination scores (i.e., participants' judged accuracy when correct minus incorrect) were computed for each location. There was a small significant positive correlation in bias scores across the internal and external knowledge conditions, $r_s(192) = .17, p = .018$. This suggests that individuals who were over/underconfident in performance when using their internal

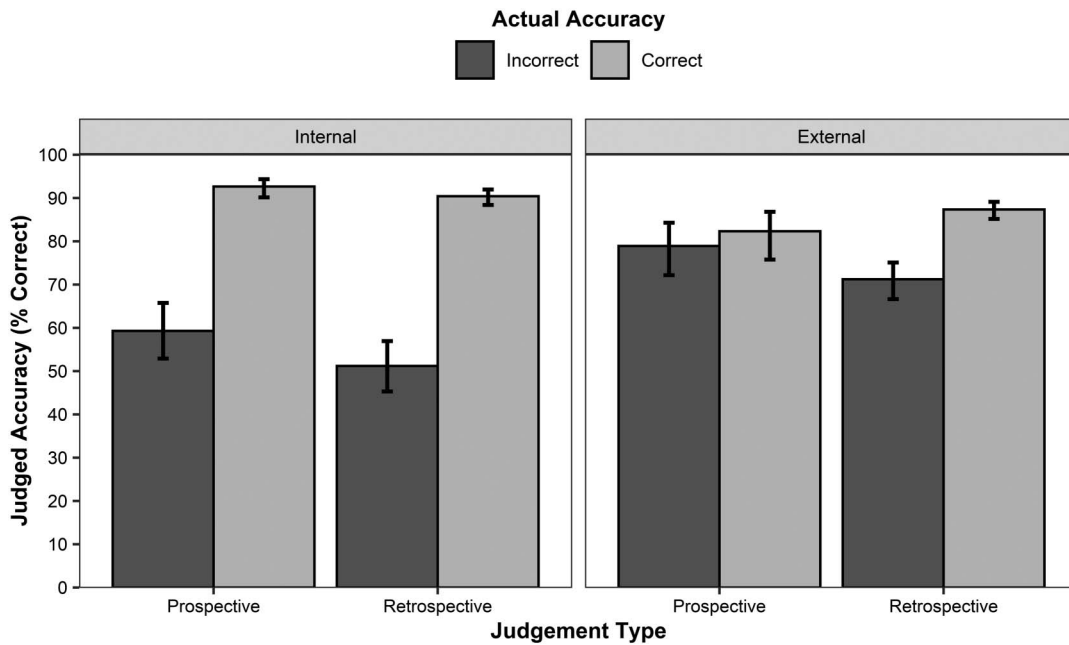
knowledge were also over/underconfident when integrating the internet. There was no correlation between discrimination ability across memory locations, $r_s(149) = -.01, p = .935$. We additionally correlated participants' Gamma coefficients in the internal and external memory conditions with one another. Overall there was no correlation in Gamma across the internal and external conditions, $r(139) = .08, p = .34$.³

General Discussion

Our day-to-day lives are filled with cognitive acts that consist of both internal and external contributions. Here, we examined individuals' metacognitions in an experiment where metacognitive judgments were provided across distributed (internal processes and the internet) and nondistributed (internal processes only) conditions. Whether individuals gave judgments prospectively or retrospectively was manipulated between subjects. First, individuals were overall underconfident when using their memory alone, and slightly underconfident prospectively and significantly overconfident retrospectively when using their memory and the internet. Second, Gamma correlations were overall higher for the internal location and retrospective judgment conditions. Interestingly, the observed interaction between location and judgment type highlights a large

³ We additionally performed two correlational analyses on Gamma coefficients across locations based on the two scenarios outlined in Footnote 2. In the first scenario where those who produced gamma coefficients of -1 were excluded, we observed a nonsignificant correlation across locations, $r(118) = .07, p = .423$. In the second scenario where those who produced gamma coefficients of -1 and/or $+1$ were excluded, we again observed a nonsignificant correlation, $r(77) = .04, p = .748$. Thus, these results did not differ from those observed originally.

Figure 3
Judged Accuracy as a Function of Actual Accuracy, Location, and Judgment Type



Note. Error bars are 95% bias corrected and accelerated (BCa) confidence intervals. Better discrimination would be represented by a larger difference between correct and incorrect accuracy within judgment types (e.g., correct answers being judged as more accurate than incorrect answers).

deficit in one form of metacognitive accuracy when making prospective judgments when using the internet. Third, participants in the external condition were less able to discriminate correct from incorrect responses than in the internal condition highlighting an additional deficit in metacognitive sensitivity. Finally, there was a significant small positive correlation in bias scores (i.e., the extent of over/underconfidence) across locations, but not in discrimination scores or Gamma correlations. In the following, we further discuss these patterns of results and situate them within the novel area of distributed metacognition.

Distributed Metacognitive Judgments

We have suggested two reasons for why metacognitive accuracy could be lower in a distributed context: (1) a lack of internally generated mnemonic cues and/or (2) a general bias to assume external cognitive aids are highly reliable. This general prediction bore out with respect to metacognitive accuracy. Gamma correlations and discrimination were overall worse in the external condition relative to the internal condition. When individuals in the external condition were incorrect they tended to think they were correct and confidence levels did not track accordingly. For instance, when individuals searched and found an incorrect answer, their average estimate of the likelihood that their answer was correct was 75%. In addition, average Gamma correlations were lowest in the external condition when individuals made prospective judgments (i.e., predicting memory performance). This metacognitive cost is likely due a combination of (1) and (2) above, given such judgments rely heavily on theory-based judgments of the internet as highly reliable

and a lack of in-task experience with the search process of the items before making a judgment.

Though metacognitive accuracy was overall worse while using the internet, external retrospective judgments were more accurate than prospective ones. This suggests that the recent experience of retrieving information from the internet provides at least some useful cues when making judgments. Recent research has pointed to multiple candidates for this source of information. Stone and Storm (2021) proposed that “search fluency” provides a feeling of strong and immediate access to information that is akin to internal retrieval fluency. Risko et al. (2016) proposed the “feeling-of-findability” as a cue reflecting preexisting beliefs of the factors that lead to relatively successful and unsuccessful searches. In the current context, both cues may work together to shore up metacognitive accuracy retrospectively. In these examples, however, both cues are subject to biases in the same way that some internal cues are affected. With respect to “search fluency,” individuals may misattribute access to answers to internal memory rather than to the external aid (Koriat, 2000; Marsh & Rajaram, 2019; Stone & Storm, 2021).

Indeed, individuals’ were significantly overconfident in their performance when making retrospective judgments after using the internet even in light of relatively better metacognitive accuracy. Increased metacognitive bias could be compounded by known shortfalls in individuals’ abilities to judge the validity of information when searching, especially when multiple sources are involved (e.g., Roulet, 2006; Stadler & Bromme, 2007; Walraven et al., 2009), further leading to their actual performance also being worse when using the internet to retrieve answers. That is, there are (at least) two biases working against individuals when using the

internet: not being able to accurately judge information and being overconfident in the information that was selected.

Taken together, the present work highlights detrimental aspects of the internet on metacognition. Future efforts should look toward identifying strategies to improve metacognitive accuracy when using the internet. This is of critical importance given the internet's ubiquity, and the roles of metacognitive ability in controlling behaviors related to learning (Nelson & Narens, 1990). Research in education and information science often describe self-regulated learning activities on the internet as "Searching as Learning" (SAL; Kuhlthau, 1991; Kuhlthau et al., 2008). Within this framework, learning is viewed as a constructive process including trying different search terms, distinguishing between relevant and nonrelevant sources, tuning knowledge structures, and identifying facts (Vakkari, 2016). Such processes can be prone to biases such as overall increases in confidence when conducting searches, including for when errors are committed (e.g., von Hoyer et al., 2019). Faulty metacognitions as they relate to SAL pose real-world risks to the control of behaviors, given miscalibration can hinder deep study of materials and retention (Dunlosky & Rawson, 2012). For instance, individuals may end their searches prematurely due to overconfidence in initial results (e.g., Bjork et al., 2013). It is especially critical to understand potential control issues like this, and similar ones, during search as misinformation runs rampant on the internet (Pennycook & Rand, 2021). Some optimism toward mitigation strategies exist, however, as recent evidence suggests some capacity to enhance domain-general metacognitive ability through adaptive methods with feedback (Carpenter et al., 2019).

Domain-General Metacognitive Ability

Though there are substantial differences between the internal and external resources available in the conditions used presently, there was a correlation across the internal and external conditions in terms of metacognitive accuracy. Thus, judgments made across the distributed and nondistributed cognitive tasks are not completely independent and may signal a common contribution to both. Specifically, bias (i.e., the difference between actual and judged accuracy) correlated significantly across the internal and external conditions. The more overconfident (or underconfident) individuals were in the internal condition the more overconfident (or underconfident) they were in the external condition, though the effect was rather small and undoubtedly requires replication. Nonetheless, this finding provides a novel extension of previous research demonstrating correlations in bias across different internal tasks (Kelemen et al., 2000). For instance, this correlation might reflect an individual's general confidence. Interestingly, although bias correlated across internal and external conditions, the ability to discriminate did not. Whether an individual was good or bad at discriminating correct from incorrect responses in the internal condition did not predict whether they would be good or bad in the external condition. This result also replicates and extends previous research (Kelemen et al., 2000) demonstrating little in the way of correlations between discrimination on different tasks within individuals. The extension here is interesting as the general tasks are ostensibly similar, but carried out using a different set of procedures rather than completely different tasks. Overall, these results are consistent with some skepticism regarding a domain general metacognitive ability (e.g., Kelemen et al., 2000; cf., Gilbert, 2015; Mazancieux et al., 2020).

Limitations of the Present Study

First, we note the limitations related to the sample used. All participants were adult U.S. citizens working on Amazon Mechanical Turk, though the participant pool is known to not be representative of the U.S. population as a whole (Arditte et al., 2016). We did not ask additional demographic questions, minimizing the generalizability of the current results. Extending lines of research in distributed metacognition to encompass individual differences, for example, across and within specific age and gender groups will be critical. For example, it has been demonstrated that older adults are slower and have more difficulties when interacting with search engines (Chevalier et al., 2015; Sharit et al., 2015). Furthermore, search behaviors can vary as a function of gender. Lorigo et al. (2006) demonstrated that males view search results further within a list and view results in a more linear order relative to females. Thus, in both cases, it may also be the case that metacognitions would vary alongside age and gender. Second, we note the limitations of the general knowledge items used in the present study. Many items presented were specific to the U.S. (e.g., "What is the last name of the man who was most responsible for photographing the U.S. Civil war?"). This might also minimize generalizability. Future work should expand question sets to include those specific to different countries, cultures, and expertise. Moreover, though renormed in 2013, the general knowledge questions pulled from Tauber et al. (2013) may be outdated. New question sets should be constructed and tested to include knowledge more relevant to younger generations. Addressing the two methodological points considered here, among others, will lead to a more complete account of distributed metacognition.

Conclusion

The present investigation represents one of the first systematic analyses of metacognition across a task completed with and without the aid of an external aid. Future work examining these types of tasks, as well as expanding to other domains, will further inform our understanding of metacognition in distributed contexts. With increases in integration between cognition and various cognitive technologies the need to understand the similarities and differences is of great theoretical and practical and interest.

References

- Ackerman, R., & Thompson, V. A. (2017). Meta-reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences*, 21(8), 607–617. <https://doi.org/10.1016/j.tics.2017.05.004>
- Arditte, K. A., Çek, D., Shaw, A. M., & Timpano, K. R. (2016). The importance of assessing clinical phenomena in Mechanical Turk research. *Psychological Assessment*, 28(6), 684–691. <https://doi.org/10.1037/pas0000217>
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37(3), 379–384. <https://doi.org/10.3758/BF03192707>
- Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, 127(1), 55–68. <https://doi.org/10.1037/0096-3445.127.1.55>
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64(1), 417–444. <https://doi.org/10.1146/annurev-psych-113011-143823>

- Boldt, A., & Gilbert, S. J. (2019). Confidence guides spontaneous cognitive offloading. *Cognitive Research: Principles and Implications*, 4(1), Article 45. <https://doi.org/10.1186/s41235-019-0195-y>
- Canty, A., & Ripley, B. (2016). *boot: Bootstrap R (S-Plus) functions* (R package version 1.3-18). <https://CRAN.R-project.org/package=boot>
- Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., & Fleming, S. M. (2019). Domain-general enhancements of metacognitive ability through adaptive training. *Journal of Experimental Psychology: General*, 148(1), 51–64. <https://doi.org/10.1037/xge0000505>
- Chevalier, A., Dommes, A., & Marquié, J. C. (2015). Strategy and accuracy during information search on the Web: Effects of age and complexity of the search questions. *Computers in Human Behavior*, 53, 305–315. <https://doi.org/10.1016/j.chb.2015.07.017>
- Clark, A. (1997). *Being there*. MIT Press.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195333213.001.0001>
- Cohen, J. (1988). *Statistical power analyses for the social sciences*. Lawrence Erlbaum Associates.
- Dunlosky, J., Mueller, M., & Tauber, S. K. (2014). The contribution of processing fluency (and beliefs) to people's judgments of learning. In D. S. Lindsay, C. M. Kelley, A. P. Yonelinas, & H. L. Roediger, III (Eds.), *Remembering: Attributions, processes, and control in human memory: Papers in honour of Larry L. Jacoby* (pp. 46–64). Psychology Press.
- Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22(4), 271–280. <https://doi.org/10.1016/j.learninstruc.2011.08.003>
- Dunn, T. L., Gaspar, C., McLean, D., Koehler, D. J., & Risko, E. F. (2021). *Distributed metacognition: Examining the metacognitions associated with retrieval from internal stores and the internet* [Data file and code]. Open Science Framework. <https://doi.org/10.17605/OSF.IO/K6JPW>
- Dunn, T. L., & Risko, E. F. (2016). Toward a metacognitive account of cognitive offloading. *Cognitive Science*, 40(5), 1080–1127. <https://doi.org/10.1111/cogs.12273>
- Ferguson, A. M., McLean, D., & Risko, E. F. (2015). Answers at your fingertips: Access to the internet influences willingness to answer questions. *Consciousness and Cognition*, 37, 91–102. <https://doi.org/10.1016/j.concog.2015.08.008>
- Fisher, M., Goddu, M. K., & Keil, F. C. (2015). Searching for explanations: How the internet inflates estimates of internal knowledge. *Journal of Experimental Psychology: General*, 144(3), 674–687. <https://doi.org/10.1037/xge0000070>
- Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: Computation, biology and function. *Philosophical Transactions of the Royal Society*, 367(1594), 1280–1286. <https://doi.org/10.1098/rstb.2012.0021>
- Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, 8, Article 443. <https://doi.org/10.3389/fnhum.2014.00443>
- Fleming, S. M., Massoni, S., Gajdos, T., & Vergnaud, J. C. (2016). Metacognition about the past and future: Quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neuroscience of Consciousness*, 2016(1), Article niw018. <https://doi.org/10.1093/nc/niw018>
- Gilbert, S. J. (2015). Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 68(5), 971–992. <https://doi.org/10.1080/17470218.2014.972963>
- Gray, W. D., Sims, C. R., Fu, W. T., & Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review*, 113(3), 461–482. <https://doi.org/10.1037/0033-295X.113.3.461>
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24(2), 95–112. <https://doi.org/10.1007/BF02289823>
- Harrell, F. E., Jr. (2016). *Hmisc: Harrell miscellaneous* (R package version 4.0-1). <https://CRAN.R-project.org/package=Hmisc>
- Heersmink, R. (2013). A taxonomy of cognitive artifacts: Function, information, and categories. *Review of Philosophy and Psychology*, 4(3), 465–481. <https://doi.org/10.1007/s13164-013-0148-1>
- Hu, X., Luo, L., & Fleming, S. M. (2019). A role for metamemory in cognitive offloading. *Cognition*, 193, Article 104012. <https://doi.org/10.1016/j.cognition.2019.104012>
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
- Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science*, 2(4), 705–715. <https://doi.org/10.1111/j.1756-8765.2010.01089.x>
- Kelemen, W. L., Frost, P. J., & Weaver, C. A., III (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory & Cognition*, 28(1), 92–107. <https://doi.org/10.3758/BF03211579>
- Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, 32(1), 1–24. <https://doi.org/10.1006/jmla.1993.1001>
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126(4), 349–370. <https://doi.org/10.1037/0096-3445.126.4.349>
- Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, 9(2 Pt. 1), 149–171. <https://doi.org/10.1006/ccog.2000.0433>
- Koriat, A., & Adiv, S. (2012). Confidence in one's social beliefs: Implications for belief justification. *Consciousness and Cognition*, 21(4), 1599–1616. <https://doi.org/10.1016/j.concog.2012.08.008>
- Koriat, A., & Bjork, R. A. (2006). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5), 1133–1145. <https://doi.org/10.1037/0278-7393.32.5.1133>
- Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5), 361–371. [https://doi.org/10.1002/\(SICI\)1097-4571\(199106\)42:5<361::AID-ASIS6>3.0.CO;2-%23](https://doi.org/10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASIS6>3.0.CO;2-%23)
- Kuhlthau, C. C., Heinström, J., & Todd, R. J. (2008). The “information search process” revisited: Is the model still useful. *Information Research*, 13(4), 13–14.
- Lawrence, M. A. (2015). *ez: Easy analysis and visualization of factorial experiments* (R package version 4.3). <https://CRAN.R-project.org/package=ez>
- Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L., & Gay, G. (2006). The influence of task and gender on search and evaluation behavior using Google. *Information Processing & Management*, 42(4), 1123–1131. <https://doi.org/10.1016/j.ipm.2005.10.001>
- Maki, R. H., Shields, M., Wheeler, A. E., & Zaccchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology*, 97(4), 723–731. <https://doi.org/10.1037/0022-0663.97.4.723>
- Marsh, E. J., & Rajaram, S. (2019). The digital expansion of the mind: Implications of internet usage for memory and cognition. *Journal of Applied Research in Memory and Cognition*, 8(1), 1–14. <https://doi.org/10.1016/j.jarmac.2018.11.001>
- Mazancieux, A., Fleming, S. M., Souchay, C., & Moulin, C. J. A. (2020). Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks. *Journal of Experimental Psychology: General*, 149(9), 1788–1799. <https://doi.org/10.1037/xge0000746>
- Mengelkamp, C., & Bannert, M. (2010). Accuracy of confidence judgments: Stability and generality in the learning process and predictive validity for

- learning outcome. *Memory & Cognition*, 38(4), 441–451. <https://doi.org/10.3758/MC.38.4.441>
- Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, 15(1), 174–179. <https://doi.org/10.3758/PBR.15.1.174>
- Michaelian, K. (2012). Is external memory memory? Biological memory and extended mind. *Consciousness and Cognition*, 21(3), 1154–1165. <https://doi.org/10.1016/j.concog.2012.04.008>
- Michaelian, K., & Sutton, J. (2013). Distributed cognition and memory research: History and current directions. *Review of Philosophy and Psychology*, 4(1), 1–24. <https://doi.org/10.1007/s13164-013-0131-x>
- Mueller, M. L., Dunlosky, J., & Tauber, S. K. (2016). The effect of identical word pairs on people's metamemory judgments: What are the contributions of processing fluency and beliefs about memory? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 69(4), 781–799. <https://doi.org/10.1080/17470218.2015.1058404>
- Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, 95(1), 109–133. <https://doi.org/10.1037/0033-2909.95.1.109>
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 1–45). Academic Press.
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>
- Proust, J. (2008). Epistemic agency and metacognition: An externalist view. *Proceedings of the Aristotelian Society*, 108(1 Pt. 3), 241–268. <https://doi.org/10.1111/j.1467-9264.2008.00245.x>
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reber, R., & Unkelbach, C. (2010). The epistemic status of processing fluency as source for judgments of truth. *Review of Philosophy and Psychology*, 1(4), 563–581. <https://doi.org/10.1007/s13164-010-0039-7>
- Revelle, W. (2016) *psych: Procedures for personality and psychological research* (R package version 1.6.9). <https://CRAN.R-project.org/package=psych>
- Risko, E. F., & Dunn, T. L. (2015). Storing information in-the-world: Metacognition and cognitive offloading in a short-term memory task. *Consciousness and Cognition*, 36, 61–74. <https://doi.org/10.1016/j.concog.2015.05.014>
- Risko, E. F., Ferguson, A., & McLean, D. (2016). On retrieving information from external knowledge stores: Feeling-of-findability, feeling-of-knowing and internet search. *Computers in Human Behavior*, 65, 534–543. <https://doi.org/10.1016/j.chb.2016.08.046>
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>
- Roulet, J. F. (2006). *The skills of document use: From text comprehension to Web-based learning*. Psychology Press. <https://doi.org/10.4324/9780203820094>
- Sharit, J., Taha, J., Berkowsky, R. W., Profita, H., & Czaja, S. J. (2015). Online information search performance and search strategies in a health problem-solving scenario. *Journal of Cognitive Engineering and Decision Making*, 9(3), 211–228. <https://doi.org/10.1177/1555343415583747>
- Siedlecka, M., Paulewicz, B., & Wierchoń, M. (2016). But I was so sure! Metacognitive judgments are less accurate given prospectively than retrospectively. *Frontiers in Psychology*, 7, Article 218. <https://doi.org/10.3389/fpsyg.2016.00218>
- Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333(6043), 776–778. <https://doi.org/10.1126/science.1207745>
- Stadler, M., & Bromme, R. (2007). Dealing with multiple documents on the WWW: The role of metacognition in the formation of documents models. *International Journal of Computer-Supported Collaborative Learning*, 2(2), 191–210. <https://doi.org/10.1007/s11412-007-9015-3>
- Sterelny, K. (2010). Minds: Extended or scaffolded? *Phenomenology and the Cognitive Sciences*, 9(4), 465–481. <https://doi.org/10.1007/s11097-010-9174-y>
- Stone, S. M., & Storm, B. C. (2021). Search fluency as a misleading measure of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(1), 53–64. <https://doi.org/10.1037/xlm0000806>
- Sutton, J. (2010). Exograms and interdisciplinarity: History, the extended mind and the civilizing process. In R. Menary (Ed.), *The extended mind* (pp. 189–225). MIT Press. <https://doi.org/10.7551/mitpress/9780262014038.003.0009>
- Tauber, S. K., Dunlosky, J., Rawson, K. A., Rhodes, M. G., & Sitzman, D. M. (2013). General knowledge norms: Updated and expanded from the Nelson and Narens (1980) norms. *Behavior Research Methods*, 45(4), 1115–1143. <https://doi.org/10.3758/s13428-012-0307-9>
- Thiede, K. W., Anderson, M., & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, 95(1), 66–73. <https://doi.org/10.1037/0022-0663.95.1.66>
- Torchiano, M. (2016). *effsize: Efficient effect size computation* (R package version 0.7.1). <https://CRAN.R-project.org/package=effsize>
- Vakkari, P. (2016). Searching as learning: A systematization based on literature. *Journal of Information Science*, 42(1), 7–18. <https://doi.org/10.1177/0165551515615833>
- von Hoyer, J., Pardi, G., Kammerer, Y., & Holtz, P. (2019). Metacognitive judgments in searching as learning (SAL) tasks: Insights on (mis-) calibration, multimedia usage, and confidence. In *Proceedings of the 1st international workshop on search as learning with multimedia information* (pp. 3–10). Association for Computing Machinery
- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. (2009). How students evaluate information and sources when searching the World Wide Web for information. *Computers & Education*, 52(1), 234–246. <https://doi.org/10.1016/j.compedu.2008.08.003>
- Ward, A. F. (2013). Supernormal: How the internet is changing our memories and our minds. *Psychological Inquiry*, 24(4), 341–348. <https://doi.org/10.1080/1047840X.2013.850148>
- Weis, P. P., & Wiese, E. (2019). Using tools to help us think: Actual but also believed reliability modulates cognitive offloading. *Human Factors*, 61(2), 243–254. <https://doi.org/10.1177/0018720818797553>
- Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12), 1–20. <https://doi.org/10.18637/jss.v021.i12>
- Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636. <https://doi.org/10.3758/BF03196322>

Received March 10, 2021

Revision received May 15, 2021

Accepted May 17, 2021 ■