Are the Contributions of Processing Experience and Prior Beliefs to Confidence Ratings Domain-General or Domain-Specific?

Xiao Hu[2], Chunliang Yang[1, 2], and Liang Luo[1, 3]

[1] Institute of Developmental Psychology, Faculty of Psychology, Beijing Normal University

[2] Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education, Faculty of Psychology, Beijing Normal University

[3] Collaborative Innovation Center of Assessment for Basic Education Quality, Beijing Normal University

Word count: 10,109

Author Note

This study was supported by the Natural Science Foundation of China (32171045, 32000742). We have no conflicts of interest to disclose. Data and scripts for the current study are available online at OSF (https://osf.io/92qjr/).

Correspondence concerning this article should be addressed to Liang Luo, Institute of Developmental Psychology and Collaborative Innovation Center of Assessment for Basic Education Quality, Beijing Normal University, Beijing 100875, China. Email: luoliang@bnu.edu.cn

Abstract

Previous studies on domain generality of metacognition showed inconsistent results about cross-domain correlation of metacognitive resolution, which might result from the varied relationship between actual performance and the information utilized during confidence rating across tasks. The current study investigated metacognitive domain generality using the Bayesian inference model for metamemory (BIM), which suggests that individuals integrate current processing experience and their prior beliefs to construct confidence ratings. Results from three experiments and a series of meta-analyses showed that the correlation between the contribution of processing experience to confidence ratings (parameter P_{exp} in BIM) across perceptual and memory domains was significantly positive, while the cross-domain correlation of metacognitive resolution (meta-*d'*/*d'*) was relatively weak. Furthermore, meta-*d'*/*d'* was related to specific task requirements, which could lead to very low cross-task correlation of meta-*d'*/*d'* even within the same cognitive domain. These results imply that the cross-domain correlation of metacognitive resolution might underestimate metacognitive domain generality, and the cognitive mechanisms underlying confidence rating process itself may be more domain-general.

Keywords: metacognition, confidence rating, domain generality, perception, memory

Are the Contributions of Processing Experience and Prior Beliefs to Confidence Ratings Domain-General or Domain-Specific?

Metacognition is the introspective ability to monitor and control one's mental processes (Metcalfe & Shimamura, 1994). Metacognition could be involved in various cognitive domains: People often evaluate their performance in different cognitive processes such as perception and memory (Fleming & Dolan, 2012). Many studies have shown that metacognitive evaluations of performance can guide subsequent learning process and actions such as help-seeking behavior (Desender et al., 2018; Metcalfe & Kornell, 2005; Pescetelli et al., 2021; Son & Metcalfe, 2000; Tullis & Benjamin, 2012). Thus, the study of cognitive mechanisms underlying metacognition has received great attention in recent years (e.g., Arbuzova et al., 2021; Carpenter et al., 2019; Fitzgerald et al., 2017; Lee et al., 2018; Mazancieux et al., 2020; Shekhar & Rahnev, 2021).

An important question in metacognition research is whether the cognitive process underlying metacognition is general or specific across different cognitive domains (Rouault et al., 2018). To explore this critical question, researchers asked participants to rate their confidence about performance in cognitive tasks from different domains (e.g., perceptual and memory tasks), and then tested whether metacognitive resolution (i.e., the degree to which confidence ratings distinguish between correct and incorrect trials in a given task; see Vuorre & Metcalfe, 2021) across different cognitive domains correlates with each other. However, previous results about whether there is domain generality of metacognitive resolution are largely inconsistent (e.g., Baird et al., 2013, 2015; McCurdy et al., 2013; Samaha & Postle, 2017).

Recently, Hu et al. (2021) proposed a Bayesian inference model (BIM) to explain the confidence rating process in memory tasks (i.e., metamemory). BIM assumes that when rating their confidence, people integrate current processing experience in the task and their prior beliefs about overall ability in performing the task through a Bayesian inference. In addition, BIM can quantitatively characterize how much processing experience and prior beliefs contribute to the confidence rating process. In the current study, we aimed to extend BIM to confidence ratings in perceptual tasks, and examine the domain generality of metacognition based on the contribution of processing experience (relative to prior beliefs) to confidence ratings across domains.

Domain Generality of Metacognition

In the field of metacognition, it is of great interest regarding whether the cognitive mechanisms of metacognition are domain-specific or domain-general. To answer this question, researchers often examine the correlation between metacognitive resolution across different domains (Rouault et al., 2018). A widely-used measure of metacognitive resolution is meta-*d'*/*d'* (also termed as metacognitive efficiency). The meta-*d'* reflects the value of Type I performance that would lead to the observed data of confidence ratings assuming the absence of noise in confidence process. By dividing meta-*d'* by the actual Type I performance *d'* in a task, meta-*d'*/*d'* models the relationship between task performance and confidence ratings when controlling the effect of task performance itself (Fleming, 2017; Maniscalco & Lau, 2012).

Previous research findings about whether there is a cross-domain correlation of metacognitive resolution are largely inconsistent. While some studies observed a significant correlation of meta-*d'*/*d'* between perceptual and memory tasks (McCurdy et al., 2013; Samaha & Postle, 2017), other studies detected no relation (Baird et al., 2013, 2015; Morales et al., 2018). A meta-analysis conducted by Rouault et al. (2018) found that the correlation between metacognitive resolution in perceptual and memory domains did not reach statistical significance. However, Mazancieux et al. (2020) recently employed a large sample size (181 participants) and observed a significant correlation between meta-*d'*/*d'* in visual perception, semantic memory and working memory tasks, even though it is worth noting that the correlation between meta-*d'*/*d'* in perceptual and episodic memory tasks was not statistically significant.

A possible explanation for these inconsistent findings is that previous studies on domain generality of metacognition used different types of perceptual and memory tasks (such as orientation vs. contrast discrimination task in perceptual domain, and short-term vs. long-term memory task in memory domain; see Baird et al., 2013; Mazancieux et al., 2020; Samaha & Postle, 2017), and metacognitive resolution may depend on the specific requirements of a given task. According to the cue-utilization theory, metacognitive resolution is determined by whether the information (or cues) utilized to inform confidence ratings could accurately reflect actual task performance

(Koriat, 1997). Across different types of tasks, the information that people utilize to construct confidence ratings may predict actual task performance to different extent, leading to variation in metacognitive resolution.

Consistent with this hypothesis, Samaha and Postle (2017) found that while meta-*d'*/*d'* in perceptual and short-term memory tasks were significantly correlated when both tasks focused on the orientation of stimulus, meta-*d'*/*d'* in two perceptual tasks were uncorrelated when one task engaged in orientation discrimination and the other engaged in contrast discrimination. Consistently, results from Arbuzova et al. (2021) revealed that while metacognitive resolution was correlated between a visual perception and a motor task, it was not correlated between two motor tasks involving different aspects of movement. These findings jointly point out to the inference that different task requirements (or different information utilized during confidence rating formation) may reduce the cross-task correlation of metacognitive resolution even when the two tasks come from the same cognitive domain. Based on these results, it is reasonable to assume that the correlation of metacognitive resolution across cognitive domains might underestimate the true domain generality of metacognition, because in perceptual and memory tasks people integrate different information when providing confidence ratings, which may predict actual performance to different extent depending on the task requirements.

Thus, it should be better to investigate metacognitive domain generality based on the measures solely reflecting the confidence rating process itself rather than the relationship between confidence ratings and actual task performance. It is possible that while the relationship between actual performance and the information utilized during confidence rating can vary across domains, the cognitive mechanisms underlying confidence rating process itself might be more domain-general.

Bayesian Inference Model for Metamemory

An available measure that solely reflects the confidence process itself comes from the recently proposed Bayesian inference model for metamemory (BIM) (Hu et al., 2021). Metamemory refers to the processes for monitoring and controlling memory activities (Nelson & Narens, 1990). One of the most influential theories in the metamemory literature is the dual-basis theory (Koriat et al., 2004). This theory proposes that metacognitive evaluation of memory performance relies on both (1) the processing experience obtained from the encoding or retrieval process, and (2) people's prior beliefs about their overall memory ability. However, few studies have proposed a formal computational model to quantitatively explain how processing experience and prior beliefs are utilized during metamemory evaluation (but see Jang et al., 2012).

To address this question, Hu et al. (2021) proposed BIM, which quantitatively characterize metamemory process based on the dual-basis theory. BIM assumes that the subjective processing experience e for each item is sampled from a distribution of processing experience for each individual, which is correlated with the distribution of objective memory strength m. During confidence rating process, people apply Bayesian inference to infer their memory strength $\hat{m}$ for each item (which may be different from the actual memory strength *m*) based on both the current processing experience *e* for each item and their overall prior beliefs about memory ability (see Figure 1). In BIM, the prior belief about memory performance is characterized by a prior belief distribution about inferred memory strength $\hat{m}$, and people believe they have higher memory performance when the mean of prior belief distribution is higher. Through the Bayesian inference process, people can obtain the posterior distribution of inferred memory strength $\hat{m}$ for each item given the processing experience *e*, which then forms the basis of confidence rating.

The contribution of processing experience and prior beliefs to the Bayesian inference process is determined by the likelihood function which encodes people's knowledge about the relationship between the inferred memory strength $\hat{m}$ and processing experience *e* (see Figure 1). When the standard deviation of likelihood function (σ_l) is smaller, people assume a closer relationship between $\hat{m}$ and *e*, and thus rely more on processing experience and less on prior beliefs during confidence rating (and vice versa). In BIM, the proportion for the contribution of processing experience (P_{exp}) and prior beliefs (P_{belief}) to confidence rating process can be calculated based on σ_l (Hu et al., 2021):

$$P_{\exp} = \frac{1}{1+\sigma_l^2}$$

$$P_{belief} = \frac{\sigma_l^2}{1+\sigma_l^2}$$

The sum of P_{exp} and P_{belief} is 1. In BIM, we typically use P_{exp} to indicate the relative contribution of processing experience and prior beliefs to confidence ratings, and the effect of processing experience on confidence is larger (and the effect of prior beliefs is smaller) when P_{exp} is higher. The P_{exp} does not reflect the strength of people's processing experience in memory task, or the estimated memory performance in people's prior beliefs. Instead, it represents how much people rely on experience and beliefs during confidence process.

Results from data simulation revealed that the value of P_{exp} affects the shape of confidence distribution, and the variance of confidence ratings across trials is smaller when P_{exp} is smaller (and vice versa) (Hu et al., 2021). This is because BIM assumes when a participant rates their confidence mainly based on prior belief rather than processing experience, their confidence ratings should be closely distributed around this prior belief about overall memory ability and the variance of the confidence distribution should be small. In contrast, when processing experience contributes more to confidence ratings, the confidence distribution should have greater variability due to variable processing experience across items. Thus, P_{exp} affects the usage of confidence scale in metamemory tasks.

Although BIM is primarily proposed to explain the cognitive mechanisms of metamemory, it is reasonable to expect that the assumption about the integration of processing experience and prior beliefs may also be suitable to metacognitive process in other cognitive domains such as perception. While previous theoretical models on perceptual metacognition mainly suggest that confidence ratings in perceptual tasks rely on processing experience (e.g., Maniscalco & Lau, 2012; Pleskac & Busemeyer, 2010; Shekhar & Rahnev, 2021), some theories point out the possible role of prior beliefs in perceptual metacognition. For example, in their second-order model, Fleming and Daw (2017) proposed that people's beliefs about the internal states during perceptual process may be malleable, which could lead to systematic change of confidence ratings. Recent empirical studies further suggest that we could develop global beliefs about task performance based on item-by-item confidence ratings in perceptual tasks, and this global belief might then affect our subsequent confidence ratings and behavior in the task (Lee et al., 2021; Rouault et al., 2019; Rouault & Fleming, 2020). However, few studies have directly investigated how much processing experience and prior beliefs contribute to perceptual metacognition. Thus, the first purpose of the current study is to extend BIM to perceptual metacognition and examine the contribution of processing experience and prior beliefs to confidence ratings (i.e., P_{exp}) across cognitive domains (perception and memory).

Compared with metacognitive resolution which represents the relationship between confidence ratings and actual performance, P_{exp} solely reflects the integration of processing experience and prior beliefs during confidence rating process itself (Hu et al., 2021). Thus, the cross-domain correlation of P_{exp} (rather than that of metacognitive resolution) may better reflect whether the cognitive mechanisms underlying confidence rating process are domain-general or domain-specific. Hence, the second purpose of the current study is to examine the correlation between P_{exp} across perception and memory domains, and compare the cross-domain correlation of P_{exp} with that of metacognitive resolution (meta-*d'*/*d'*).

The Current Study

The current study contained three experiments and a series of meta-analyses. In Experiments 1 and 2, participants were instructed to perform both a perceptual task and a memory task, and give item-by-item confidence ratings. The correlation of P_{exp} between the two tasks was computed to determine whether the cognitive process underlying the integration of processing experience and prior beliefs during confidence rating is domain-general or domain-specific. We also examined the correlation between metacognitive resolution (meta-*d'*/*d'*) across perceptual and memory domains. The reason for conducting two experiments was to test whether the cross-domain correlation of P_{exp} and meta-*d'*/*d'* generalizes to perceptual and memory tasks with different difficulty levels. In addition, we compared the cross-domain correlation of P_{exp} with that of meta-*d'*/*d'*. We expected that the correlation of P_{exp} across domains should be greater than that of meta-*d'*/*d'*, because the cross-domain correlation of meta-*d'*/*d'* might underestimate the domain generality of metacognition due to varied relationship between actual performance and the information utilized in confidence rating process across tasks.

Furthermore, in order to directly illustrate that meta-*d'*/*d'* might be affected by specific task requirements, we then conducted Experiment 3 in which participants were asked to perform two different perceptual tasks with confidence ratings. Finally, a series of meta-analyses, which synthesized data from six published experiments in the Confidence Database (Rahnev et al., 2020), was performed to further investigate the cross-domain correlation of P_{exp} across different types of perceptual and memory tasks, and to compare the domain generality of P_{exp} and meta-*d'*/*d'*.

Experiment 1

Method

The experimental materials, data and scripts for all experiments can be accessed at the OSF project page: https://osf.io/92qjr/ (Hu, 2022).

Participants

Participants in Experiment 1 were recruited online via the Prolific website (https://prolific.ac/). The sample size was determined using sequential hypothesis testing with Bayes factors (Dienes, 2016). Specifically, we selected a minimum sample size of 20, and defined the stopping rule as the point at which the Bayes factor (BF_{10}) for the correlation of P_{exp} between perceptual and memory tasks was greater than 10 (implying strong support for the alternative hypothesis) or lower than 1/10 (implying strong support for the null hypothesis) (Jeffreys, 1961). We calculated the Bayes factor after running 20 participants, and again after each additional 2 participants. The stopping rule was reached at 68 participants (35 female; age: M = 29.46 years, SD = 6.47). Data from another one participant were excluded due to that the *d'* (representing task performance; see the Data Analysis section for details) for this participant was smaller than 0 (i.e., lower than the chance level) in memory task.

Participants signed informed consent prior to the experiment, and received monetary compensation (£8.5) and a bonus (up to £1, dependent on task performance) after the experiment. All participants spoke English as the first language and reported normal or corrected-to-normal vision. All procedures were approved by the Ethics Committee at Beijing Normal University.

Materials and Procedure

Participants were asked to perform a perceptual task and a memory task. Task order was randomly assigned for each participant. Both tasks were programmed using Gorilla (Anwyl-Irvine et al., 2020).

The perceptual task used in Experiment 1 was a color discrimination task (see Figure 2A). In this task, a fixation was first presented for 500 ms in each trial, after which two rectangles were presented on the screen for 2000 ms. Each rectangle was split into two areas filled by either orange or blue color (same as the stimuli used in Lei et al., 2020). For one of the two rectangles, the sizes of blue and orange areas were equal. For the other one, the blue area was larger than the orange area (the difference between blue and orange area was between 2.5% and 16% of the total rectangle area, randomly defined in each trial). The positions (left or right) for the two rectangles were randomly determined in each trial. After stimuli presentation, participants needed to press the number key 1 or 2 to decide which rectangle contained larger blue area. The chosen answer would then be highlighted for 500 ms. Finally, participants were required to press a number key to indicate their confidence regarding answer correctness. The confidence scale ranged from 1 (not confident at all) to 6 (extremely confident). The chosen number in the confidence rating scale would then be highlighted for 500 ms. There was no time pressure on decision making or confidence rating. There were 200 trials in the perceptual task. Here we used a two-alternative-forced-choice (2AFC) task rather than the yes/no task (deciding which color area was larger in each rectangle) in Lei et al. (2020) because previous study suggests 2AFC task is more suitable for the study of domain generality of metacognition (Lee et al., 2018).

The experimental materials for the memory task were 400 English words selected from the MRC Psycholinguistic Database (Coltheart, 1981). All words contained 4-8 letters and 1-3 syllables. The Kucera and Francis word frequency for all words was between 8 and 90 (M = 32.38; SD = 20.83), and the ratings of familiarity (M = 537.66; SD = 39.72), concreteness (M = 561.10; SD = 47.46) and imagability (M = 565.74; SD = 43.10) were between 450 and 650. The words were randomly divided into two halves for each participant, in which 200 words were used as study materials in the learning phase, with the other 200 words served as lures in the 2AFC recognition memory test.

The procedure for the memory task was similar to that in previous studies (Fleming et al., 2014; McCurdy et al., 2013; Palmer et al., 2014). The memory task contained four blocks, and in each block participants were presented with 50 words simultaneously on the screen (arranged in 10 rows and 5 columns) and asked to memorize as many words as possible (see Figure 3A). The exposure duration of the words was 30, 50, 70 or 90 s in the four blocks, and the order of the four durations was randomized across participants. When there were 10 s left of the learning phase, participants were notified by a countdown clock presented below the 50 words.

Following the learning phase in each block, participants completed 50 trials in the 2AFC recognition memory test (see Figure 3C). After presenting the fixation for 500 ms, a learned word and a new word were simultaneously shown on the screen, and participants needed to decide which word had been learned. Then they were asked to rate their confidence by pressing a number key ranging from 1 to 6.

Data Analysis

We first used paired-sample *t* tests to compare task performance (represented by *d'* in signal detection theory) between the perceptual and memory tasks, and examined the correlation between *d'* in two tasks. We also investigated the effect of cognitive domain (perception vs. memory) on mean confidence level, and the cross-domain correlation of mean confidence.

We then used the recently developed *HMeta-d* package to estimate meta-*d'*/*d'*, which represents metacognitive resolution (Fleming, 2017). We separately estimated the meta-*d'*/*d'* in each task for each participant using the code *fit_meta_d_mcmc.m* in MATLAB. In *HMeta-d* package, the single-participant estimation of meta-*d'* was performed using Markov chain Monte Carlo (MCMC) methods implemented in JAGS (Plummer, 2003). We fit the model with 3 chains and each chain contained 11,000 samples. We discarded 1,000 samples per chain for burn-in, resulting in 30,000 stored samples in total. The estimated value of meta-*d'* for each participant was then divided by *d'*, which was calculated based on signal detection theory (Fleming, 2017). A paired-sample *t* test was conducted to compare the meta-*d'*/*d'* between the perceptual and memory tasks. We also examined the Pearson correlation coefficient of meta-*d'*/*d'* between the two tasks. We focused on the single-participant estimation of meta-*d'*/*d'* in the current study because the estimation of meta-*d'*/*d'* for each participant was needed when comparing the cross-domain correlation of P_{exp} and that of meta-*d'*/*d'* (see below for details). However, for completeness, we also conducted hierarchical Bayesian analysis to estimate the correlation between log meta-*d'*/*d'* across domains at group level (Fleming, 2017).

Next, we used BIM to estimate the parameter P_{exp} in each task for each participant based on maximum likelihood estimation (Hu et al., 2021). According to the assumptions of BIM, we computed the likelihood of observing the current data (task performance and confidence ratings) given the value of each parameter in BIM, including P_{exp} and the other parameters unrelated to the current study (see Hu et al., 2021 for the calculation). Then we fit BIM to the data separately from each participant, and found the parameter value that maximized the likelihood function. We examined the difference in P_{exp} between the perceptual and memory tasks, and the cross-domain correlation of P_{exp}.

We conducted both frequentist and Bayesian hypothesis tests when performing paired-sample *t* tests and correlation analyses described above. All Bayesian hypothesis tests were performed via JASP (http://www.jasp-stats.org). When examining the cross-domain correlation of meta-*d'*/*d'* and P_{exp}, we also performed Bayes factor robustness check in JASP to explore how Bayes factors changed with the shape of prior distribution for correlation coefficient (results from the Bayes factor robustness check in Experiments 1-3 are reported in Section S1 of Supplemental Materials).

Finally, we quantitatively compared the cross-domain correlation of P_{exp} and that of meta-*d'*/*d'* using the statistical method provided by Dunn and Clark (1969). We first computed the cross-domain correlation of P_{exp} (denoted by r_{12}) and that of meta-*d'*/*d'* (denoted by r_{34}), and performed Fisher's *r*-to-*z* transformation on r_{12} and r_{34} (the transformed correlation coefficients are denoted by z_{12} and z_{34}). We also calculated the correlation between P_{exp} and meta-*d'*/*d'* in the perceptual task (denoted by r_{13}), between P_{exp} and meta-*d'*/*d'* in the memory task (denoted by r_{24}), between P_{exp} in perceptual task and meta-*d'*/*d'* in memory task (denoted by r_{14}), and between P_{exp} in memory task and meta-*d'*/*d'* in perceptual task (denoted by r_{23}). Then we could construct a *Z* statistic for the difference between the *z*-transformed cross-domain correlation of P_{exp} and that of meta-*d'*/*d'*:

$$Z = \frac{z_{12} - z_{34}}{\sqrt{\frac{2(1-c)}{N-3}}}$$

In the equation above, *N* represents the sample size, and *c* is calculated as (Dunn & Clark, 1969):

$$c = [\frac{r_{12}r_{34}}{2}\left(r_{13}^{2} + r_{14}^{2} + r_{23}^{2} + r_{24}^{2}\right) + \left(r_{13}r_{24} + r_{23}r_{14}\right) - \left(r_{13}r_{23}r_{34} + r_{14}r_{24}r_{34} + r_{12}r_{13}r_{14} + r_{12}r_{23}r_{24}\right)] / \left(1 - r_{12}^{2}\right)\left(1 - r_{34}^{2}\right)$$

We could obtain the *p* value for the *Z* statistic by comparing the value of the *Z* statistic to standard normal distribution.

Results

Task performance and mean confidence

Task performance (represented by *d'*) was better in the memory task than in the perceptual task, $t(67) = 4.98$, $p < .001$, Cohen's $d = 0.60$, $BF_{10} > 100$ (see Figure 4A). In addition, *d'* in the perceptual and memory tasks did not correlate with each other, $r = -.01$, $p = .945$, $BF_{10} = 0.15$, indicating that there was no relationship between the task performance in perceptual and memory domains.

Mean confidence was also higher in the memory task than in the perceptual task, $t(67) = 3.55$, $p < .001$, Cohen's $d = 0.43$, $BF_{10} = 34.74$ (see Figure 4B). There was a significant correlation between mean confidence in the two tasks, $r = .58$, $p < .001$, $BF_{10} > 100$.

Metacognitive resolution

Metacognitive resolution, represented by meta-*d'*/*d'*, was greater in the memory task than that in the perceptual task, $t(67) = 6.90$, $p < .001$, Cohen's $d = 0.84$, $BF_{10} > 100$ (see Figure 4C). Of critical interests, the cross-domain correlation of meta-*d'*/*d'* based on single-participant estimation was statistically detectable according to *p* value (although the Bayes factor was inconclusive), $r = .27$, $p = .029$, $BF_{10} = 1.56$ (see Figure 5A). For completeness, we also estimated the cross-domain correlation of meta-*d'*/*d'* at group level based on the hierarchical Bayesian model in *HMeta-d* package. However, results from the hierarchical Bayesian model revealed that the cross-domain correlation of meta-*d'*/*d'* did not reach significance, $\rho = .20$, 95% credible interval (CrI) [-.22, .58]. These results suggested that although there might exist domain generality of metacognitive resolution, the relationship between metacognitive resolution across domains should be weak.

Contribution of processing experience to confidence ratings

Next, we turned to investigate the contribution of processing experience to confidence ratings (represented by the parameter P_{exp} in BIM) in perceptual and memory domains. Results revealed that processing experience contributed more to confidence ratings (i.e., P_{exp} was higher) in the memory task than in the perceptual task, $t(67) = 8.84$, $p < .001$, Cohen's $d = 1.07$, $BF_{10} > 100$ (see Figure 4D). Furthermore, the correlation between P_{exp} in the perceptual and memory tasks was significant, $r = .37$, $p = .002$, $BF_{10} = 14.80$ (see Figure 5B), indicating the contribution of processing experience to confidence ratings was domain-general.

Difference in domain generality between P_{exp} and meta-d'/d'

We then quantitatively compared the domain generality of metacognitive resolution (meta-*d'*/*d'*) and that of the contribution of processing experience to metacognition (P_{exp}). Although the correlation between P_{exp} in perceptual and memory tasks ($r = .37$) was numerically higher than the correlation between meta-*d'*/*d'* across domains ($r = .27$), this difference was not statistically detectable, $Z = 0.69$, $p = .490$.

Discussion

Although Experiment 1 revealed that the correlation between meta-*d'*/*d'* across perceptual and memory domains was significantly higher than zero based on single-participant estimation, the correlation coefficient from the hierarchical Bayesian model did not reach statistical significance. According to Fleming (2017), the hierarchical Bayesian model is able to take into account both within- and between-participant uncertainty in model fitting, and improve the fit of group-level parameters such as cross-domain correlation. However, single-participant estimation with 200 trials for each participant is also accurate (Fleming, 2017). Future studies should further investigate the relationship between the estimated value of meta-*d'*/*d'* based on single-participant estimation and hierarchical Bayesian model. Overall, the results in Experiment 1 suggested that there might exist weak domain generality of metacognitive resolution, which is consistent with some of previous studies (Mazancieux et al., 2020; Rouault et al., 2018). According to Mazancieux et al. (2020), a larger sample size is needed to reliably obtain the domain generality of meta-*d'*/*d'*. But even with data from 181 participants, Mazancieux et al. (2020) did not find statistically significant correlation between meta-*d'*/*d'* in perceptual task and episodic memory task (which was also used in the current study), suggesting the relationship between metacognitive resolution in these two domains might be difficult to detect.

On the other hand, our results showed reliably positive correlation between the contribution of processing experience to confidence ratings (P_{exp}) across perceptual and memory tasks, suggesting that the cognitive process underlying the integration of processing experience and prior beliefs during metacognition might be domain-general. In both perceptual and memory tasks, people may have prior beliefs about their overall ability in the task, and also need to monitor the processing experience during the task when they are asked to rate their confidence (Koriat et al., 2004; Lee et al., 2021; Rouault et al., 2019; Yang et al., 2021). By fitting BIM to data from perceptual and memory tasks, we found the extent to which experience and beliefs contributed to confidence process seemed be correlated across domains.

Our hypothesis was that the domain generality of P_{exp} might be greater than that of meta-*d'*/*d'*, because the varied relationship between actual performance and the information (e.g., processing experience) utilized during confidence rating process across tasks might reduce the domain generality of metacognitive resolution. However, although the cross-domain correlation of P_{exp} was numerically higher than that of meta-*d'*/*d'*, this difference was not statistically detectable. One possibility is that the difference in domain generality between P_{exp} and meta-*d'*/*d'* might be small in Experiment 1, and there might be a lack of sufficient statistical power when comparing the cross-domain correlation of P_{exp} and meta-*d'*/*d'*.

Before drawing a firm conclusion about the difference in domain generality between P_{exp} and meta-*d'*/*d'*, we conducted Experiment 2 to conceptually replicate the results of Experiment 1 using perceptual and memory tasks with different task difficulty.

Experiment 2

Method

Participants

Participants in Experiment 2 were recruited online via the Prolific website. As in Experiment 1, we planned to collect data until the Bayes factor (BF_{10}) for the correlation of P_{exp} between perceptual and memory tasks was greater than 10 or lower than 1/10. However, the BF_{10} for the cross-domain correlation of P_{exp} was higher than 10 when there were only 26 participants. In order to obtain a more stable pattern of the domain generality of P_{exp} and meta-*d'*/*d'*, we decided to continue collecting data until 68 participants were recruited (39 female; age: M = 28.34 years, SD = 6.32), which was the same sample size as in Experiment 1. Data from another seven participants were excluded due to that their *d'* in memory task was equal to or smaller than 0. All participants spoke English as the first language and reported normal or corrected-to-normal vision. Participants signed informed consent, and received monetary compensation (£8.5) and a bonus (up to £1). All procedures were approved by the Ethics Committee at Beijing Normal University.

Materials and Procedure

As in Experiment 1, participants in Experiment 2 performed both a perceptual task and a memory task, and task order was randomly assigned for each participant. The perceptual task in Experiment 2 was a dot discrimination task, which is similar to the task in Mazancieux et al. (2020) (see Figure 2B). After presenting a fixation for 500 ms in each trial, two circles were shown on the screen for 700 ms. One of the circles always contained 50 dots, and the number of dots in the other circle varied between 51 and 75, randomly defined in each trial. Participants needed to decide which circle contained more dots, and then rated their confidence by pressing a number key from 1 to 6. There were 200 trials in the perceptual task.

The experimental materials for the memory task in Experiment 2 were the same as those in Experiment 1. However, a more difficult memory task was employed in Experiment 2, in which participants needed to first learn all of the 200 words, and then completed 200 trials in the 2AFC recognition memory test with confidence rating. During the learning phase, words were presented one by one for 1500 ms each, and the order of words was randomized across participants. A fixation was presented for 500 ms prior to the presentation of each word (see Figure 3B).

Data Analysis

Data analysis in Experiment 2 was the same as that in Experiment 1.

Results

Task performance and mean confidence

Task performance (*d'*) was greater in the perceptual task than in the memory task, $t(67) = 11.17$, $p < .001$, Cohen's $d = 1.35$, $BF_{10} > 100$ (see Figure 6A). As in Experiment 1, *d'* in the two tasks did not correlate with each other, $r = .01$, $p = .923$, $BF_{10} = 0.15$, suggesting that there was no domain generality of task performance.

Mean confidence was also higher in the perceptual task than in the memory task, $t(67) = 7.41$, $p < .001$, Cohen's $d = 0.90$, $BF_{10} > 100$ (see Figure 6B).

Furthermore, the cross-domain correlation of mean confidence reached statistical significance (although the Bayes factor was inconclusive), $r = .29$, $p = .015$, $BF_{10} = 2.74$.

Metacognitive resolution

Metacognitive resolution (meta-*d'*/*d'*) was greater in the memory task than that in the perceptual task, $t(67) = 2.65$, $p = .010$, Cohen's $d = 0.32$, $BF_{10} = 3.35$ (see Figure 6C). Furthermore, there was no reliable correlation between meta-*d'*/*d'* in the two tasks, $r = .11$, $p = .386$, $BF_{10} = 0.22$ (see Figure 7A). However, Figure 7A suggests that there were extreme values in the meta-*d'*/*d'* of the memory task. The occurrence of extreme values was due to very low performance in the memory task ($d' < 0.1$) for these participants. In order to reduce the effect of extreme values on the cross-domain correlation, we examined the non-parametric Kendall's τ_b correlation between meta-*d'*/*d'* in the perceptual and memory tasks, which was statistically detectable based on *p* value (although the Bayes factor was inconclusive), $\tau_b = .17$, $p = .039$, $BF_{10} = 1.28$.

We then estimated the cross-domain correlation of meta-*d'*/*d'* at group level based on the hierarchical Bayesian model in *HMeta-d* package. As in Experiment 1, results from the hierarchical Bayesian model revealed that the cross-domain correlation of meta-*d'*/*d'* did not reach significance, $\rho = .30$, 95% credible interval (CrI) [-.11, .69].

Contribution of processing experience to confidence ratings

As in Experiment 1, processing experience contributed more to confidence ratings (i.e., P_{exp} was higher) in the memory task than in the perceptual task, $t(67) = 6.21$, $p < .001$, Cohen's $d = 0.75$, $BF_{10} > 100$ (see Figure 6D). Furthermore, the correlation between P_{exp} in the perceptual and memory tasks was reliable, $r = .64$, $p < .001$, $BF_{10} > 100$ (see Figure 7B), supporting domain generality of the contribution of processing experience to confidence ratings.

Difference in domain generality between P_{exp} and meta-d'/d'

Finally, we quantitatively compared the domain generality of meta-*d'*/*d'* and that of P_{exp}. The result revealed that the cross-domain correlation of P_{exp} ($r = .64$) was higher than that of meta-*d'*/*d'* ($r = .11$), and this difference was statistically detectable, $Z = 3.87$, $p < .001$. However, we should note that the Pearson correlation between meta-*d'*/*d'* in perceptual and memory tasks might be affected by the extreme values in the meta-*d'*/*d'* from the memory task, which might exaggerate the difference between the cross-domain correlation of P_{exp} and meta-*d'*/*d'*. Thus, we then compared the cross-domain Kendall's τ_b correlation of P_{exp} and meta-*d'*/*d'*, which also indicated that the cross-domain correlation of P_{exp} was reliably higher, $Z = 3.33$, $p < .001$.[1]

Discussion

The task difficulty in Experiment 2 was different from that in Experiment 1: participants performed better in the perceptual task than memory task in Experiment

[1] We first transformed Kendall's τ_b into Pearson's r according to Gilpin (1993), and then compared the transformed cross-domain Pearson's r of P_{exp} and meta-*d'*/*d'* using the method provided by Dunn and Clark (1969).

2, while task performance in Experiment 1 was greater in the memory task than in the perceptual task. In addition, results from the comparison between task performance in the two experiments revealed that the perceptual task was easier while the memory task was more difficult in Experiment 2 than those in Experiment 1 (ts > 4.96, ps < .001, BF_{10} > 100). Regardless of these changes in task difficulty among experiments, most of the main results in Experiment 1 were successfully replicated in Experiment 2. First, although the single-participant estimation of meta-*d'*/*d'* was correlated between perceptual and memory tasks (based on Kendall's τ_b correlation), the correlation coefficient from the hierarchical Bayesian model did not reach statistical significance. These results further suggest there might exist weak domain generality of metacognitive resolution. Second, a reliably positive correlation between P_{exp} across domains was obtained, supporting the domain generality of the cognitive process underlying the integration of processing experience and prior beliefs during metacognition across various levels of task difficulty.

Unlike Experiment 1, the cross-domain correlation of P_{exp} was significantly greater than that of meta-*d'*/*d'* in Experiment 2. One possibility is that difference between the domain generality of P_{exp} and meta-*d'*/*d'* may be related to the particular task type, which varied between the two experiments. More data from various types of perceptual and memory tasks are needed to investigate whether the overall difference between the cross-domain correlation of P_{exp} and meta-*d'*/*d'* was reliable.

Before further comparing the domain generality of P_{exp} and meta-*d'*/*d'*, we conducted Experiment 3 in which participants were asked to perform two different perceptual tasks with confidence ratings, including the color discrimination task in Experiment 1 and the dot discrimination task in Experiment 2. The aim of Experiment 3 was to directly examine the effect of specific task requirements on meta-*d'*/*d'*. Because both tasks in Experiment 3 were within the same perceptual domain, it is reasonable to expect that the cross-task correlation of meta-*d'*/*d'* in Experiment 3 should be significantly higher than that across perceptual and memory tasks in Experiments 1 and 2, as shown in previous study (Rouault et al., 2018). However, we should also note that the specific task requirements in the two perceptual tasks are quite different, and the information (or cues) utilized during confidence rating process might predict actual task performance to different extent in the two tasks (Koriat, 1997), which could lead to variation in metacognitive resolution and reduce its generality across tasks. If meta-*d'*/*d'* was affected by specific task requirements, then it is also possible that the correlation between meta-*d'*/*d'* in two perceptual tasks might be low (Arbuzova et al., 2021; Samaha & Postle, 2017).

Experiment 3

Method

Participants

Participants in Experiment 3 were recruited online via the Prolific website. We collected data from 68 participants (34 female; age: M = 29.68 years, SD = 6.27), which was the same sample size as in Experiments 1 and 2. Data from another one participant were excluded due to constant confidence rating for all trials in the dot discrimination task. All participants spoke English as the first language and reported normal or corrected-to-normal vision. Participants signed informed consent, and received monetary compensation (£7.5) and a bonus (up to £1). All procedures were approved by the Ethics Committee at Beijing Normal University.

Materials and Procedure

Participants in Experiment 3 performed two perceptual tasks, including the color discrimination task used in Experiment 1 and the dot discrimination task in Experiment 2. The task order was randomly assigned for each participant.

Data Analysis

Data analysis in Experiment 3 was similar to that in Experiments 1 and 2 with the following exceptions. First, we analyzed the results from two different perceptual tasks in Experiment 3, rather than one perceptual and one memory task (as in the first two experiments). Second, we compared the cross-task correlation of meta-*d'*/*d'* in Experiment 3 and that in Experiments 1 and 2. Similarly, we examined the difference between the cross-task correlation of P_{exp} in Experiment 3 and that in Experiments 1 and 2. The Z statistic for the difference between the correlation coefficients in two independent experiments is calculated as (Fisher, 1925):

$$Z = \frac{z_1 - z_2}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}$$

In the equation above, z_1 and z_2 are the z-transformed correlation in the two experiments. The N_1 and N_2 represent the sample sizes in each experiment.

Results

Task performance and mean confidence

Task performance (*d'*) was greater in the dot discrimination task than in the color discrimination task, $t(67) = 22.62$, $p < .001$, Cohen's $d = 2.74$, $BF_{10} > 100$ (see Figure 8A). In addition, *d'* in the two tasks was correlated with each other (although the Bayes factor was inconclusive), $r = .29$, $p = .017$, $BF_{10} = 2.43$.

Mean confidence was also higher in the dot discrimination task than in the color discrimination task, $t(67) = 8.84$, $p < .001$, Cohen's $d = 1.07$, $BF_{10} > 100$ (see Figure 8B). Furthermore, there was a reliable correlation between the mean confidence in the two tasks, $r = .73$, $p < .001$, $BF_{10} > 100$.

Metacognitive resolution

Metacognitive resolution (meta-*d'*/*d'*) did not differ between color discrimination and dot discrimination tasks, $t(67) = 1.21$, $p = .232$, Cohen's $d = 0.15$, $BF_{10} = 0.27$ (see Figure 8C). Furthermore, there was no reliable correlation between meta-*d'*/*d'* in the two tasks, $r = .14$, $p = .243$, $BF_{10} = 0.30$ (see Figure 9A). We also computed the non-parametric Kendall's τ_b correlation between meta-*d'*/*d'* in the two tasks, and found a statistically non-significant correlation, $\tau_b = .07$, $p = .380$, $BF_{10} = 0.23$.[2] As in Experiments 1 and 2, we then estimated the cross-task correlation of meta-*d'*/*d'* at group level based on the hierarchical Bayesian model in *HMeta-d*

[2] For completeness, we computed the cross-domain Kendall's τ_b correlation of meta-*d'*/*d'* in Experiment 1, which did not reach statistical significance, $\tau_b = .14$, $p = .103$, $BF_{10} = 0.58$. In addition, the cross-task Kendall's τ_b correlation of P_{exp} was statistically detectable in all of the three experiments, τ_bs $> .22$, *p*s $< .01$, BF_{10}s > 5.

package, which revealed that the cross-task correlation of meta-*d'*/*d'* did not reach significance, $\rho = .02$, 95% credible interval (CrI) [-.47, .52].

Next, we compared the cross-task correlation of meta-*d'*/*d'* in Experiment 3 with that in Experiments 1 and 2. Results indicated that the correlation between meta-*d'*/*d'* in the two perceptual tasks in Experiment 3 ($r = .14$) did not reliably differ from that across the perceptual and memory tasks in Experiment 1 ($r = .27$), $Z = 0.72$, $p = .471$. Similarly, the cross-task correlation of meta-*d'*/*d'* did not differ between Experiments 2 ($r = .11$) and 3, $Z = 0.21$, $p = .830$.

Contribution of processing experience to confidence ratings

Processing experience contributed more to confidence ratings (i.e., P_{exp} was higher) in the dot discrimination task than in the color discrimination task (although the Bayes factor was inconclusive), $t(67) = 2.60$, $p = .011$, Cohen's $d = 0.32$, $BF_{10} = 2.98$ (see Figure 8D). Furthermore, the correlation between P_{exp} in the two tasks was reliable, $r = .48$, $p < .001$, $BF_{10} > 100$ (see Figure 9B).

We then compared the cross-task correlation of P_{exp} in Experiment 3 and that in Experiments 1 and 2. Results revealed that the correlation between P_{exp} in the two perceptual tasks in Experiment 3 ($r = .48$) did not reliably differ from that across the perceptual and memory tasks in Experiment 1 ($r = .37$), $Z = 0.79$, $p = .432$, or from the cross-task correlation of P_{exp} in Experiment 2 ($r = .64$), $Z = 1.37$, $p = .169$.

Difference in task generality between P_{exp} and meta-d'/d'

Finally, we quantitatively compared the cross-task correlation of meta-*d'*/*d'* and that of P_{exp}. The result revealed that the cross-task correlation of P_{exp} (r = .48) was higher than that of meta-*d'*/*d'* (r = .14), and this difference was statistically detectable, Z = 2.31, p = .021.

Discussion

Experiment 3 indicated that the metacognitive resolution (represented by meta-*d'*/*d'*) was not reliably correlated between color discrimination and dot discrimination tasks, although there might exist weak correlation between meta-*d'*/*d'* in memory task and each of the two perceptual tasks (as shown in Experiments 1 and 2). Interestingly, meta-*d'*/*d'* in the two perceptual tasks did not correlate with each other even when the performance (*d'*) was correlated across tasks. Further analysis revealed that the cross-task correlation of meta-*d'*/*d'* in Experiment 3 did not significantly differ from that in Experiments 1 and 2.

In our view, these results do not suggest that the generality of cognitive mechanisms underlying metacognition across perceptual and memory domains is similar to (or even higher than) that within the perceptual domain, which is theoretically unlikely. Instead, the very low correlation between meta-*d'*/*d'* in the two perceptual tasks might result from the different task requirements (comparing color area or dot number in two visual stimuli). Participants might utilize different kinds of cues during confidence rating formation when the task requirements varied, which could predict actual ask performance to different extent (Koriat, 1997). As stated in Arbuzova et al. (2021), specific first-order task demands may play a crucial role in our ability to distinguish correct and incorrect response in a task during metacognition process, and should be considered when we try to make inference about domain generality of metacognition based on metacognitive resolution. Thus, we should not conclude whether the cognitive process underlying metacognition is general or specific simply on the basis of the cross-domain correlation of metacognitive resolution.

On the other hand, results from Experiments 1-3 revealed that there was a reliably positive cross-task correlation of P_{exp} both between perceptual and memory domains, and within the perceptual domain. Thus, the cognitive mechanisms underlying the integration of processing experience and prior beliefs during confidence rating process should be general across tasks and domains. However, the current study only used two different perceptual tasks and one type of memory task (single-word recognition), and whether similar domain generality of P_{exp} could be obtained in other types of perceptual and memory tasks remains unknown. In addition, more data are needed to investigate whether the domain generality of P_{exp} was reliably greater than that of meta-*d'*/*d'*. Thus, a series of meta-analyses was performed to examine the domain generality of P_{exp} across different kinds of perceptual and memory tasks, and further explore the potential difference in cross-domain correlation between P_{exp} and meta-*d'*/*d'*.

Meta-analysis

Method

The data and scripts for the meta-analyses can be accessed at the OSF project page: https://osf.io/92qjr/ (Hu, 2022).

Experiments

In order to further investigate the domain generality of the contribution of processing experience (and prior beliefs) to confidence ratings (i.e., P_{exp}), and to examine the difference in domain generality between P_{exp} and metacognitive resolution (meta-*d'*/*d'*), a series of meta-analyses was conducted to integrate data from the Confidence Database to increase statistical power. The Confidence Database is a large database containing data of task performance and confidence ratings from published and unpublished studies (for details, see Rahnev et al., 2020). To our knowledge, there are only six experiments in the Confidence Database which asked the same group of participants to perform both perceptual and memory tasks, and give item-by-item confidence ratings (Mazancieux et al., 2018; Sadeghi et al., 2017; Samaha et al., 2016; Samaha & Postle, 2017; Schmidt et al., 2019; Ye et al., 2018). Thus, here we performed a meta-analysis on data from these six experiments. In addition, we performed another meta-analysis which included the data from our Experiments 1 and 2.

Table 1 shows the experimental paradigm for the six experiments, in which different types of perceptual and memory tasks were used.[3] Furthermore, while some of the experiments asked participants to complete only one perceptual and one

[3] Mazancieux et al. (2018) is the preprint version of Mazancieux et al. (2020).

memory task, other experiments contained multiple perceptual or memory tasks.

Data Analysis

All meta-analyses were performed using random-effects models via the *metafor* package in R (Viechtbauer, 2010). Data from seven participants in Mazancieux et al. (2018) were excluded, because one participant gave the constant confidence rating across all trials which caused problem in the estimation of P_{exp} (Hu et al., 2021), and the other six participants had *d'* equal to or lower than 0. In addition, data from one participant in Samaha et al. (2016) were excluded due to *d'* lower than 0, and one participant from Schmidt et al. (2019) was excluded due to missing data. Furthermore, the continuous confidence scale used in Schmidt et al. (2019) was converted to a 7-point scale with equal length in each confidence bin in order to compute meta-*d'*/*d'*. Here we divided continuous confidence ratings into 7 bins to make sure no confidence value lay just at the edge of bins (Hu et al., 2021).

We first fit BIM separately to data from each perceptual and memory task in each experiment, and performed meta-analysis on the correlation between the estimated value of P_{exp} in BIM across domains. We calculated the correlation coefficient between P_{exp} in perceptual and memory tasks for each experiment, and then transformed the correlation coefficients into Fisher's *z* values which were used in subsequent meta-analysis. For the experiments containing multiple perceptual or memory tasks, we calculated the correlation between P_{exp} in each perceptual and each memory task, and then averaged the Fisher's *z* values from different tasks (see Section

S2 in Supplemental Materials for calculation of the standard error for averaged Fisher's *z* values). In addition, some experiments contained several within-participant conditions. For example, Schmidt et al. (2019) measured task performance and confidence ratings before and after meditation training, and Ye et al. (2018) applied transcranial magnetic stimulation over different brain areas. In this case, we separately calculated the Fisher's *z* value for each experimental condition, and then averaged the *z* values across conditions.

Next, we conducted meta-analysis on the cross-domain correlation of meta-*d'*/*d'*. For each experiment, we used *HMeta-d* package to obtain single-participant estimation of the meta-*d'*/*d'* in each task for each participant (as in Experiments 1-3), and calculated the correlation between meta-*d'*/*d'* across perceptual and memory tasks, which was then transformed into Fisher's *z* values. Finally, we calculated the difference between the *z*-transformed cross-domain correlation of P_{exp} and meta-*d'*/*d'* in each experiment, and performed meta-analysis on this difference.

In all of the meta-analyses reported below, effect sizes (i.e., *z*-transformed correlation coefficients) across experiments were weighted by the inverse of their variance, and the variance of true effect sizes (i.e., τ^2) was estimated using restricted maximum likelihood estimation. Furthermore, we assessed heterogeneity amongst effect sizes across experiments using the *Q* statistic, and performed PET-PEESE analysis to examine publication bias (Stanley & Doucouliagos, 2014).

Results and Discussion

Domain generality of P_{exp}

Results from the meta-analysis revealed that the overall correlation between P_{exp} across perceptual and memory tasks was significant across the six experiments, $r = .45$, 95% confidence interval (CI) [.36, .54] (see Figure 10 for the *z*-transformed correlation), which was consistent with the results found in Experiments 1 and 2. Including the data from Experiments 1 and 2 into the meta-analysis resulted into the same pattern, $r = .48$, 95% CI [.38, .57]. These results further support our hypothesis that the contribution of processing experience and prior beliefs to confidence ratings is domain-general, suggesting the cognitive process underlying the integration of experience and beliefs during metacognition is similar across different cognitive domains.

Domain generality of meta-d'/d'

When examining the cross-domain correlation of meta-*d'*/*d'*, the meta-analytic results showed that the overall correlation between meta-*d'*/*d'* in perceptual and memory tasks from the six experiments was just above the significance level, $r = .13$, 95% CI [.01, .25] (see Figure 11 for the *z*-transformed correlation). Adding the data from Experiments 1 and 2 into the meta-analysis resulted into the same pattern, $r = .15$, 95% CI [.07, .24]. These results suggest a significant but weak domain generality of metacognitive resolution across perceptual and memory domains.

Difference between domain generality of P_{exp} and meta-d'/d'

Finally, we found the *z*-transformed cross-domain correlation of P_{exp} was significantly greater than that of meta-*d'*/*d'* in the meta-analysis on the six experiments, $z = 0.33$, 95% CI [.17, .47] (see Figure 12). Including the data from Experiments 1 and 2 into the meta-analysis did not change the result pattern, $z = 0.35$, 95% CI [.19, .49]. These results indicate that the proportion for the contribution of processing experience (and prior beliefs) to metacognition was more domain-general than metacognitive resolution.

The heterogeneity amongst the effect sizes across experiments was not statistically detectable in any of the meta-analyses reported above, *Q*s < 12.11, *p*s > .09. Furthermore, results from the PET-PEESE analysis revealed little need to worry about publication bias in all meta-analyses, *p*s > .32.

General Discussion

Previous studies employed metacognitive resolution to investigate domain generality of metacognition, and they showed inconsistent results about whether metacognition is domain-general or domain-specific (e.g., Baird et al., 2013, 2015; McCurdy et al., 2013; Samaha & Postle, 2017). The current study examined metacognitive domain generality using the Bayesian inference model for metamemory (BIM), which suggests when people rate their confidence about task performance, they integrate processing experience obtained from the task and their prior beliefs about overall ability through a Bayesian inference (Hu et al., 2021). In the current study, we applied BIM to perceptual metacognition, and examined whether the contribution of processing experience and prior beliefs to metacognition is general or specific across perceptual and memory domains.

Results from Experiments 1 and 2 indicated that the contribution of processing experience to confidence ratings (parameter P_{exp} in BIM) was reliably correlated across perceptual and memory domains. Furthermore, there might exist a weak cross-domain correlation of metacognitive resolution (meta-*d'*/*d'*). We then conducted Experiment 3 which surprisingly showed a non-significant correlation between meta-*d'*/*d'* in two perceptual tasks, suggesting metacognitive resolution might be affected by specific task requirements. Finally, a series of meta-analyses, which was performed on data from six published experiments, validated the significant correlation between P_{exp} in perceptual and memory tasks, and revealed greater domain generality of P_{exp} than that of meta-*d'*/*d'*.

To our knowledge, the current study is the first to extend BIM to perceptual metacognition, and suggests that perceptual metacognition is based on the integration of processing experience and prior beliefs. Previous studies on perceptual metacognition mainly focused on the effect of processing experience (such as stimulus evidence and response time for decision making) on confidence ratings (e.g., van den Berg et al., 2016; Zylberberg et al., 2012), and the investigation of global beliefs in perceptual metacognition literature has just begun (Fleming & Daw, 2017; Lee et al., 2021; Rouault et al., 2019; Seow et al., 2021). For example, the second-order model proposed by Fleming and Daw (2017) suggests that confidence ratings in perceptual tasks may be systematically affected by both perceptual evidence and beliefs about perceptual process. Furthermore, Rouault et al. (2019) indicates that during perceptual tasks people could build global beliefs about overall performance, which could affect subsequent decision making process. The current study supports the second-order model and further quantitatively suggests that both processing experience and prior beliefs contribute to perceptual metacognition.

Furthermore, our results revealed significant correlation between P_{exp} across perceptual and memory domains. Although the integration of processing experience and prior beliefs during metamemory has been extensively discussed (for reviews, see Bjork et al., 2013; Yang et al., 2021), few have examined whether similar cognitive mechanisms could account for metacognition in other domains. The findings documented here indicate that the extent to which experience and beliefs contribute to confidence ratings showed consistent tendency across perceptual and memory tasks, suggesting the cognitive process underlying the integration of experience and beliefs during metacognition is domain-general.

In the current study, we detected a weak correlation between meta-*d'*/*d'* across perceptual and memory domains both in Experiments 1-2 and the meta-analysis on data from six published experiments. One possible explanation for the weak cross-domain correlation of metacognitive resolution is that metacognitive resolution relates people's confidence ratings and task performance, and thus the domain generality of metacognitive resolution is inevitably affected by the specific design and requirements of the tasks across domains. During metacognition in perceptual and memory tasks, participants could utilize different types of information, which might predict the actual task performance to different extent and lead to weak domain generality of metacognitive resolution. For example, participants could integrate different cues about stimulus when evaluating their memory or perceptual performance (e.g., frequency and familiarity of words, and spatial information in images), and the degree to which each cue affected the accuracy of confidence ratings might vary across participants and tasks (Koriat, 1997).

To further illustrate the effect of task demands on metacognitive resolution, we conducted Experiment 3 which revealed a statistically non-detectable correlation between meta-*d'*/*d'* in two perceptual tasks. This result suggests that there may not be task generality of metacognitive resolution even within the same cognitive domain when the specific requirements of the two tasks are quite different. Thus, the cross-domain correlation of metacognitive resolution should be carefully interpreted when the task requirements vary across cognitive domains, and simply concluding whether there exists metacognitive domain generality based on the cross-domain correlation of metacognitive resolution (as did in many previous studies) is inappropriate.

The reported meta-analyses also showed that the cross-domain correlation of P_{exp} was reliably higher than that of meta-*d'*/*d'*. Contrary to meta-*d'*/*d'*, P_{exp} solely reflects the confidence process itself, and indicates how processing experience and prior beliefs are integrated during confidence rating formation (Hu et al., 2021). Thus, the extent of the domain generality of P_{exp} may better reflect whether the cognitive mechanisms underlying confidence rating process are domain-general or domain-specific. Our results suggest that the cognitive process underlying the utilization of processing experience and prior beliefs during metacognition is general across different cognitive domains and task requirements. Thus, when investigating metacognitive domain generality, we should try to eliminate the confounds introduced by the varied relationship between confidence ratings and actual performance across tasks, and focus on the confidence rating process itself.

Although the current study observed a positive cross-domain correlation of P_{exp}, it did not investigate why the contribution of processing experience to confidence ratings was correlated across perceptual and memory domains. One possibility lies in the utilization of response time in the tasks as a cue during confidence rating. Previous studies suggest that response time is related to processing fluency in the task, which is one of the most important processing experiences considered in metacognitive process (e.g., Koriat et al., 2008; Koriat & Ackerman, 2010). Thus, it is possible that the extent of the reliance of confidence ratings on response time could be correlated across cognitive domains, which might then lead to domain generality of P_{exp}. We performed additional meta-analyses to explore this possibility, and the results showed that P_{exp} was reliably related to the effect of response time on confidence, which was indeed correlated between perceptual and memory domains (see Section S3 in Supplemental Materials). Future studies should further investigate the relationship between individual differences in P_{exp} and the utilization of response time during confidence formation.

Readers need to be cautious about the weak domain generality of metacognitive resolution found in this study, because here we used the single-participant estimation of meta-*d'*/*d'* from the *HMeta-d* package to represent the metacognitive resolution for each participant (Fleming, 2017). Although single-participant estimation is more suitable for meta-analysis and can be used to compute the difference between cross-domain correlation of P_{exp} and meta-*d'*/*d'*, hierarchical Bayesian model might produce more accurate estimation of meta-*d'*/*d'* by taking into account the uncertainty in parameter estimation at both individual and group levels (Fleming, 2017). Furthermore, results based on the single-participant and hierarchical estimation of meta-*d'*/*d'* were slightly different in Experiments 1 and 2: while there was a significant but weak correlation between single-participant estimation of meta-*d'*/*d'* in perceptual and memory tasks, hierarchical estimation of the cross-domain correlation of meta-*d'*/*d'* did not reach statistical significance. Future studies should try to compare the cross-domain correlation of P_{exp} and meta-*d'*/*d'* based on estimation from hierarchical Bayesian models.

The current study only investigated metacognition across perceptual and memory domains. In fact, people can evaluate their performance in tasks from other cognitive domains such as nociception, voluntary movement, reasoning, and so on (Arbuzova et al., 2021; Beck et al., 2019; Double & Birney, 2019). Future studies should examine the correlation of P_{exp} across these cognitive domains, and further quantitatively compare the extent of generality of metacognitive resolution and that of P_{exp} in these domains.

Future studies are also encouraged to use P_{exp} to examine the generality of metacognitive process across sensory modalities (e.g., vision and audition) within the perceptual domain. Previous research findings indicate that compared with the cross-domain correlation of metacognitive resolution, the correlation between metacognitive resolution in different sensory modalities is typically stronger (Rouault et al., 2018). However, some studies also reported inconsistent results about whether there is significant cross-modality correlation of metacognitive resolution (e.g., Ais et al., 2016; Faivre et al., 2018), which might be due to the variation of task requirements (as shown in our Experiment 3). In contrast, the cross-modality correlation of P_{exp} may better reflect the generality of cognitive mechanisms underlying confidence rating process itself. Based on the results from the current study, it is reasonable to speculate that the contribution of processing experience and prior beliefs to confidence should be general across sensory modalities. When examining the generality of metacognitive process both between and within domains, future studies should use discrete confidence scale with a relatively large number of points (e.g., 6 points or more; Rahnev et al., 2020) or continuous confidence scale, because the estimation accuracy of both metacognitive resolution and P_{exp} is improved when the number of scale points increases (Higham & Higham, 2019; Hu et al., 2021).

Conclusion

In conclusion, the current study demonstrated that while the domain generality of metacognitive resolution is weak, the cognitive process underlying the integration of processing experience and prior beliefs during confidence rating formation is more general across perceptual and memory domains.

Context of the Research

This study was inspired by the Bayesian inference model for metamemory (BIM) proposed by Hu et al. (2021). We speculated that metacognitive resolution (e.g., meta-*d'*/*d'*), which has been widely used in previous studies, may not be a suitable measure for investigating domain generality of metacognition, because the relationship between confidence ratings and task performance is affected by specific task requirements (Arbuzova et al., 2021; Samaha & Postle, 2017). In contrast, the parameter P_{exp} in BIM only reflects the contribution of processing experience and prior beliefs to confidence ratings, and is unrelated to task performance (Hu et al., 2021). In fact, this study is just an initial attempt to reduce the confounds introduced by the varied relationship between confidence and performance across tasks when examining metacognitive domain generality. We hope future studies can develop other measures that solely reflect the cognitive mechanisms underlying confidence rating process itself, which can then be utilized to investigate domain generality of metacognition.

References

Ais, J., Zylberberg, A., Barttfeld, P., & Sigman, M. (2016). Individual consistency in the accuracy and distribution of confidence judgments. *Cognition, 146*, 377–386. https://doi.org/10.1016/J.COGNITION.2015.10.006

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Arbuzova, P., Peters, C., Röd, L., Koß, C., Maurer, H., Maurer, L. K., Müller, H., Verrel, J., & Filevich, E. (2021). Measuring metacognition of direct and indirect parameters of voluntary movement. *Journal of Experimental Psychology: General*, Advance online publication. https://doi.org/10.1037/xge0000892

Baird, B., Cieslak, M., Smallwood, J., Grafton, S. T., & Schooler, J. W. (2015). Regional white matter variation associated with domain-specific metacognitive accuracy. *Journal of Cognitive Neuroscience*, *27*(3), 440–452. https://doi.org/10.1162/jocn_a_00741

Baird, B., Smallwood, J., Gorgolewski, K. J., & Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *Journal of Neuroscience*, *33*(42), 16657–16665. https://doi.org/10.1523/JNEUROSCI.0786-13.2013

Beck, B., Peña-Vivas, V., Fleming, S., & Haggard, P. (2019). Metacognition across sensory modalities: Vision, warmth, and nociceptive pain. *Cognition*, *186*, 32–

41. https://doi.org/10.1016/j.cognition.2019.01.018

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

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

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497–505. https://doi.org/10.1080/14640748108400805

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/tqmp.01.1.p042

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective Confidence Predicts Information Seeking in Decision Making. *Psychological Science*, 095679761774477. https://doi.org/10.1177/0956797617744771

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology*, *72*, 78–89. https://doi.org/10.1016/j.jmp.2015.10.003

Double, K. S., & Birney, D. P. (2019). Do confidence ratings prime confidence?

Psychonomic Bulletin and Review, *26*, 1035–1042. https://doi.org/10.3758/s13423-018-1553-3

Dunn, O. J., & Clark, V. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association*, *64*(325), 366–377. https://doi.org/10.1080/01621459.1969.10500981

Faivre, N., Filevich, E., Solovey, G., Kühn, S., & Blanke, O. (2018). Behavioral, modeling, and electrophysiological evidence for supramodality in human metacognition. *Journal of Neuroscience*, *38*(2), 263–277. https://doi.org/10.1523/JNEUROSCI.0322-17.2017

Fisher, R. A. (1925). *Statistical methods for research workers*. Oliver and Boyd.

Fitzgerald, L. M., Arvaneh, M., & Dockree, P. M. (2017). Domain-specific and domain-general processes underlying metacognitive judgments. *Consciousness and Cognition*, *49*, 264–277. https://doi.org/10.1016/j.concog.2017.01.011

Fleming, S. M. (2017). HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness*, *2017*(1), nix007. https://doi.org/10.1093/nc/nix007

Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general bayesian framework for metacognitive computation. *Psychological Review*, *124*(1), 91–114. https://doi.org/10.1037/rev0000045

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability.

Philosophical Transactions of the Royal Society B: Biological Sciences, *367*(1594), 1338–1349. https://doi.org/10.1098/rstb.2011.0417

Fleming, S. M., Ryu, J., Golfinos, J. G., & Blackmon, K. E. (2014). Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain*, *137*(10), 2811–2822. https://doi.org/10.1093/brain/awu221

Fowler, J. W. (2011). *Density Function for the Sum of Correlated Random Variables*. http://web.ipac.caltech.edu/staff/fmasci/home/astro_refs/SumOfCorrelatedRVs.pdf

Gilpin, A. R. (1993). Table for conversion of Kendall's tau to Spearman's rho within the context of measures of magnitude of effect for meta-analysis: *Educational and Psychological Measurement*, *53*(1), 87–92. https://doi.org/10.1177/0013164493053001007

Higham, P. A., & Higham, D. P. (2019). New improved gamma: Enhancing the accuracy of Goodman–Kruskal's gamma using ROC curves. *Behavior Research Methods*, *51*, 108–125. https://doi.org/10.3758/s13428-018-1125-5

Hu, X. (2022). *Domain generality of Pexp*. https://osf.io/92qjr/

Hu, X., Zheng, J., Su, N., Fan, T., Yang, C., Yin, Y., Fleming, S. M., & Luo, L. (2021). A Bayesian inference model for metamemory. *Psychological Review*, *128*(5), 824–855. https://doi.org/10.1037/rev0000270

Jang, Y., Wallsten, T. S., & Huber, D. E. (2012). A stochastic detection and retrieval model for the study of metacognition. *Psychological Review*, *119*(1), 186–200. https://doi.org/10.1037/a0025960

Jeffreys, H. (1961). *Theory of probability, (3rd ed.)*. Oxford University Press.

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., & Ackerman, R. (2010). Choice latency as a cue for children's subjective confidence in the correctness of their answers. *Developmental Science*, *13*(3), 441–453. https://doi.org/10.1111/J.1467-7687.2009.00907.X

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, *133*(4), 643–656. https://doi.org/10.1037/0096-3445.133.4.643

Koriat, A., Nussinson, R., Bless, H., & Shaked, N. (2008). Information-based and experience-based metacognitive judgments: Evidence from subjective confidence. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of Metamemory and Memory* (pp. 117–135). https://doi.org/10.4324/9780203805503

Lee, A. L. F., de Gardelle, V., & Mamassian, P. (2021). Global visual confidence. *Psychonomic Bulletin and Review*, *28*, 1233–1242. https://doi.org/10.3758/s13423-020-01869-7

Lee, A. L. F., Ruby, E., Giles, N., & Lau, H. (2018). Cross-domain association in metacognitive efficiency depends on first-order task types. *Frontiers in Psychology, 9*, 2464. https://doi.org/10.3389/fpsyg.2018.02464

Lei, W., Chen, J., Yang, C., Guo, Y., Feng, P., Feng, T., & Li, H. (2020). Metacognition-related regions modulate the reactivity effect of confidence ratings on perceptual decision-making. *Neuropsychologia*, *144*, 107502. https://doi.org/10.1016/j.neuropsychologia.2020.107502

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430. https://doi.org/10.1016/j.concog.2011.09.021

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

Mazancieux, A., Fleming, S., Souchay, C., & Moulin, C. (2018). *Retrospective confidence judgments across tasks: domain-general processes underlying metacognitive accuracy*. https://doi.org/10.31234/osf.io/dr7ba

McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P., & Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *Journal of Neuroscience*, *33*(5), 1897–1906. https://doi.org/10.1523/JNEUROSCI.1890-12.2013

Metcalfe, J., & Kornell, N. (2005). A Region of Proximal Learning model of study time allocation. *Journal of Memory and Language*, *52*(4), 463–477. https://doi.org/10.1016/j.jml.2004.12.001

Metcalfe, J., & Shimamura, A. P. (1994). *Metacognition: knowing about knowing*. MIT press.

Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-General and Domain-Specific Patterns of Activity Supporting Metacognition in Human Prefrontal Cortex. *Journal of Neuroscience*, *38*(14), 3534–3546. https://doi.org/10.1523/jneurosci.2360-17.2018

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In H. B. Gordon (Ed.), *Psychology of Learning and Motivation: Vol. Volume 26* (pp. 125–173). Academic Press. https://doi.org/10.1016/S0079-7421(08)60053-5

Palmer, E. C., David, A. S., & Fleming, S. M. (2014). Effects of age on metacognitive efficiency. *Consciousness and Cognition*, *28*, 151–160. https://doi.org/10.1016/j.concog.2014.06.007

Pescetelli, N., Hauperich, A.-K., & Yeung, N. (2021). Confidence, advice seeking and changes of mind in decision making. *Cognition*, *215,* 104810. https://doi.org/10.1016/j.cognition.2021.104810

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, *117*(3), 864–901. https://doi.org/10.1037/a0019737

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, 20–22. https://doi.org/10.1.1.13.3406

Rahnev, D., Desender, K., Lee, A. L. F., Adler, W. T., Aguilar-Lleyda, D., Akdoğan, B., Arbuzova, P., Atlas, L. Y., Balcı, F., Bang, J. W., Bègue, I., Birney, D. P., Brady, T. F., Calder-Travis, J., Chetverikov, A., Clark, T. K., Davranche, K., Denison, R. N., Dildine, T. C., … Zylberberg, A. (2020). The Confidence database. *Nature Human Behaviour, 4*(3), 317–325. https://doi.org/10.1038/s41562-019-0813-1

Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature Communications, 10*, 1141. https://doi.org/10.1038/s41467-019-09075-3

Rouault, M., & Fleming, S. M. (2020). Formation of global self-beliefs in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 117*(44), 27268–27276. https://doi.org/10.1073/pnas.2003094117

Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human metacognition across domains: Insights from individual differences and neuroimaging. *Personality Neuroscience, 1*, e17. https://doi.org/10.1017/pen.2018.16

Sadeghi, S., Ekhtiari, H., Bahrami, B., & Ahmadabadi, M. N. (2017). Metacognitive deficiency in a perceptual but not a memory task in methadone maintenance patients. *Scientific Reports, 7*(1), 1–8. https://doi.org/10.1038/s41598-017-06707-w

Samaha, J., Barrett, J. J., Sheldon, A. D., LaRocque, J. J., & Postle, B. R. (2016). Dissociating perceptual confidence from discrimination accuracy reveals no influence of metacognitive awareness on working memory. *Frontiers in Psychology, 7*, 851. https://doi.org/10.3389/fpsyg.2016.00851

Samaha, J., & Postle, B. R. (2017). Correlated individual differences suggest a common mechanism underlying metacognition in visual perception and visual short-term memory. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1867), 20172035. https://doi.org/10.1098/rspb.2017.2035

Schmidt, C., Reyes, G., Barrientos, M., Langer, Á. I., & Sackur, J. (2019). Meditation focused on self-observation of the body impairs metacognitive efficiency. *Consciousness and Cognition*, *70*, 116–125. https://doi.org/10.1016/j.concog.2019.03.001

Seow, T. X. F., Rouault, M., Gillan, C. M., & Fleming, S. M. (2021). How local and global metacognition shape mental health. *Biological Psychiatry*, *90*(7), 436–446. https://doi.org/10.1016/J.BIOPSYCH.2021.05.013

Shekhar, M., & Rahnev, D. (2021). The nature of metacognitive inefficiency in perceptual decision making. *Psychological Review*, *128*(1), 45–70. https://doi.org/10.1037/rev0000249

Siedlecka, M., Skóra, Z., Paulewicz, B., Fijałkowska, S., Timmermans, B., & Wierzchoń, M. (2019). Responses improve the accuracy of confidence judgements in memory tasks. *Journal of Experimental Psychology: Learning Memory and Cognition*, *45*(4), 712–723. https://doi.org/10.1037/xlm0000608

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-Time allocation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *26*(1), 204–221. https://doi.org/10.1037/0278-7393.26.1.204

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78. https://doi.org/10.1002/JRSM.1095

Tullis, J. G., & Benjamin, A. S. (2012). Consequences of restudy choices in younger and older learners. *Psychonomic Bulletin & Review*, *19*(4), 743–749. https://doi.org/10.3758/s13423-012-0266-2

van den Berg, R., Anandalingam, K., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016). A common mechanism underlies changes of mind about decisions and confidence. *ELife*, *5*, e12192. https://doi.org/10.7554/eLife.12192

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Vuorre, M., & Metcalfe, J. (2021). Measures of relative metacognitive accuracy are confounded with task performance in tasks that permit guessing. *Metacognition and Learning*, Advance online publication. https://doi.org/10.1007/S11409-020-

09257-1

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

Ye, Q., Zou, F., Lau, H., Hu, Y., & Kwok, S. C. (2018). Causal evidence for mnemonic metacognition in human precuneus. *Journal of Neuroscience, 38*(28), 6379–6387. https://doi.org/10.1523/JNEUROSCI.0660-18.2018

Zylberberg, A., Barttfeld, P., & Sigman, M. (2012). The construction of confidence in a perceptual decision. *Frontiers in Integrative Neuroscience, 6*, 79. https://doi.org/10.3389/fnint.2012.00079

Table 1

Experimental paradigms for the six experiments included in the meta-analyses

Experiments	Sample size	Confidence scale	Perceptual task	Memory task
Mazancieux et al., 2018	181	11-point	Dot discrimination	Episodic memory task Semantic memory task Working memory task
Sadeghi et al., 2017	50	6-point	Contrast discrimination	Episodic memory task
Samaha et al., 2016	15	4-point	Orientation discrimination	Working memory task
Samaha & Postle, 2017, Experiment 3	20	4-point	Orientation discrimination Contrast discrimination	Working memory task
Schmidt et al., 2019	27	Continuous (1-6)	Contrast discrimination	Episodic memory task
Ye et al., 2018	18	4-point	Resolution discrimination	Episodic memory task

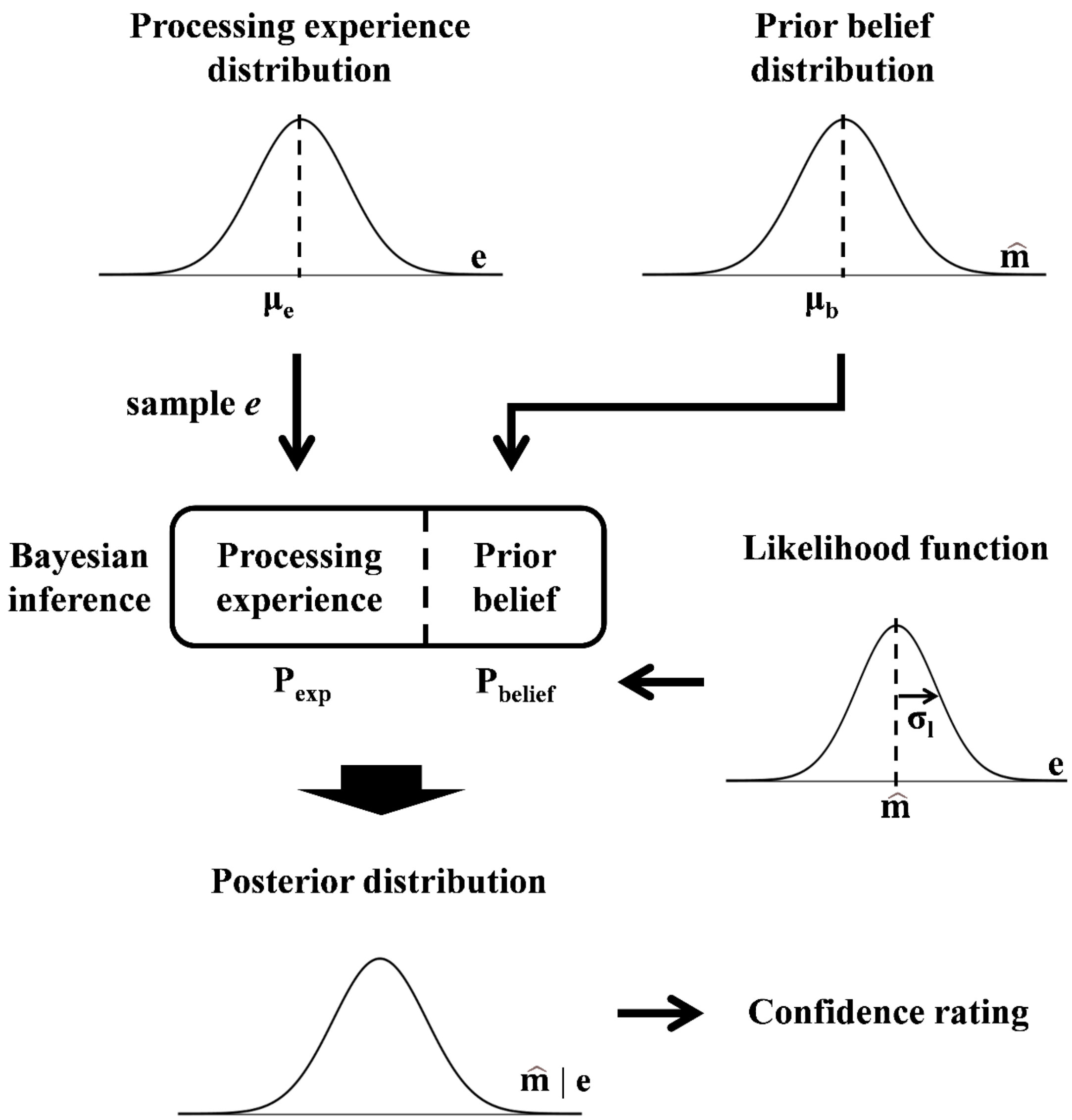

Figure 1. Schematic illustration of the Bayesian inference model for metamemory (BIM). During confidence rating process, people apply Bayesian inference to infer memory strength for each item based on both the processing experience and prior beliefs about memory ability. The contribution of processing experience and prior beliefs to confidence is determined by P_{exp} and P_{belief}, respectively, both of which are related to the standard deviation of likelihood function.

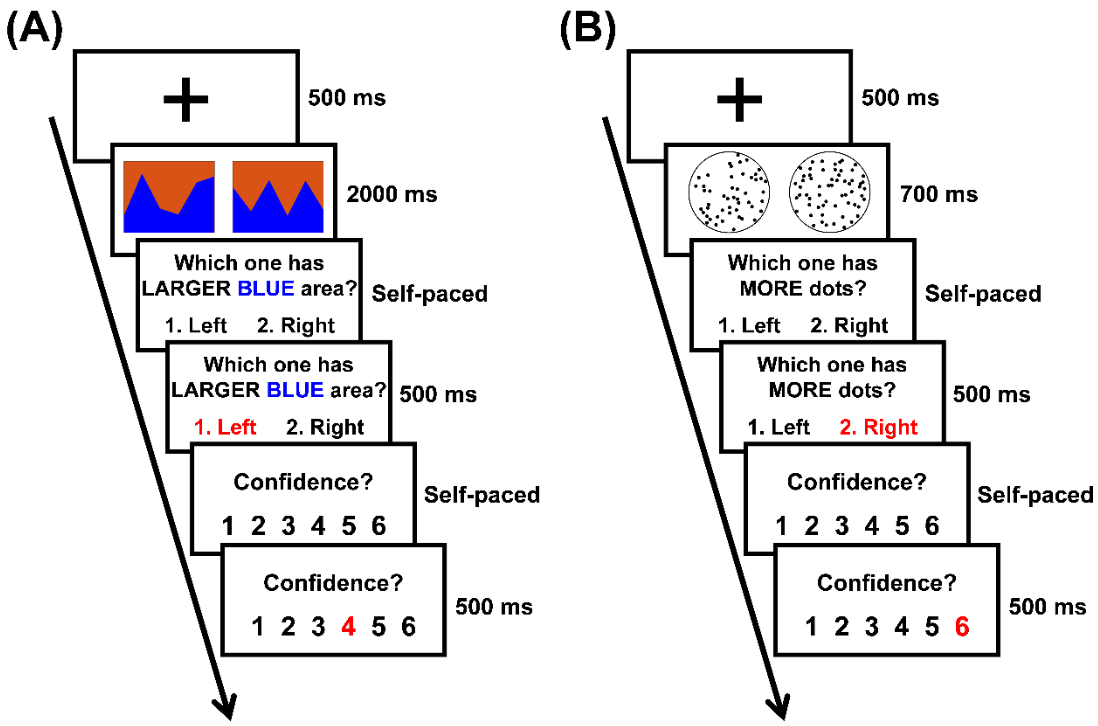

Figure 2. Procedure for the perceptual tasks in Experiments 1-3. Participants performed a color discrimination task in Experiment 1 (A), and a dot discrimination task in Experiment 2 (B). In Experiment 3, participants performed both perceptual tasks.

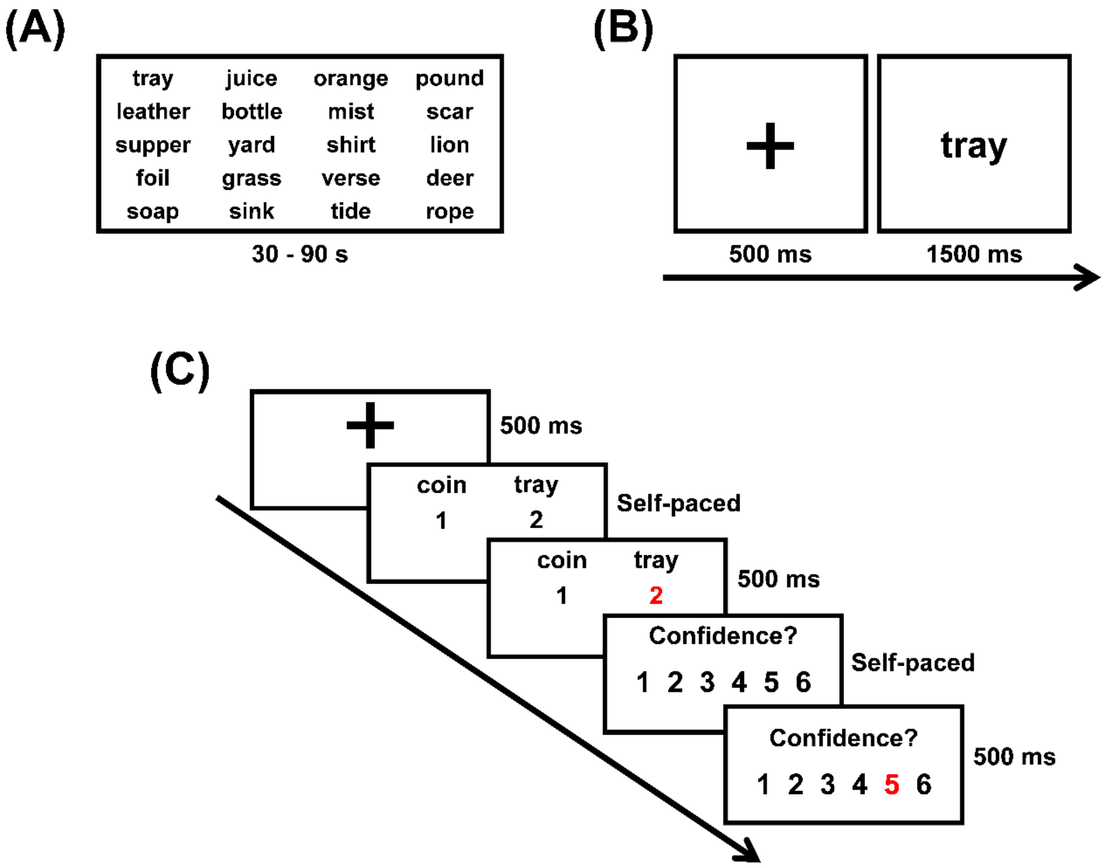

Figure 3. Procedure for the memory tasks in Experiments 1 and 2. The learning phase was different between Experiments 1 (A) and 2 (B), but the recognition memory test (C) was the same.

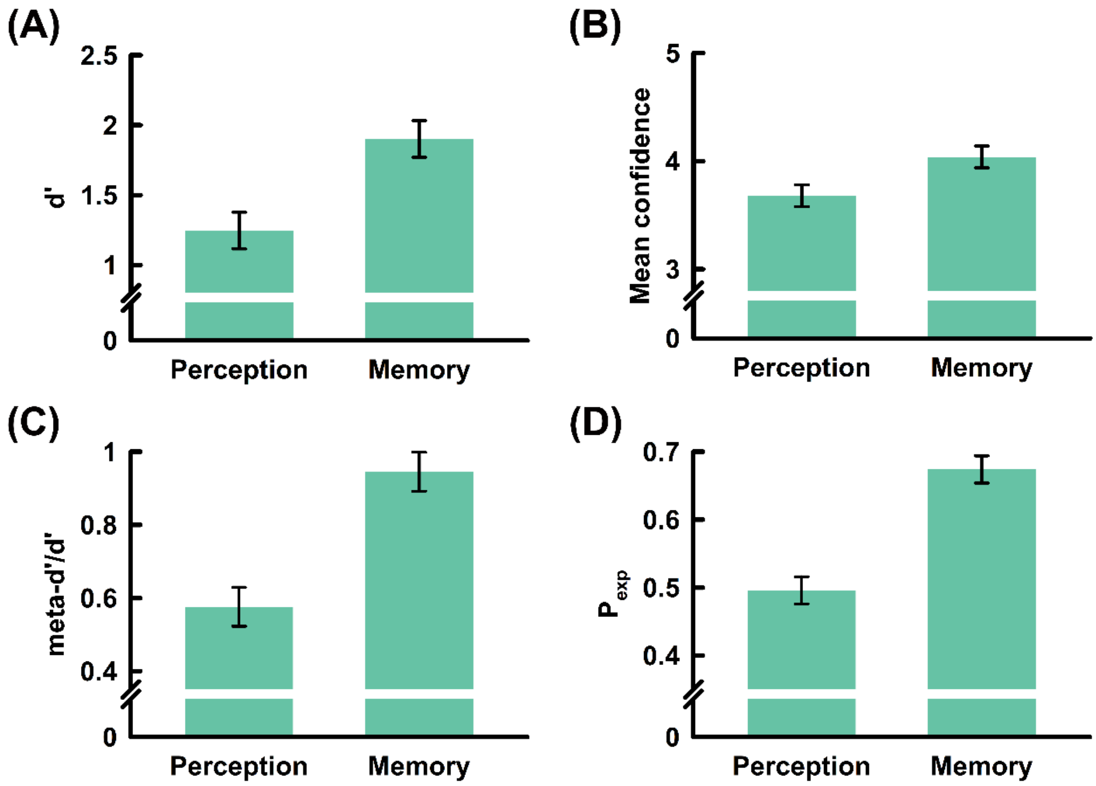

Figure 4. (A) Task performance (*d'*), (B) mean confidence, (C) metacognitive resolution (meta-*d'*/*d'*) and (D) contribution of processing experience to confidence ratings (P_{exp}) in Experiment 1. All of the four measures were significantly higher in memory task than those in perceptual task. Error bars represent within-participant 95% confidence interval (Cousineau, 2005).

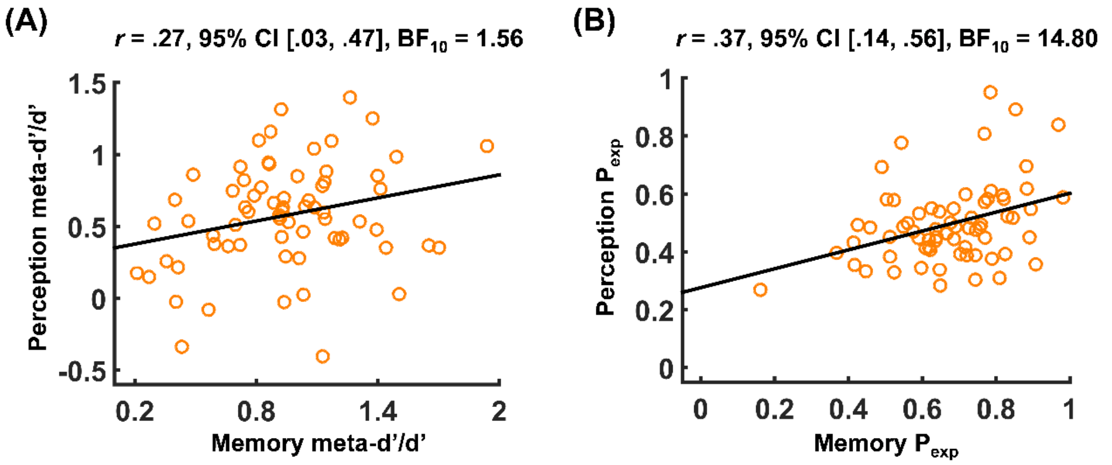

Figure 5. Relationship between meta-*d'*/*d'* in perceptual and memory tasks (A), and between P_{exp} in perceptual and memory tasks (B) in Experiment 1. Pearson correlation coefficients (and their associated statistical results) are reported in the figure.

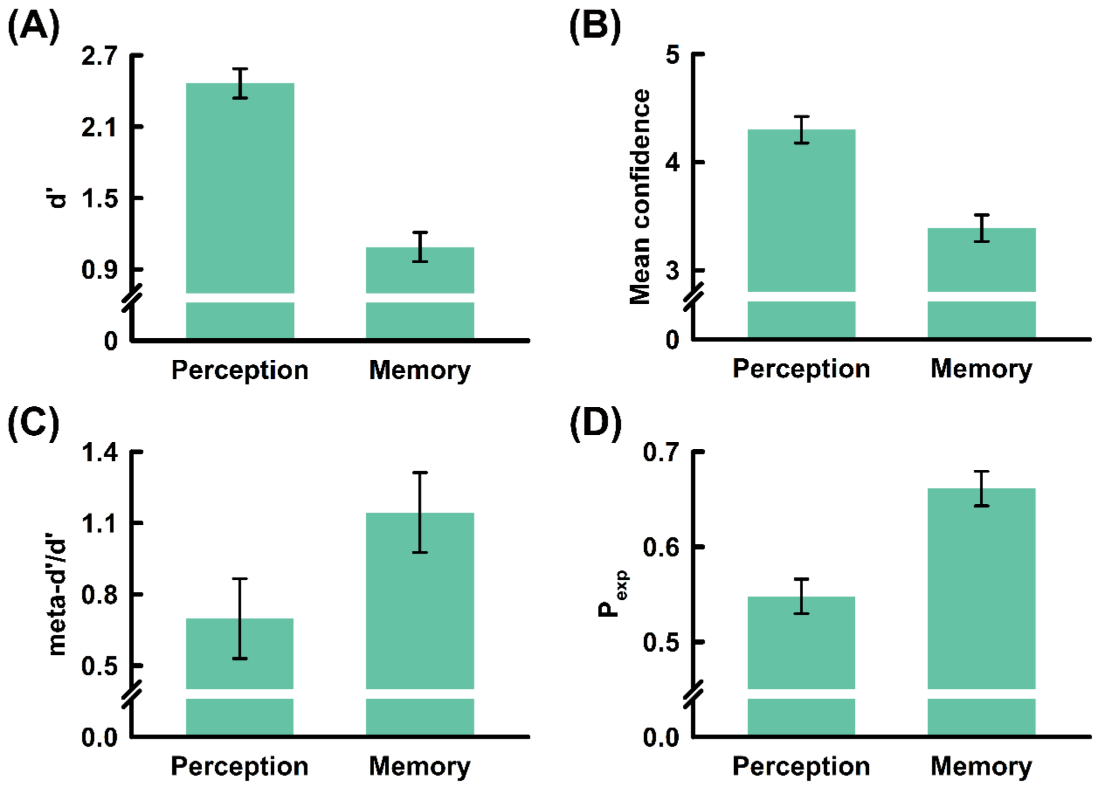

Figure 6. (A) Task performance (*d'*), (B) mean confidence, (C) metacognitive resolution (meta-*d'*/*d'*) and (D) contribution of processing experience to confidence ratings (P_{exp}) in Experiment 2. The *d'* and mean confidence were significantly higher in perceptual task than those in memory task, while the meta-*d'*/*d'* and P_{exp} were significantly higher in memory task. Error bars represent within-participant 95% confidence interval (Cousineau, 2005).

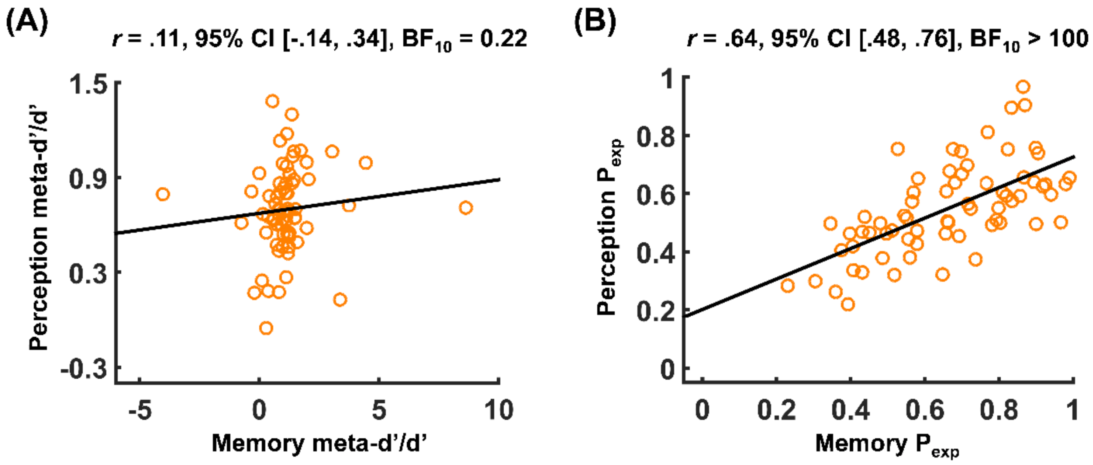

Figure 7. Relationship between meta-*d'*/*d'* in perceptual and memory tasks (A), and between P_{exp} in perceptual and memory tasks (B) in Experiment 2. Pearson correlation coefficients (and their associated statistical results) are reported in the figure.

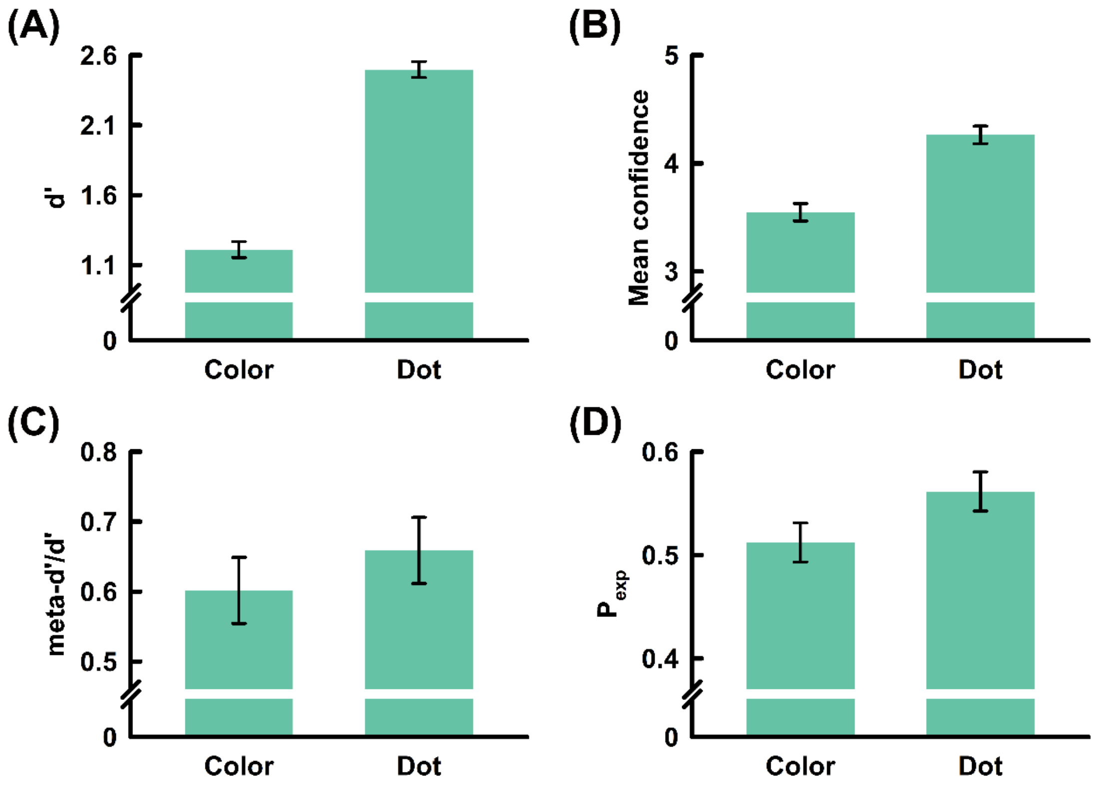

Figure 8. (A) Task performance (*d'*), (B) mean confidence, (C) metacognitive resolution (meta-*d'*/*d'*) and (D) contribution of processing experience to confidence ratings (P_{exp}) in Experiment 3. The *d'*, mean confidence and P_{exp} were significantly higher in dot discrimination task than those in color discrimination task (although the Bayes factor for P_{exp} was inconclusive). The meta-*d'*/*d'* did not reliably differ between the two tasks. Error bars represent within-participant 95% confidence interval (Cousineau, 2005).

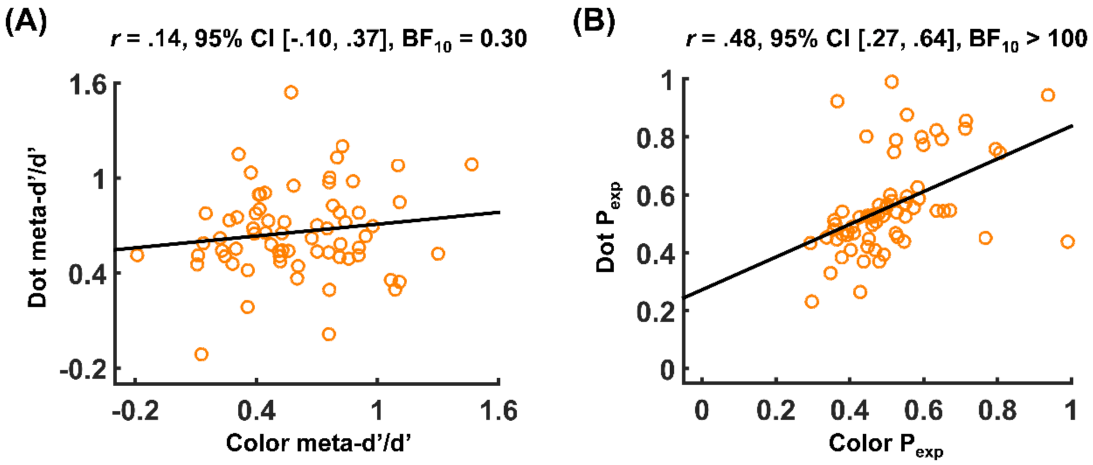

Figure 9. Relationship between meta-*d'*/*d'* in color and dot discrimination tasks (A), and between P_{exp} in color and dot discrimination tasks (B) in Experiment 3. Pearson correlation coefficients (and their associated statistical results) are reported in the figure.

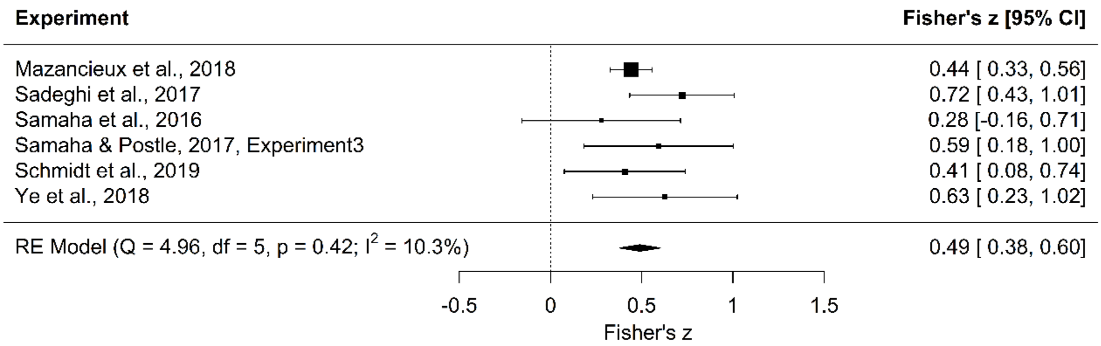

Figure 10. Forest plot for the meta-analysis on the cross-domain correlation of P_{exp} in the six experiments. Error bars represent 95% confidence interval.

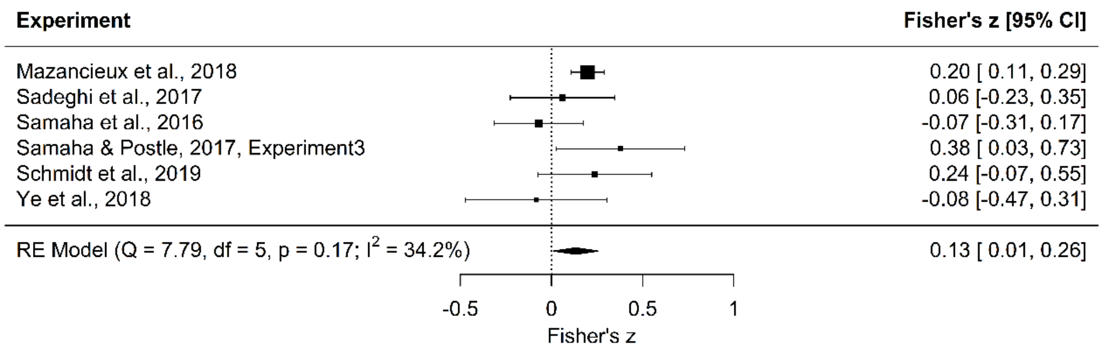

Figure 11. Forest plot for the meta-analysis on the cross-domain correlation of meta-*d'*/*d'* in the six experiments. Error bars represent 95% confidence interval.

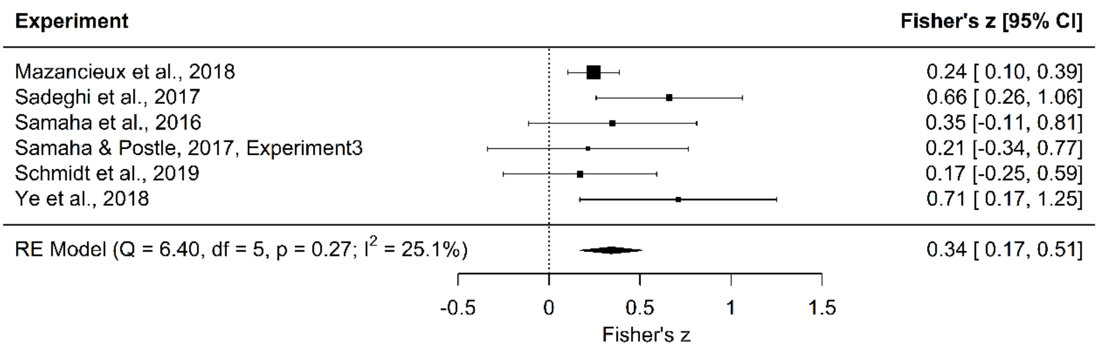

Figure 12. Forest plot for the meta-analysis on the difference between cross-domain correlation of P_{exp} and meta-*d'*/*d'* in the six experiments. Error bars represent 95% confidence interval.

Supplemental Materials

S1. Bayes Factor Robustness Check

In this section, we report the results from Bayes factor robustness check for the cross-task correlation of meta-*d'*/*d'* and P_{exp} in Experiments 1-3 (between a perceptual and a memory task in Experiments 1-2, and between two perceptual tasks in Experiment 3). The Bayes factor robustness check can examine how Bayes factor changes with the shape of prior distribution. In JASP, the prior distribution for correlation coefficient is stretched beta distribution, which can be seen as a beta distribution stretched from the 0-1 range to the range between -1 and 1. The shape of stretched beta prior (i.e., whether the prior distribution is narrow or wide) can be changed by altering the stretched beta prior width. By default, the value of stretched beta prior width is 1, indicating a flat prior.

Figure S1 shows the results from Bayes factor robustness check for the cross-task correlation of meta-*d'*/*d'*. In Experiment 1, the Bayes factor was inconclusive across a wide range of stretched beta prior width (the maximum value of BF_{10} was just above 3). In Experiments 2 and 3, the Bayes factors tended to support null hypothesis (H0), and BF_{10} decreased with increasing stretched beta prior width. Overall, there is a lack of evidence supporting positive cross-task correlation of meta-*d'*/*d'*.

Figure S2 shows the results from Bayes factor robustness check for the cross-task correlation of P_{exp}. In all of the three experiments, BF_{10} strongly supported alternative hypothesis H1 (i.e., $BF_{10} > 10$) for most value of stretched beta prior width, indicating a robust positive cross-task correlation of P_{exp} both between perceptual and memory domains (in Experiments 1-2) and within the perceptual domain (in Experiment 3).

S2. Standard Error for Averaged *z*-transformed Correlation Coefficients

When performing the meta-analyses in the current study, we averaged the *z*-transformed correlation coefficients between each perceptual and each memory task if there were multiple perceptual or memory tasks in one experiment, or if there were several within-participant conditions. In this section, we discuss how to calculate the standard error for the averaged *z*-transformed correlation coefficients.

Fisher's *r*-to-*z* transformation could be performed using the following equation:

$$z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$$

The standard error for the *z*-transformed correlation coefficient was equal to:

$$\sigma_z = \frac{1}{\sqrt{N-3}}$$

in which *N* represents sample size.

We can average multiple *z* values obtained from the same participant group:

$$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$$

in which n is the number of z values. The standard error for the averaged z value is (Fowler, 2011):

$$\sigma_{\bar{z}} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}\sigma_{Zi}\sigma_{Zj}} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\rho_{ij}}{N-3}}$$

In the equation above, ρ_{ij} is the correlation between z_i and z_j, and equal to 1 when $i = j$. Thus, we need to compute the value of ρ_{ij} when i is not equal to j.

Suppose x_1, x_2, x_3 and x_4 are four different variables obtained from the same group of participants. Here we use r_{ij} to represent the correlation coefficient between x_i and x_j, and z_{ij} denotes the z-transformed correlation. The equation for calculating the correlation between z_{12} and z_{34} is provided by Dunn and Clark (1969) (see also the Method section for Experiments 1):

$$\rho = [\frac{r_{12}r_{34}}{2}(r_{13}{}^2 + r_{14}{}^2 + r_{23}{}^2 + r_{24}{}^2) + (r_{13}r_{24} + r_{23}r_{14}) - (r_{13}r_{23}r_{34} + r_{14}r_{24}r_{34} + r_{12}r_{13}r_{14} + r_{12}r_{23}r_{24})] / (1 - r_{12}{}^2)(1 - r_{34}{}^2)$$

We can compute the standard error for the averaged z-transformed correlation coefficients based on the equations above.

S3. Examining The Relationship Between P_{exp} and Effect of Response Time on Confidence Ratings

In this section, we examined whether P_{exp} was related to the utilization of response time in the task during confidence rating process. To address this issue, we performed meta-analyses on data from eight experiments, including the six published experiments from the Confidence Database and Experiments 1-2 in the current study.

For each experiment, we first cleaned response time data in perceptual and memory tasks for each participant. Specifically, we excluded trials with response time that was shorter than 100 ms or differed by more than 3 standard deviations from the mean in each task. Next, for each participant, trial-by-trial confidence ratings were regressed on response time separately in each task, and the regression coefficient b was computed. Then we examined the correlation between individual differences in b and P_{exp} separately for perceptual and memory tasks (the value of P_{exp} was estimated based on the cleaned data), and performed meta-analysis on the z-transformed correlation.

Figures S3 and S4 show the meta-analytical results about the z-transformed correlation between b and P_{exp} in perceptual and memory tasks, respectively. The overall correlation between b and P_{exp} was significantly negative both in perceptual tasks, r = -.25, 95% confidence interval (CI) [-.37, -.13], and in memory tasks, r = -.25, 95% CI [-.38, -.11]. Previous studies indicate that higher confidence rating is associated with shorter response time in the task (e.g., Koriat & Ackerman, 2010; Siedlecka et al., 2019), and the regression coefficient b should be more negative when response time has a larger influence on confidence. Thus, the current meta-analyses suggest that processing experience contributed more to confidence ratings when the effect of response time on confidence was larger. Furthermore, heterogeneity amongst the effect sizes was statistically detectable in the meta-analysis for memory tasks, Q = 16.02, p = .025, but not for perceptual tasks, Q = 12.00, p = .101. The PET-PEESE analysis revealed no significant publication bias in any of the meta-analyses, ps > .30. We also performed the meta-analyses above using the value of P_{exp} estimated based on all data rather than cleaned data, which did not change the results.

Finally, we conducted a meta-analysis to examine whether the effect of response time on confidence ratings was domain-general. Results revealed that the overall correlation between the regression coefficient b in perceptual and memory domains was reliably positive, r = .36, 95% CI [.17, .52] (see Figure S5 for the z-transformed correlation). Thus, it is possible that the domain generality of the utilization of response time during confidence rating process might underlie the domain generality of P_{exp}. Furthermore, heterogeneity amongst the effect sizes was statistically detectable, Q = 33.16, p < .001. The PET-PEESE analysis revealed no significant publication bias, p > .80.

References

Dunn, O. J., & Clark, V. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association*, *64*(325), 366–377. https://doi.org/10.1080/01621459.1969.10500981

Fowler, J. W. (2011). *Density Function for the Sum of Correlated Random Variables*. http://web.ipac.caltech.edu/staff/fmasci/home/astro_refs/SumOfCorrelatedRVs.p

df

Koriat, A., & Ackerman, R. (2010). Choice latency as a cue for children's subjective confidence in the correctness of their answers. *Developmental Science*, *13*(3), 441–453. https://doi.org/10.1111/J.1467-7687.2009.00907.X

Siedlecka, M., Skóra, Z., Paulewicz, B., Fijałkowska, S., Timmermans, B., & Wierzchoń, M. (2019). Responses improve the accuracy of confidence judgements in memory tasks. *Journal of Experimental Psychology: Learning Memory and Cognition*, *45*(4), 712–723. https://doi.org/10.1037/xlm0000608

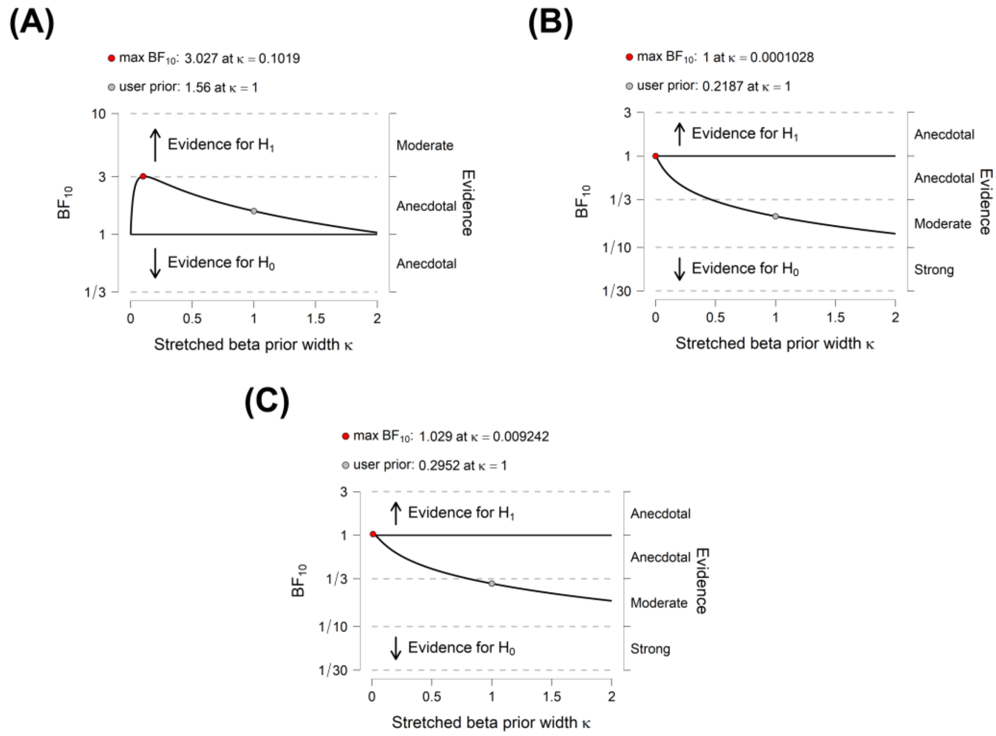

Figure S1. Results from Bayes factor robustness check for the cross-task correlation of meta-*d'*/*d'* in Experiments 1 (A), 2 (B) and 3 (C).

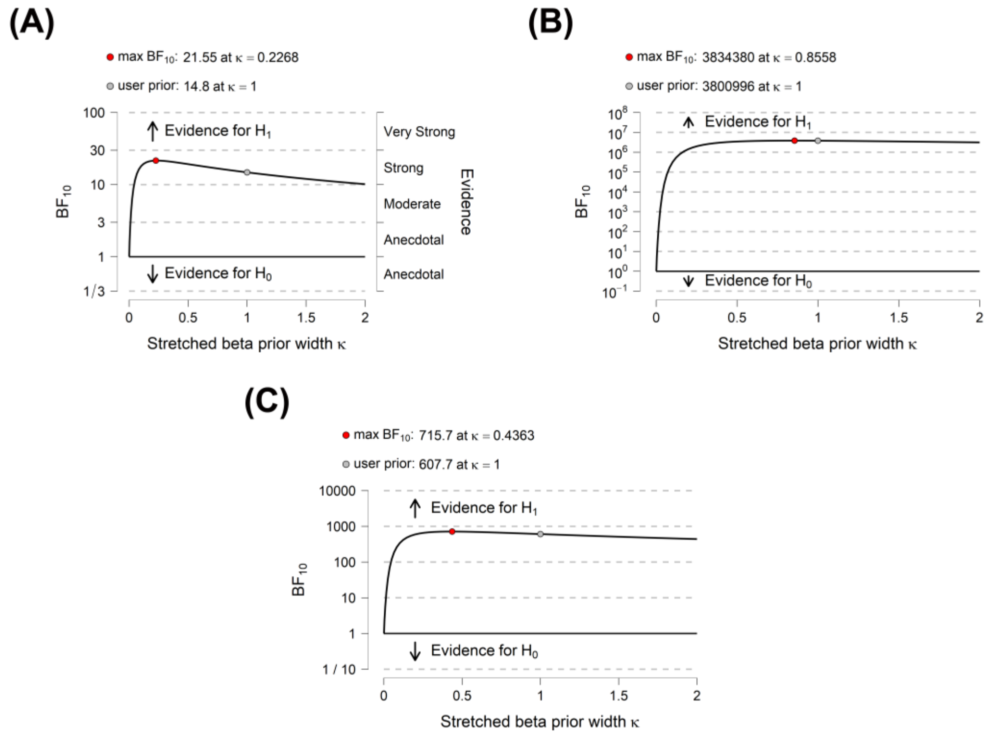

Figure S2. Results from Bayes factor robustness check for the cross-task correlation of P_{exp} in Experiments 1 (A), 2 (B) and 3 (C).

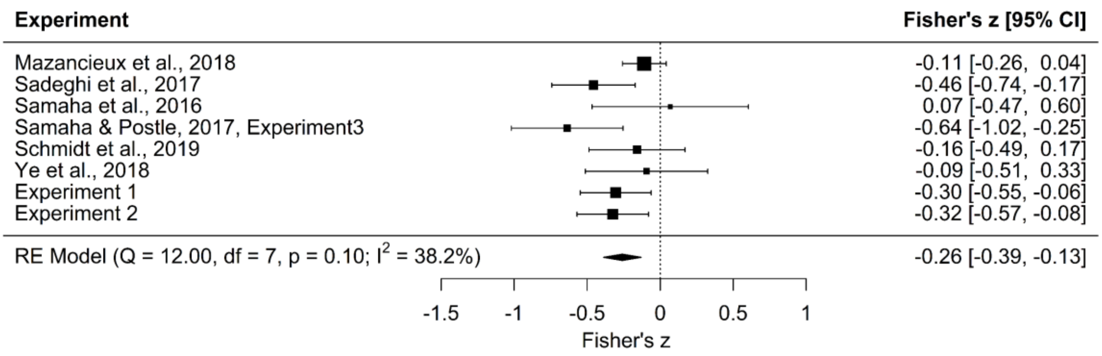

Figure S3. Forest plot for the meta-analysis on the correlation between P_{exp} and the regression coefficient *b* of confidence ratings on response time in perceptual tasks. Error bars represent 95% CI.

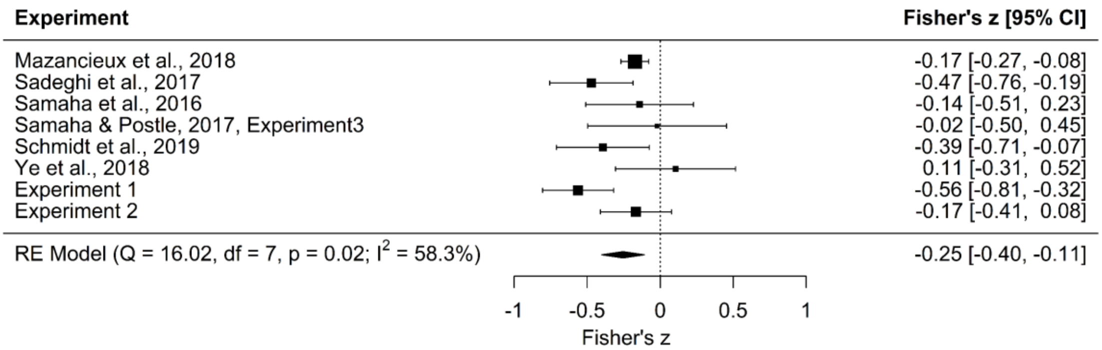

Figure S4. Forest plot for the meta-analysis on the correlation between P_{exp} and the regression coefficient *b* of confidence ratings on response time in memory tasks. Error bars represent 95% CI.

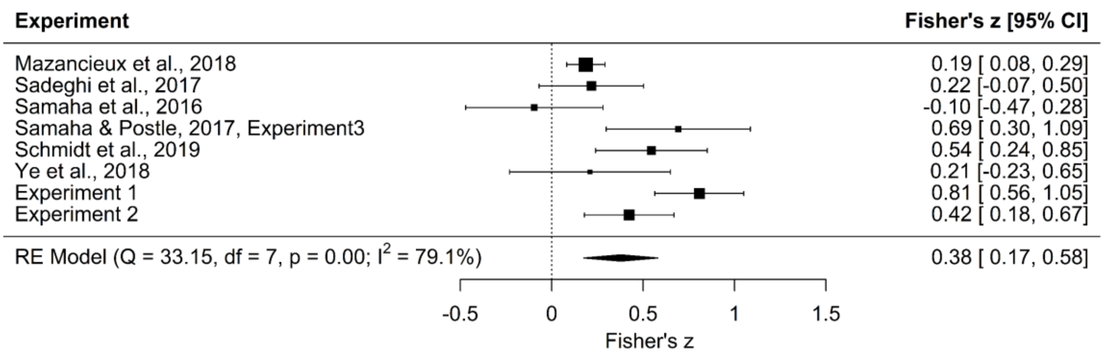

Figure S5. Forest plot for the meta-analysis on the correlation between the regression coefficient *b* of confidence ratings on response time in perceptual and memory domains. Error bars represent 95% CI.