

# Anatomical Coupling between Distinct Metacognitive Systems for Memory and Visual Perception

Li Yan McCurdy,<sup>1</sup> Brian Maniscalco,<sup>1</sup> Janet Metcalfe,<sup>1</sup> Ka Yuet Liu,<sup>2</sup> Floris P. de Lange,<sup>3</sup> and Hakwan Lau<sup>1,3</sup>

<sup>1</sup>Department of Psychology, Columbia University, New York, New York 10027, <sup>2</sup>Department of Sociology, University of California, Los Angeles, Los Angeles, California 90095, and <sup>3</sup>Radboud University Nijmegen, Donders Institute for Brain, Cognition, and Behaviour, 6500 HE Nijmegen, The Netherlands

A recent study found that, across individuals, gray matter volume in the frontal polar region was correlated with visual metacognition capacity (i.e., how well one's confidence ratings distinguish between correct and incorrect judgments). A question arises as to whether the putative metacognitive mechanisms in this region are also used in other metacognitive tasks involving, for example, memory. A novel psychophysical measure allowed us to assess metacognitive efficiency separately in a visual and a memory task, while taking variations in basic task performance capacity into account. We found that, across individuals, metacognitive efficiencies positively correlated between the two tasks. However, voxel-based morphometry analysis revealed distinct brain structures for the two kinds of metacognition. Replicating a previous finding, variation in visual metacognitive efficiency was correlated with volume of frontal polar regions. However, variation in memory metacognitive efficiency was correlated with volume of the precuneus. There was also a weak correlation between visual metacognitive efficiency and precuneus volume, which may account for the behavioral correlation between visual and memory metacognition (i.e., the precuneus may contain common mechanisms for both types of metacognition). However, we also found that gray matter volumes of the frontal polar and precuneus regions themselves correlated across individuals, and a formal model comparison analysis suggested that this structural covariation was sufficient to account for the behavioral correlation of metacognition in the two tasks. These results highlight the importance of the precuneus in higher-order memory processing and suggest that there may be functionally distinct metacognitive systems in the human brain.

## Introduction

What is the neural basis of metacognition, i.e., the introspective ability to monitor one's own mental processes (Metcalfe and Shimamura, 1994; Shimamura, 2008)? In a recent study, Fleming et al. (2010) reported a positive correlation between individuals' metacognitive capacity and anterior prefrontal cortex (aPFC) gray matter volume. Metacognitive capacity was quantified by measuring how well one's trial-by-trial confidence judgments discriminate between correct and incorrect responses on psychophysical tasks (Galvin et al., 2003).

One important question that arises is whether such putative metacognitive mechanisms in the aPFC are task specific or not. On one hand, the PFC has interconnections with virtually all sensory, motor, and memory systems (Miller and Cohen, 2001), so it is possible that the structural correlates in the aPFC reported by Fleming et al. (2010) reflect a general mechanism for various kinds of metacognitive behavior. Supporting this notion, Song et

al. (2011) recently demonstrated a positive behavioral correlation between metacognitive capacities on two different visual tasks. Using different task paradigms and methodology, other studies also support this claim (Schraw et al., 1995; Veenman et al., 1997; Schraw and Nietfeld, 1998; Veenman and Verheij, 2003; Veenman and Beishuizen, 2004). However, there is also empirical evidence and theoretical ideas to the contrary, suggesting distinct mechanisms involved with different kinds of metacognition (Glaser et al., 1992; Kelemen et al., 2000; Weaver and Kelemen, 2002; Schnyer et al., 2004; Pannu et al., 2005).

In this study, we tested whether metacognitive capacity on a word memory task and a visual perception task was behaviorally correlated across individuals, and whether they depended on the same neural structures. To do so, a technical challenge had to be overcome. It is known that metacognitive capacity is constrained by basic task performance (e.g., visual discrimination accuracy) (Galvin et al., 2003; Rotello et al., 2008). Therefore, measurements of memory and visual metacognition could be contaminated by variations in basic task performances. In the study by Fleming et al. (2010), this problem was circumvented by titrating the physical stimulus to keep basic visual task performance constant. However, such titration is relatively difficult to achieve in a memory task.

Here we used a recently developed psychophysical measure of metacognitive capacity to address this problem (Maniscalco and Lau, 2012). This new measure, known as meta- $d'$ , has the advantage of being expressed in the same units as the signal-to-noise ratio units for the standard signal-detection theoretic measure  $d'$ .

Received April 18, 2012; revised Aug. 29, 2012; accepted Nov. 23, 2012.

Author contributions: B.M., J.M., and H.L. designed research; L.Y.M. performed research; L.Y.M., B.M., K.Y.L., and H.L. analyzed data; L.Y.M., J.M., F.P.d.L., and H.L. wrote the paper.

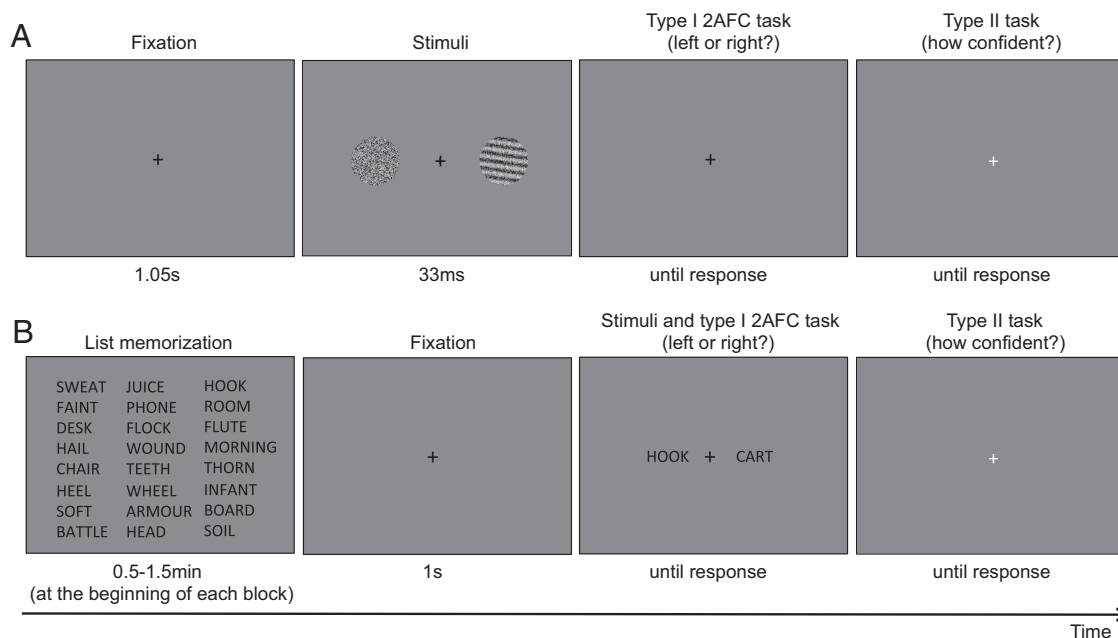
This work was supported by Templeton Foundation Grant 15462-SC104 (H.L.). We thank Ben Shababo for programming the memory experiment. We thank Barbara Knowlton and Anthony Wagner for valuable discussion, and Jesse Rissman for comments on an earlier version of the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Li Yan McCurdy, Department of Psychology, Columbia University, New York, NY 10027. E-mail: liyanmccurdy@gmail.com.

DOI:10.1523/JNEUROSCI.1890-12.2013

Copyright © 2013 the authors 0270-6474/13/331897-10\$15.00/0



**Figure 1.** Behavioral tasks. Participants performed both 2-AFC tasks. **A**, Visual task. Participants viewed two circular stimuli that were presented simultaneously to the left and right of fixation; one stimulus contained only visual noise, and the other contained a grating embedded in noise. Participants performed a 2-AFC judgment, indicating which stimulus (left or right) contained the grating. Subsequently, participants rated how confident they were that their 2-AFC judgment was correct using a 4-point scale (not shown on the screen). Participants were constrained to provide both responses within 5 s (see Materials and Methods). **B**, Memory task. At the beginning of each block of 50 trials, participants studied a list of words arranged in 10 rows and 5 columns (an 8 row  $\times$  3 column is shown here for ease of display). In each trial, participants viewed two words presented simultaneously to the left and right of fixation; one word had been presented on the study list and the other had not. Participants performed a 2-AFC judgment, indicating which word (left or right) was on the previously studied list. Subsequently, participants rated how confident they were that their 2-AFC judgment was correct using a 4-point scale (not shown on the screen). Both responses had to be provided within 5 s.

(which measures task performance capacity in a basic task, e.g., visual discrimination). Thus, although we cannot experimentally control for variability in basic task performance, this is easily corrected by normalizing meta- $d'$  by  $d'$ .

Using this method, we found a positive correlation between participants' metacognitive efficiency on the visual perception task and memory task. With this finding, we conducted a voxel-based morphometry (VBM) analysis to identify the neuroanatomical differences between participants that could explain our behavioral findings. We found that memory and visual metacognition depended on distinct neural structures that anatomically covaried in volume across individuals.

## Materials and Methods

**Participants.** Thirty-four healthy participants (18 females; 18–38 years old; mean age,  $25 \pm 5.1$  years) participated in this study for payment of €8 per hour. They were recruited from Radboud University Nijmegen (Nijmegen, The Netherlands). Participants had normal or corrected-to-normal vision and had no history of neurological or psychiatric disorders. Written informed consent was obtained from all participants. The research was approved by the local ethics committee where the experiment was performed (Commissie Mensgebonden Onderzoek Committee on Research Involving Human Subjects, region Arnhem-Nijmegen, The Netherlands).

**Stimuli and procedures.** All stimuli and experiments were programmed in MATLAB (MathWorks) using Psychtoolbox (Brainard, 1997). Participants sat 60 cm away from the personal computer screen. For every trial of the visual task, two circular stimuli ( $3^\circ$  diameter) were shown after a fixation period of 1.05 s. Stimuli were centered at  $4^\circ$  left and right of fixation for 33 ms. One stimulus contained visual noise (grayscale intensity values for each pixel were selected uniformly at random), and the other stimulus contained a grating (2 cycles/ $^\circ$ ) of random orientation embedded in visual noise (Fig. 1). The prominence of the grating with respect to the noise was titrated for each subject by adjusting luminance

contrast for the grating, using the QUEST threshold estimation procedure (Watson and Pelli, 1983). All stimuli were set to a constant overall level of 90% Michelson contrast.

For the word memory test, English words were generated using the Medical Research Council Psycholinguistic Database (Wilson, 1988). These standard nouns were four to eight letters long, had one to three syllables, and had a familiarity, concreteness, and imaginability rating of 400–700 each. At the beginning of each block of 50 trials of the memory task, 50 words (Calibri font, size 24) were presented simultaneously on the screen for 0.5, 1, or 1.5 min to create three levels of difficulty in which participants performed at neither chance nor ceiling. Participants were instructed to memorize as many words on the list as possible during the study period. A small notice appeared at the bottom of the screen to inform them when there was 10 s left to study the list. After the study period, a series of trials probing memory for the word list was presented. In each trial, two words were presented  $\sim 5.4^\circ$  to the left and right of fixation. One of these words had been presented on the study list ("old"), and the other word had not been presented previously ("new").

Both tasks required two responses per trial. First, participants provided a two-alternative forced-choice (2-AFC) judgment with regard to the spatial configuration of the stimuli on the screen (Fig. 1). For the visual task, participants indicated whether the stimulus containing the grating was located to the left or right of fixation. For the memory task, participants indicated whether the "old" studied word was located to the left or right of fixation. After the 2-AFC judgment, participants rated how confident they were that their 2-AFC judgment was correct, using a 4-point scale. No feedback was given about the participants' performance during the main experiment. To make these responses, participants used their left hand to press a key (either the "1" or "2" key on the number row of a standard QWERTY keyboard) to indicate whether they saw the grating (or studied word) to the left or right of fixation, respectively. After making this decision, the fixation cross changed color, prompting a response regarding their confidence. Participants pressed one of four keys ("7," "8," "9," or "0") using their right hand to indicate confidence in correctness of the 2-AFC judgment. Participants were in-

Should we do this as well?

How is our setup with keys?

Very Nice

structured that the scale represented confidence in a relative way (e.g., that a confidence rating of 4 would indicate high confidence relative to the typical level of confidence experienced in this task rather than a high absolute level of confidence) to encourage them to use the whole confidence scale. Both responses (2-AFC and confidence rating) had to be entered within 5 s of stimulus offset. If both responses were not entered within 5 s, the next trial commenced automatically, and that trial was omitted from data analysis.

Participants completed practice blocks for the visual and memory tasks before the actual experiment. The practice block for the visual task had 56 trials, in which trial-by-trial feedback was given regarding their response. After the practice block for the visual task, participants completed a calibration block consisting of 120 trials to determine the threshold level of grating contrast (relative to visual noise contrast) that yielded a target level of task performance of 80% correct for each subject. During the calibration, the contrast of the grating relative to the contrast in the visual noise in the target was adjusted continuously between trials using the QUEST threshold estimation procedure (Watson and Pelli, 1983). Three independent threshold estimates were acquired, with 40 randomly ordered trials contributing to each, and the median estimate of these was used in the main experiment. Fifty practice trials were conducted for the memory task. No threshold determination was performed for the memory test, but three different sets of study times were used for each subject, and each subject completed two or three blocks using each study time. The order of the study times and the word lists used were randomized across participants. Participants alternated between blocks of the visual task (102 trials per block) and blocks of the memory task (50 trials per block). A total of 510 ( $5 \times 102$ ) trials were completed for the visual task, and 400 ( $8 \times 50$ ) trials were completed for the memory task. The entire experiment was conducted over 2 days to reduce subject fatigue.

**Image acquisition.** A 1.5 T Avanto MR-scanner (Siemens), using a 32-channel head coil, was used to acquire the T1-weighted anatomical MRI images (176 slices; echo time, 2.95 ms; repetition time, 2250 ms; voxel size, 1 mm isotropic).

**Data analyses.** Metacognitive capacity was quantified using the measure meta- $d'$  (Maniscalco and Lau, 2012). It is known that measures of metacognitive capacity are constrained by basic task performance (Galvin et al., 2003; Maniscalco and Lau, 2012). As an intuitive example, if one performs a basic task (e.g., visual discrimination) at chance levels, one's confidence ratings theoretically would not be able to distinguish between the correct and incorrect trials, because the correct trials are only correct by chance guessing. Using meta- $d'$  as a measure of metacognitive capacity has the advantage of allowing us to circumvent this problem. Because meta- $d'$  is expressed in the same units of signal-to-noise ratio as  $d'$ , it is possible to directly compare meta- $d'$  to  $d'$  to assess an observer's metacognitive capacity. Meta- $d'$  is defined such that, if an observer is metacognitively ideal according to signal-detection theory, then meta- $d'$  equals  $d'$ , whereas suboptimal metacognition entails that meta- $d'$  is less than  $d'$ . Therefore, metacognitive efficiency can be assessed by dividing meta- $d'$  by  $d'$ , which takes into account the variability of  $d'$  across individuals when assessing metacognitive performance. Meta- $d'$  was calculated using the MATLAB code available at <http://www.columbia.edu/~bsm2105/type2sdt> (Maniscalco and Lau, 2012). Regression analysis was performed using “sregress” add-on function in Stata 12 (Verardi and Croux, 2009).

**VBM analyses.** VBM preprocessing was performed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>). Similar to the preprocessing protocol used by Fleming et al. (2010), the scans were first segmented into gray matter, white matter, and CSF in native space. VBM DARTEL (diffeomorphic anatomical registration through exponentiated lie algebra) algorithm (Ashburner, 2007) was used to increase the accuracy of inter-subject alignment, by aligning and warping the gray matter images to an iteratively improved template. The DARTEL template was then registered to Montreal Neurological Institute stereotactic space, and then gray matter images were modulated such that their tissue volumes were preserved. Images were smoothed using an 8 mm full-width at half-maximum Gaussian kernel.

The resultant preprocessed gray matter images were used in separate multiple regression design matrices in SPM8 to determine which brain

regions were correlated with the measures of visual or memory metacognitive efficiency. Participants' gender was included as a covariate. Proportional scaling was used to account for global brain volume variability across participants. A binary mask (template  $>0.3$ ) was used to exclude significant clusters outside the brain to limit the search volume to voxels likely to contain gray matter.

Two separate  $T$ -statistic maps reflecting the correlation between gray matter volume and memory or visual meta- $d'/d'$  were generated. An initial threshold of  $p < 0.001$  uncorrected was used, and clusters were identified. Expressing both memory and visual meta- $d'/d'$  in the same regression model yielded similar clusters. These clusters were used to define regions of interest (ROIs) using MarsBar version 0.42 software (marsbar.sourceforge.net). Small-volume correction was done to the clusters of interest by defining a 10 mm sphere at the five peak voxel coordinates presented by Fleming et al. (2010) that were found to be associated with their measure of metacognitive capacity [left aPFC, ( $-20, 53, 12$ ); right aPFC, ( $24, 65, 18$ ), ( $33, 50, 9$ ); dorsolateral PFC, ( $36, 39, 21$ ); precuneus, ( $6, -57, 18$ )]. The supplementary motor area (SMA) ROI that was used as a control was defined using the automated anatomical labeling ROI library (Tzourio-Mazoyer et al., 2002).

**Model comparison.** To compare between the models for the two metacognitive systems, we tested four structural equation models (see Fig. 5A) representing the conceptual relationships between the two types of metacognitive efficiencies and the two brain regions. Using structural equation models allowed us to model the effects of independent variables on more than one dependent variable simultaneously while allowing the errors to correlate. This is equivalent to fitting a seemingly unrelated regression model with maximum likelihood estimation (StataCorp, 2011). The two independent variables used were the PFC and precuneus gray matter volumes, and the two dependent variables were the visual and memory metacognitive efficiency measures. Model I assumes that both the PFC and precuneus measures affect both metacognitive performances. Model II assumes that the precuneus affects both metacognitive performances, whereas the PFC is only related to metacognitive performance of the visual task. Model III assumes that the PFC affects both types of metacognitive performance, whereas the precuneus affects only the memory metacognitive performance. Model IV is the most parsimonious model, which proposes the PFC affects visual metacognitive performance and the precuneus affects memory metacognitive performance (see Fig. 5A). The models were fitted using Stata12 (StataCorp).

The models were compared using the following indices. First,  $\chi^2$  statistics were used to compare the deviance ( $-2LL$ ) of Models I to IV against the most parsimonious model, Model IV. An insignificant difference in the  $\chi^2$  statistics indicates that the two models under comparison are statistically not different in terms of goodness of fit. Root mean square error of approximation (RMSEA) is a measure of the discrepancy between the predicted and the observed data per degree of freedom, and a cutoff of 0.05 or below has been suggested to indicate a good fit (Hu and Bentler, 1999). Although the RMSEA captures the difference between the predicted and the actual data, comparative fit index (CFI) measures the discrepancy of the target model with a null model, which assumes independence between the variables. A CFI value of 0.93 or larger has been suggested to indicate a good fit (Byrne, 1994). RMSEA and CFI are less sensitive to sample size (Fan et al., 1999) and are often used together because they measure different aspects of the goodness of fit of a model. Last, we used two information theoretic methods: the Bayesian information criterion (BIC) and the Akaike information criterion (AIC). Both criteria for model selection quantify goodness of model fit by rewarding close fits to the data and penalizing model complexity, as indexed by the number of free model parameters. One difference between the AIC and BIC is that the penalty term added for the number of parameters in the model is larger in the latter than the former. The AIC analysis was included in addition to the BIC analysis because the former has been shown to have theoretical advantages over the latter (Burnham and Anderson, 2002). Specifically, we calculated the BIC and AIC values for each model, as defined by the following equations:

$$BIC_i = -2\log(L_i) + k\log(n),$$

$$AIC_i = -2\log(L_i) + 2k,$$

How many blocks can we run?

We do it in one go?

Comment on HRT

→

I would need help getting these numbers out

Metacognitive efficacy, as measured by m-ratio (Barrett method, using the meta-SDT package in R by Craddock) <https://osf.io/yhv2k>



where  $L$  is the maximum likelihood for the model  $i$ ,  $k$  is the number of free parameters to be estimated, and  $n$  is the sample size.

This value was then used to calculate the Akaike weights, as defined by the following:

$$\text{Akaike weight}(w) = \frac{e^{-0.5(\text{AIC}_i - \text{AIC}_{\min})}}{\sum_{r=1}^R e^{-0.5(\text{AIC}_r - \text{AIC}_{\min})}},$$

where  $\text{AIC}_{\min}$  is the AIC value of the best candidate model.

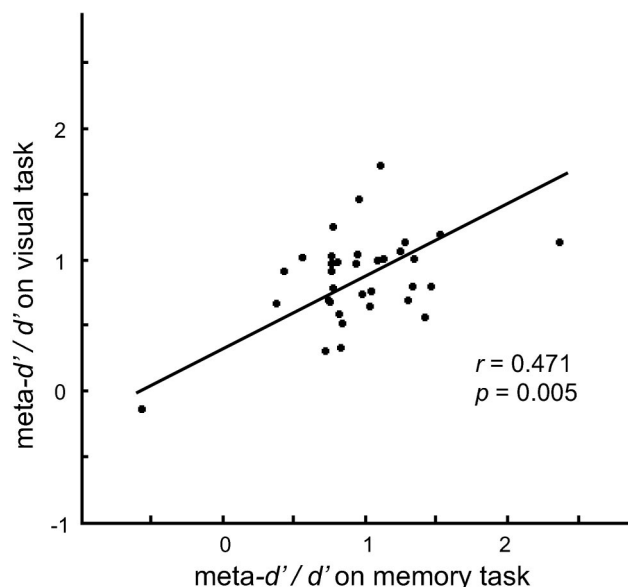
Akaike weights (Akaike, 1973) are another method for determining the strength of evidence for each model and are interpreted as a measure of conditional probabilities for each model (Wagenmakers and Farrell, 2004). The probability that the particular model is the best model is defined by its ability to minimize the Kullback–Leibler divergence—the distance between a reference distribution and the distribution generated by a model—given the data and the set of other models (Kullback and Leibler, 1951).

## Results

On average, subjects missed 1.68 trials in the visual task (0.33% of all visual trials) and 12.35 trials in the memory task (3.09% of all memory trials). All missed trials were omitted from analysis. As expected, participants' basic task performance was not correlated across the memory and visual tasks ( $r = -0.032$ ,  $p = 0.855$ ).

Each participant's confidence level averaged across all trials on each task taken separately was strongly correlated with their averaged basic performance ( $d'$ ) on that task (memory task,  $r = 0.565$ ,  $p < 0.001$ ; and visual task,  $r = 0.564$ ,  $p < 0.001$ ). We also found that participants had a similar confidence bias across modalities, i.e., participants who tend to rate confidence as high on the visual task also tend to do so on the memory task ( $r = 0.484$ ,  $p = 0.004$ ). These are factors that could potentially influence the calculation of metacognitive capacity using traditional methods, introducing potential confounds into such measures (Galvin et al., 2003), although our recently developed measure, meta- $d'$ , is immune to these potential confounds (Maniscalco and Lau, 2011).

To control for the influence of basic task performance, we divided meta- $d'$  by  $d'$  to calculate each subject's metacognitive efficiency (also referred to as meta- $d'/d'$ ). This estimates the amount of signal strength that is available for metacognition, expressed as a fraction of the amount of signal strength that is available for the primary discrimination task. This ratio has a value of 0 when one is metacognitively "blind" and a value of 1 when one is metacognitively "ideal" according to signal-detection theory (i.e., when all information available for the primary discrimination task is also available for metacognition). We calculated the meta- $d'/d'$  value for each modality for every subject and found that the difference between the average visual meta- $d'/d'$  was not significantly different from the average memory meta- $d'/d'$  ( $p = 0.152$ ). One subject had negative meta- $d'/d'$  values on both tasks. This is because the subject's meta- $d'$  values were slightly below zero, despite above-chance  $d'$  performance. Although this seems to suggest that this subject had "negative metacognitive ability," i.e., that the subject's confidence ratings carried negative information about his/her performance, this negative value is likely to be attributable to estimation error. That is, the true value was likely to be zero or slightly higher than zero, but the influence of estimation error yielded an estimate with a negative value of small magnitude (memory meta- $d'/d' = -0.562$ , visual meta- $d'/d' = -0.143$ ). Excluding these data could introduce biases into the analysis by selectively removing data



**Figure 2.** Behavioral data. Metacognitive efficiency (as quantified by meta- $d'/d'$ ) on the memory task was positively correlated with metacognitive efficiency on the visual task ( $r = 0.471$ ,  $p = 0.005$ ). One may worry that this effect was driven by potential outliers, because estimates of meta- $d'/d'$  here seem to produce extreme and implausible values in some individuals. To ensure that this effect is genuine and robust, we used a robust regression, as well as a nonparametric Spearman's rank correlation test, both of which downplay the effects of outliers. Significant results were obtained using both methods.

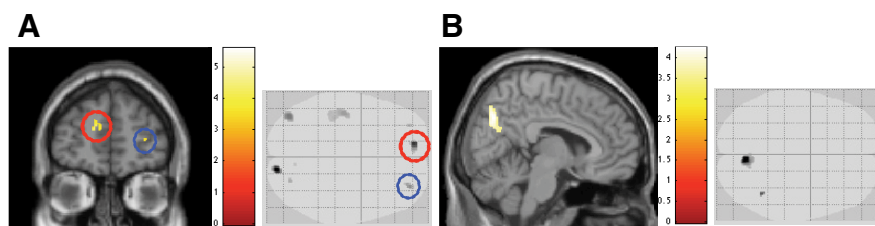
points with negative estimation error while not excluding data points with positive estimation error. For this reason, it is conventional in standard signal-detection analysis to not exclude negative  $d'$  values from analysis. Thus, we did not exclude this subject's data from analysis despite the negative meta- $d'$  value.

A significant across-subject correlation between the meta- $d'/d'$  value for memory and visual metacognition was found ( $r = 0.471$ ,  $p = 0.005$ ; Fig. 2). A Deming regression was plotted to show the linear relationship between visual and memory metacognition. The Deming regression is an errors-in-variables model for finding the best-fit line while accounting for errors in the dependent and independent variables (Deming, 1943). The simple linear regression was not suitable for plotting the regression line and assessing the precise quantitative relationship between the two variables because it assumes error only exists for the dependent variable.

To ensure that the results were not influenced by potential outliers, we ran a robust regression to determine the significance of the correlation. Robust statistics are designed specifically to deal with the issue of outliers in a principled and objective manner. Specifically, we ran an "S-estimation" regression. It is an example of robust estimators that are resistant to outliers because they use estimators that minimize a measure of dispersion of the residuals that is less sensitive to extreme values than the variance (Rousseeuw and Yohai, 1984). This makes S-estimation regression more resistant to outliers compared with the standard ordinary least-squares method, which tends to distort parameters' estimation in the presence of outliers (Verardi and Croux, 2008). This analysis produced a positive regression line that was statistically significant ( $p = 0.035$ ). We do note that the significance value obtained from the robust regression is substantially lower than that of the Pearson's correlation analysis, which suggests that the true correlation strength uncontaminated by outlier influence, although significant, is likely to be lower than the  $r$  value of 0.471 as reported in Figure 2.

Negative  $d'$ ?  
What to do?

Plan what  
to do  
with this



**Figure 3.** Gray matter volume correlations with metacognitive efficiency. **A**, Statistical ( $T$ ) maps shown for both a standard overlay and axial “glass brain,” showing areas in which gray matter volume correlates positively with meta- $d'/d'$  on the visual task. **B**, Statistical ( $T$ ) maps for positive correlation with meta- $d'/d'$  on the memory task, for both standard overlay and axial glass brain. The significant cluster was found in the precuneus region. All images were thresholded at  $p < 0.001$  uncorrected for display purposes, but the circled clusters (in **A**) and the precuneus (in **B**) pass small volume correction for multiple comparisons.

As a verification that meta- $d'/d'$  was not confounded by the basic task performance, we found no correlation between metacognitive efficiency (meta- $d'/d'$ ) and basic task performance ( $d'$ ) for both the visual ( $r = -0.169$ ,  $p = 0.341$ ) and the memory ( $r = -0.213$ ,  $p = 0.227$ ) tasks. Thus, despite differences in basic task performance within participants, we were able to control for it and to uncover the unbiased correlations across participants between metacognitive efficiency in the two different tasks.

We went on to relate this variability in metacognitive efficiency to inter-individual differences in brain structure. We found that participants' metacognitive efficiency (meta- $d'/d'$ ) on the visual task was positively correlated with regions in aPFC (Fig. 3A). Two of these clusters correspond with those regions identified by Fleming et al. (2010) (their Table S2, the third and fifth clusters listed as positively correlated with metacognitive capacity) and statistically they survived small-volume correction for multiple comparisons [peak voxel coordinate for left aPFC,  $(-12, 54, 16)$ ,  $T = 5.01$ , cluster familywise error (FWE)-corrected  $p = 0.023$ ; peak voxel coordinate for right aPFC,  $(32, 50, 7)$ ,  $T = 4.04$ , cluster FWE-corrected  $p = 0.025$ ].

Likewise, we conducted a similar analysis on the metacognitive efficiency (meta- $d'/d'$ ) on the memory task and found that it was positively correlated with the precuneus [peak voxel coordinates,  $(8, -64, 24)$ ,  $T = 3.55$ , cluster FWE corrected  $p = 0.031$ ]. This cluster also passed small-volume correction based on the peak coordinates of the results of Fleming et al. (2010, their Table S2, second cluster). Hence, we were able to identify two general regions associated with the two measures of metacognitive efficiency: two clusters in the aPFC were correlated with visual metacognition, and a cluster in the precuneus was correlated with memory metacognition.

After conducting the whole-brain analysis, we used ROI analyses to further investigate the pattern of the data. The main purpose of this analysis was not to make new statistical inferences (which would have been problematic and circular because of the way the ROIs were identified) but to exploit the higher sensitivity in ROI analyses so as to reveal weaker trends in the data for illustration and motivation for the modeling analysis described below. The three identified clusters were used to define ROIs using the MarsBar toolbox (Brett et al., 2002). To obtain the most robust estimate of aPFC volume, we combined both aPFC clusters in the region to produce an average; all subsequent analyses refer to these combined data as aPFC. Interestingly, there was a significant positive correlation between the volume of the precuneus ROI and visual metacognitive efficiency ( $r = 0.375$ ,  $p = 0.029$ ), although this effect was weak and could only be observed using this ROI approach of analysis. At whole-brain SPM analyses, even at a liberal threshold of 0.05 uncorrected, we did not find

any voxels in this region that showed a correlation with visual metacognitive efficiency (meta- $d'/d'$ ). For the aPFC, there was no correlation with memory metacognitive efficiency ( $r = 0.218$ ,  $p = 0.215$ ), even with this sensitive ROI approach.

Our data suggest that the two metacognitive processes (memory and visual) depend on relatively distinct brain structures (precuneus and aPFC, respectively). However, how can this finding account for the behavioral correlation between memory and visual metacognitive efficiencies across participants? One possibility is that, although the two metacognitive systems are essentially distinct, there is some degree of functional overlap, perhaps particularly in the precuneus in which we found a weak effect for visual metacognition (only using the ROI analysis). It thus could be that the precuneus contains a common mechanism for both memory and visual metacognition, which could explain the behavioral correlation across individuals.

The results from an interaction analysis using SPM support this notion. The interaction analysis was run by including both sets of metacognition values as predictors in the general linear model (GLM). (Subjects' gender was also included as a covariate.) We looked for regions that were significantly more correlated with visual metacognition than with memory metacognition, by using a contrast with a positive unit coefficient for the visual metacognition term and the negative coefficient for the memory metacognition term [i.e., setting a  $T$  contrast in the GLM as  $(1, -1)$ , respectively]. We found that small-volume correction for a cluster in the left aPFC voxel was statistically significant (cluster-level FWE-corrected  $p = 0.025$ ). This suggests that the aPFC was more involved with visual metacognition than memory metacognition.

Looking at the converse set of correlations (i.e., regions that were more correlated with visual metacognition than memory metacognition) yielded few voxels even at  $p < 0.001$  uncorrected, and they were located in neither the frontal polar nor precuneus area. This is in line with our finding that precuneus volume was significantly correlated with both memory and visual metacognition, and thus the precuneus was not implicated more in memory metacognition than visual metacognition. Thus, it appears that the precuneus correlation with memory metacognition was not significantly greater than with visual metacognition and that the aPFC correlated with visual metacognition was significantly greater than with memory metacognition.

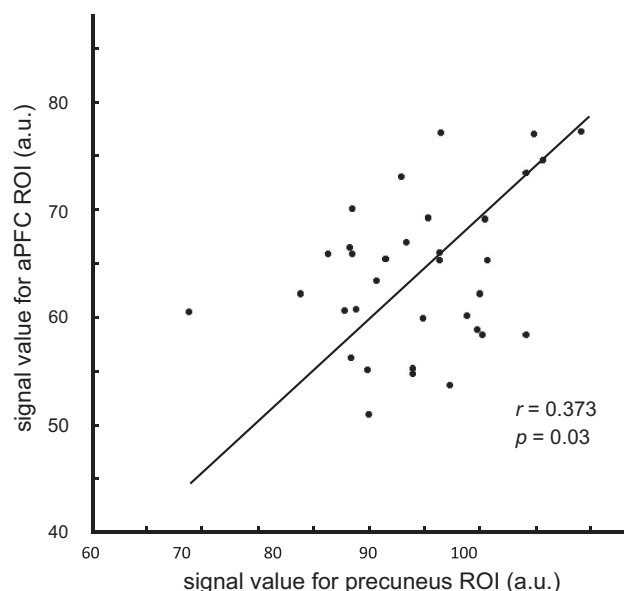
However, we also found that, across individuals, the precuneus and the aPFC covaried in volume, i.e., participants with bigger volume in one region tend to have bigger volume in the other ( $r = 0.373$ ,  $p = 0.03$ ; Fig. 4). As with Figure 2, we plotted a Deming regression to visualize the linear relationship between the two brain regions and then applied robust regression analysis to this plot to determine whether this correlation was statistically significant or not, while taking into account the effect of potential outliers. The robust regression analysis produced a statistically significant regression plot ( $p < 0.001$ ). This anatomical covariability could also potentially account for our behavioral results, even if the two systems are functionally distinct. Notably, this covariability might account for the generality of metacognitive skills observed by some research teams (Schraw et al., 1995; Veenman et al., 2004).

One concern was that the volume of other, task-unrelated brain regions might also be correlated with aPFC and precuneus volumes. To determine whether the correlation of aPFC volume and precuneus volume was anatomically specific rather than reflective of a widespread correlation of brain volumes across multiple regions, we selected the SMA as our control area and found that there was no significant correlation between the gray matter volume in the precuneus and the gray matter volume in this control area ( $r = 0.170$ ,  $p = 0.336$ ), and neither was there a correlation between the gray matter volume in our aPFC region and the control ( $r = -0.060$ ,  $p = 0.738$ ). This suggests that these interactions were relatively specific to the two identified brain regions.

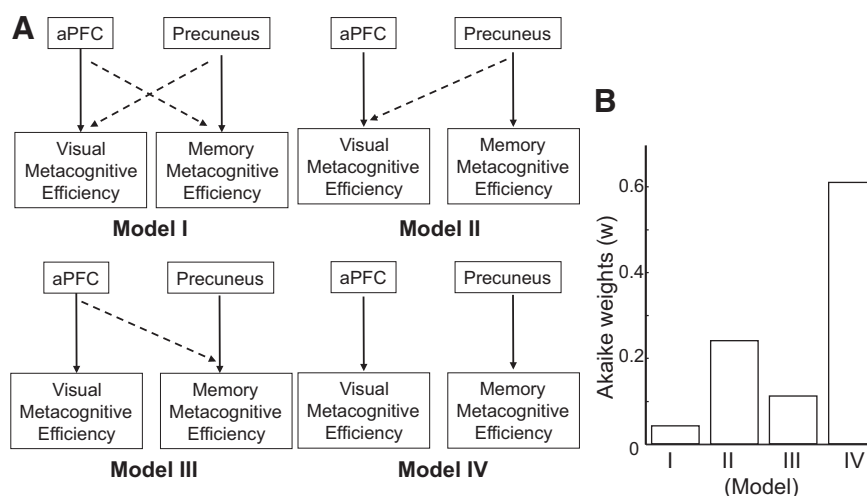
We also ran two separate GLM analyses, using precuneus volume or aPFC volume as a predictor, and found no significant clusters elsewhere in the brain at a significance level of (uncorrected)  $p = 0.001$ . Although this did not reveal other areas that were shown to be significantly correlated in volume, the lack of results could merely be attributable to a limit of statistical power. Nonetheless, both these analyses suggest that the relationship between the two types of metacognition and the two identified brain regions was relatively specific.

With this confirmation that these two specific brain regions were correlated with the two kinds of metacognition, there still remained the task of determining whether the correlation between the two metacognitive tasks was attributable to “functional crosstalk” (as suggested by the weak correlation between the precuneus volume and visual metacognition) or that the behavioral correlation was merely attributable to the anatomical correlation as seen in Figure 4. To distinguish between these two possibilities (functional vs anatomical coupling), we considered four formal models for the two metacognitive systems, in which there are varying degrees of functional crosstalk (Fig. 5A), using structural equation modeling. The results are reported in Table 1. First, the precuneus–visual metacognitive efficiency path was statistically insignificant when included in the models (Models I and II,  $p = 0.20$  in both models). The aPFC–memory metacognitive efficiency path was highly insignificant when included (Models I and III,  $p = 0.905$  and  $p = 0.945$ ) and the standardized parameter estimates were close to zero, indicating there was little relationship between aPFC and memory performance in our data. In other words, there was no statistical evidence to suggest the crosstalk paths, especially the aPFC–memory metacognitive efficiency path, in our data.

The same conclusion was reached by looking at the results of the likelihood ratio tests. The  $\chi^2$  statistics of Models I–III, which included one or both of the crosstalk paths, did not differ significantly from that of most parsimonious model (Model IV). The RMSEA values were lower than the cut point of 0.05 for both Models II and IV, and CFI values were above the cut point of 0.93, indicating that both models were good



**Figure 4.** Gray matter volume covaries between precuneus and frontal polar region (aPFC). The gray matter volumes in the two ROIs were plotted against each other for each subject. Gray matter volume in the aPFC and precuneus are positively correlated with gray matter volume in the precuneus ROI.



**Figure 5.** Model comparison. **A**, Schematic description of the models. These models characterize the possible ways of explaining the positive behavioral correlation between memory and visual metacognitive efficiencies. In Model I, the solid arrows indicate that the aPFC is mainly functionally responsible for visual metacognitive efficiency, and the precuneus is mainly functionally responsible for memory metacognitive efficiency. The dashed lines indicate that there may be some degree of functional crosstalk between the two systems, such that the precuneus may also be partially responsible for visual metacognitive efficiency and the aPFC may also be partially responsible for memory metacognitive efficiency. Models II and III are variants in which the crosstalk is one-sided. In Model IV, there is no functional crosstalk, i.e., the precuneus is functionally responsible for only memory but not visual metacognitive efficiency, and the PFC is functionally responsible for only visual but not memory metacognitive efficiency. In this model, the behavioral correlation between visual and memory metacognitive efficiency is accounted for entirely by the covariation in volume between the precuneus and the PFC across individuals (as shown in Fig. 4). **B**, Model statistics and comparison. AIC model selection was used to estimate which of the four models (in **A**) was most likely. AIC quantifies goodness of model fits by rewarding close fits to data while penalizing model complexity (number of model parameters), because more complex models are prone to overfitting. Akaike weights calculated from the AIC values can be interpreted as the probability that that particular model is best. Thus, Model IV fits the data best.

fits according to these measures. However, given that both models fit the data, Model IV was the better model given the likelihood ratio test result. Although the CFI of Model III was above the cut point, its RMSEA did not meet the cut point of  $<0.05$ , suggesting that Model III showed a larger discrepancy



**Table 1. Statistical significance of each path in each model**

	Model I	Model II	Model III	Model IV
<b>Structural</b>				
Intercept–visual ME				
Standard intercept	−4.421*	−4.337*	−2.693*	−2.665*
SE	1.613	1.601	1.117	1.049
aPFC–visual ME				
Standard $\beta$	0.499*	0.504*	0.569*	0.567*
SE	0.007	0.123	0.116	0.111
Precuneus–visual ME				
Standard $\beta$	0.188	0.185		
SE	0.147	0.144		
Intercept–memory ME				
Standard intercept	−5.652*	−5.727*	−5.177*	−5.128*
SE	1.379	1.230	1.564	1.402
Precuneus–memory ME				
Standard $\beta$	0.630*	0.623*	0.578*	0.582*
SE	0.117	0.105	0.126	0.113
aPFC–memory ME				
Standard $\beta$	−0.017		0.010	
SE	0.146		0.149	
<b>Mean <math>\pm</math> SE</b>				
aPFC	64.178 $\pm$ 1.196			
Precuneus	89.588 $\pm$ 1.218			
<b>Variance (standard variance)</b>				
$\epsilon$ (visual ME)	0.645	0.642	0.676	0.679
$\epsilon$ (memory ME)	0.611	0.611	0.661	0.661
<b>Covariance (Corr)</b>				
aPFC and Precuneus	0.373*			
$\epsilon$ (visual ME) and $\epsilon$ (memory ME)	0.390*	0.390*	0.400*	0.400*
<b>Model statistics</b>				
<i>n</i>	34	34	34	34
AIC	509.021	507.034	508.59	506.595
BIC	530.390	526.877	538.433	526.911
RMSEA	0.000**	0.000	0.129	0.000
CFI	1.000**	1.000	0.982	1.000
<b>Likelihood ratio test against Model IV</b>				
$\chi^2$ (df)	1.57 (2)	1.56 (1)	0.00 (1)	
<i>p</i>	0.455	0.212	0.945	

The table reports whether each path (from a particular brain region to a particular kind of metacognition) is statistically significant. ME, Metacognitive efficiency; Corr, correlation. \* $p < 0.05$ . \*\* indicates the saturated model. Standardized  $\beta$  values represent standardized coefficient values. RMSEA expresses fit per degree of freedom of the model; values of RMSEA of  $<0.08$  imply an acceptable model fit and values of  $<0.05$  imply a good fit. *p* values indicate that, for all four models, the paths from the aPFC to visual metacognition and from the precuneus to memory metacognition are statistically significant. Crosstalk paths, i.e., from the precuneus to visual metacognition (Model II), from the aPFC to memory metacognition (Model III), or both (Model I), are not statistically significant. Thus, Model IV is the most parsimonious model in which all paths are statistically significant, and there is no evidence to suggest that crosstalk between the aPFC and precuneus causes the behavioral correlation between visual and memory metacognition.

from the observed data than the other models. (Model I was a saturated model and necessarily had a CFI of 1 and RMSEA of 0).

We also calculated the AIC (an information criterion-based measure) for each model and found that Model IV had the lowest value, indicating that Model IV had the highest model likelihood and/or lowest model complexity. Finally, we tabulated the Akaike weights (Fig. 5B), which can be interpreted as the probability that a model is the correct one. That data indicated that Model IV was the most probable model given our data, followed by Model II. In summary, our model comparison reaffirmed that Model IV was the best-fitting model, i.e., that the two metacognitive systems operate mostly independently, and that the behavioral correlation between memory and visual metacognition seems to be sufficiently explained by the fact that the precuneus and aPFC covary in volume across individuals.

Finally, given the two distinct correlations (aPFC–visual metacognition and precuneus–memory metacognition), we considered what the nature of this relationship was like. With the

behavioral and anatomical correlations plotted in Figures 2 and 4, respectively, one consideration was whether the behavioral results could be explained simply by the anatomical relationship. We found that the Deming regression line in Figure 2 was significantly different from zero ( $p < 0.001$ ) and unity ( $p < 0.001$ ). This corresponds with the derived slope of 0.553, which indicates that a unit increase in memory metacognition leads to a smaller increase in visual metacognition. However, the Deming regression for Figure 4 depicting the anatomical correlation was not statistically different from unity ( $p < 0.01$ ), because the regression line had a slope of 0.953. This suggests that the relationship between the anatomical and behavioral correlation may not simply be identical, because the sub-unity behavioral relationship in metacognition is not attributable to the unity anatomical correlation. It thus appears that, although the two brain regions seem to operate independently to bring about the two forms of metacognition, the direct mapping between the gray matter volume of the brain regions to metacognition may not be completely realistic.

## Discussion

In this study, we found that an individual's metacognitive ability measured on a visual perception task and on a word memory task were correlated. We also found that these measures of metacognitive ability, although correlated, were attributed to distinct brain regions and that the volumes of these regions seemed to have somewhat independent correlations with the two metacognitive abilities, respectively. The fact that the two relevant brain regions (precuneus and aPFC) correlated in volumes across individuals seemed to be sufficient to account for the behavioral correlation between memory and visual metacognitive efficiencies.

Our findings lead to a potentially interesting interpretation of previous results. The aPFC is often associated with metacognition (Christoff et al., 2003; Ramnani and Owen, 2004; Fleming and Dolan, 2012; Fleming et al., 2012), and there is indirect evidence that the metacognitive mechanisms in the aPFC may be general for different visual tasks (Song et al., 2011). As mentioned in the Introduction, although some studies have suggested that there may be distinct metacognitive mechanisms underlying different tasks, content, and cognitive processes, it has also been suggested, mainly on theoretical grounds, that each cognitive process may be monitored by its own metacognitive system (Nelson and Narens, 1990; Dodson et al., 1998; Kelemen et al., 2000; Weaver and Kelemen, 2002). Other researchers have also found evidence to the contrary, that there may be a general mechanism for different kinds of metacognition. Within the context of this controversy, our results point to a potential resolution: although our study is one that examines the structural relationships in the brain to explain the behavioral correlation and thus we cannot conclude anything definitive about whether different types of metacognition are functionally distinct and mutually exclusive, one possibility that would resolve this controversy would be that the different brain regions involved with different kinds of metacognition are not functionally but anatomically coupled.

That the frontal polar region and precuneus seem to be responsible for visual metacognition and memory metacognition, respectively, is compatible with our knowledge of the anatomical connections of these areas. The frontal polar region receives input from higher cognitive areas (Ramnani and Owen, 2004), including the dorsolateral PFC, which in turn has been characterized as the apex of the visual processing hierarchy (among other things), receiving input from secondary sensory processing

regions from both the dorsal and the ventral stream (Young, 1992). The precuneus, located in the medial parietal region, is densely connected with the middle temporal lobe, which is heavily involved in memory processing (Wagner et al., 2005). Thus, from an anatomical standpoint, it is plausible that these two regions are situated in positions to perform higher-order monitoring functions for the visual and memory modalities, respectively.

However, somewhat puzzling is the fact that previous studies on memory metacognition seem to point to the PFC rather than the precuneus as an important substrate (Janowsky et al., 1989; Schnyer et al., 2004; Pannu and Kaszniak, 2005; Modirrousta and Fellows, 2008; Chua et al., 2009). Some of these studies are concerned with prospective metacognitive judgments (Schnyer et al., 2004), which may differ from the kind of metacognition concerned here, which is based on post-trial confidence ratings. However, Pannu et al. (2005) also reported deficits in post-trial confidence-based memory metacognition after frontal lobe damage. In another brain imaging study (Yokoyama et al., 2010), it was found that the level of brain activation in the frontal polar region reflects metacognitive capacity as assessed with post-trial confidence in a memory task.

Perhaps such apparent differences between these previous findings and our result are partly attributable to the difference in method. VBM aims to reveal the variability in brain volume in relevant neural structures that reflect individual differences in performance. It does not strictly follow that the identified region is more active during the relevant tasks in individuals showing superior performance, because the relationship between VBM structural results and functional brain imaging findings remains complex and incompletely understood (Kanai and Rees, 2011). A multidisciplinary approach incorporating different techniques, such as diffusion tensor imaging, which identifies anatomical connectivity between brain regions, would be required to gain a clearer understanding of the functional and anatomical connectivity in the human brain (for review, see Ramnani et al., 2004). Also, the lack of significant correlation between memory metacognition and prefrontal volume is a negative finding that could have been attributable to the limitations in statistical power.

Regarding the relevant neuropsychological data, although Pannu et al. (2005) reported that damage to the PFC can lead to impairments in post-trial confidence-based metacognition in memory processes, it is also possible that such prefrontal damage may have distal impact on posterior regions, including the precuneus. We would like to clarify that our present findings (a relatively specific correlation between precuneus gray matter volume and memory metacognition) do not imply that the sole function of the precuneus is to provide memory metacognition, nor does it imply that memory metacognition solely arises from precuneus and not from other brain regions. However, based on our findings, our prediction would be that, even when memory performance is matched between precuneus lesion patients and controls, the lesion patients would show additional deficits in metacognitive performance. However, this empirical issue is yet to be resolved because damage to this region is relatively rare, and, when it happens, such patients may not be tested specifically for memory metacognition. It has been noted that this cortical area receives relatively little attention, perhaps in part as a result of its hidden location and the relative paucity of data from lesion studies (Cavanna and Trimble, 2006).

One other possible reason for the discrepancy in results could be because of the differences in how metacognition is assessed

empirically (i.e., task structure) and analytically (i.e., what measure is used to quantify metacognition). Most of these studies mentioned measure metacognitive capacity by assessing the raw correlations between accuracy and confidence, which may be influenced by primary task performance and thus may reflect a combination of basic task performance and metacognitive performance (Galvin et al., 2003), whereas here in this study, we use metacognitive efficiency (i.e.,  $\text{meta-}d'/d'$ ) to quantify metacognition, which reflects pure metacognitive performance without contamination of basic task performance. We also used a 2-AFC task to assess memory performance, which is relatively uncommon in previous studies.

Nonetheless, we should note that the precuneus is also identified in the VBM study of Fleming et al. (2010), although the focus there was on visual processing. There is also substantial evidence of precuneus involvement in memory processes, especially regarding confidence judgments (Yonelinas et al., 2005) and other self- and awareness-related processes (Kircher et al., 2000; Ochsner et al., 2004; Cavanna and Trimble, 2006). In general, the medial parietal lobe is implicated in various studies on facilitating episodic memory retrieval (Tulving et al., 1994; Henson et al., 1999; Wagner et al., 2005); lesions to parietal midline structures have been shown to lead to amnesia (Valenstein et al., 1987). What is lacking is specific evidence, collected using quantitative methods, to relate functional activity in this region to memory metacognition. However, this may well be because studies on this specific topic are relatively rare (for example, compared to studies on episodic retrieval), and future research can hopefully address this gap in the literature. We hope that the present study represents a step toward this goal.

Finally, we also acknowledge some limitations in experimental design that prevented us from exploring several important issues. First, because we sought in part to replicate the VBM study by Fleming et al. (2010), we followed their design and focused on post-trial confidence as our measure of metacognitive capacity. Others have investigated other forms of metacognition that involve prospective judgments, such as feeling of knowing (Hart, 1965), judgment of learning (Arbuckle, 1969), or other types of summary judgments of performance, e.g., an overall sense of agency or control (Miele et al., 2011). At least in the memory domain, it is known that some aspects of prospective are different from retrospective metacognitive judgments (Costermans et al., 1992), and hence different judgments may depend on dissociable neural substrates to some extent (Schnyer et al., 2004; Pannu et al., 2005; but see Choa et al., 2009). Therefore, the interpretation of the present results is restricted to post-trial confidence-based metacognition and may not generalize to prospective metacognitive judgments.

To conclude, using our novel psychophysical measure of metacognition,  $\text{meta-}d'$ , we were able to provide an extension of the previous finding that metacognition between two visual perception experiments is correlated (Song et al., 2011). This measure allowed us to test metacognitive ability in modalities whose stimuli cannot be adjusted to control for fluctuations in objective performance. Although this positive correlation between metacognitive ability may *prima facie* suggest a general mechanism behind metacognition that is independent of modality, a closer look at the anatomical correlates across individuals suggests that metacognition in the two different modalities is associated with two distinct structural regions that operate relatively independently of each other.



## References

- Akaike H (1973) Information theory and an extension of the maximum likelihood principle (Petrov BN, Csaki F, eds), pp 267–281. Second International Symposium on Information Theory. Budapest: Akademiai Kiado.
- Arbuckle T (1969) Discrimination of item strength at time of presentation. *J Exp Psychol* 8:126–131.
- Ashburner J (2007) A fast diffeomorphic image registration algorithm. *Neuroimage* 38:95–113. [CrossRef Medline](#)
- Brett M, Anton JL, Valabregue R, Poline JB (2002) Region of interest analysis using an SPM toolbox. *Neuroimage* 16:2.
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach, Ed 2. Berlin: Springer.
- Byrne BM (1994) Structural equation modeling with EQS and EQS/Windows. Basic concepts, applications, and programming. Thousand Oaks, CA: Sage.
- Cavanna AE, Trimble MR (2006) The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129:564–583. [CrossRef Medline](#)
- Christoff K, Ream JM, Geddes LP, Gabrieli JD (2003) Evaluating self-generated information: anterior prefrontal contributions to human cognition. *Behav Neurosci* 117:1161–1168. [CrossRef Medline](#)
- Chua EF, Schacter DL, Sperling RA (2009) Neural correlates of metamemory: a comparison of feeling-of-knowing and retrospective confidence judgments. *J Cogn Neurosci* 21:1751–1765. [CrossRef Medline](#)
- Costermans J, Lories G, Ansary C (1992) Confidence level and feeling of knowing in question answering: the weight of inferential processes. *J Exp Psychol* 18:142–150. [CrossRef](#)
- Deming WE (1943) Statistical adjustment of data. New York: Wiley.
- Dodson CS, Holland PW, Shimamura AP (1998) On the recollection of specific- and partial-source information. *J Exp Psychol Learn Mem Cogn* 24:1121–1136. [CrossRef Medline](#)
- Fan X, Thompson B, Wang L (1999) The effects of sample size, estimation methods, and model specification on SEM fit indices. *Struct Equ Model* 6:56–83. [CrossRef](#)
- Fleming SM, Dolan RJ (2012) The neural basis of metacognitive ability. *Philos Trans R Soc Lond B Biol Sci* 367:1338–1349. [CrossRef Medline](#)
- Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G (2010) Relating introspective accuracy to individual differences in brain structure. *Science* 329:1541–1543. [CrossRef Medline](#)
- Fleming SM, Dolan RJ, Frith CD (2012) Metacognition: computation, biology and function. *Philos Trans R Soc Lond B Biol Sci* 367:1280–1286. [CrossRef Medline](#)
- Galvin SJ, Podd JV, Drga V, Whitmore J (2003) Type 2 tasks in the theory of signal detectability: discrimination between correct and incorrect decisions. *Psychon Bull Rev* 10:843–876. [CrossRef Medline](#)
- Glaser R, Schauble L, Raghavan K, Zeitz C (1992) Scientific reasoning across different domains. In: Computer-based learning environments and problem solving (de Corte E, Linn MC, Mandl H, Verschaffel L, eds), NATO ASI series F, Vol 84, pp, 345–371. Heidelberg: Springer.
- Hart J (1965) Memory and the feeling-of-knowing experience. *J Educ Psychol* 56:208–216. [CrossRef](#)
- Henson RN, Rugg MD, Shallice T, Josephs O, Dolan RJ (1999) Recollection and familiarity in recognition memory: an event-related functional magnetic resonance imaging study. *J Neurosci* 19:3962–3972. [Medline](#)
- Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model* 6:1–55. [CrossRef](#)
- Janowsky JS, Shimamura AP, Kritchevsky M, Squire LR (1989) Cognitive impairment following frontal lobe damage and its relevance to human amnesia. *Behav Neurosci* 103:548–560. [CrossRef Medline](#)
- Kanai R, Rees G (2011) The structural basis of inter-individual differences in human behaviour and cognition. *Nat Rev Neurosci* 12:231–242. [CrossRef Medline](#)
- Kelemen WL, Frost PJ, Weaver CA 3rd (2000) Individual differences in metacognition: evidence against a general metacognitive ability. *Mem Cognit* 28:92–107. [CrossRef Medline](#)
- Kircher TT, Senior C, Phillips ML, Benson PJ, Bullmore ET, Brammer M, Simmons A, Williams SC, Bartels M, David AS (2000) Towards a functional neuroanatomy of self processing: effects of faces and words. *Brain Res Cogn Brain Res* 10:133–144. [Medline](#)
- Kullback S, Leibler RA (1951) On information and sufficiency. *Ann Math Stat* 22:79–86. [CrossRef](#)
- Maniscalco B, Lau H (2012) A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Conscious Cogn* 21:422–430. [CrossRef Medline](#)
- Metcalfe J, Shimamura AP (1994) Metacognition: knowing about knowing. Cambridge, MA: Massachusetts Institute of Technology.
- Miele DB, Wager TD, Mitchell JP, Metcalfe J (2011) Dissociating neural correlates of action monitoring and metacognition of agency. *J Cogn Neurosci* 23:3620–3636. [CrossRef Medline](#)
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202. [CrossRef Medline](#)
- Modirrousta M, Fellows LK (2008) Medial prefrontal cortex plays a critical and selective role in “feeling of knowing” meta-memory judgments. *Neuropsychologia* 46:2958–2965. [CrossRef Medline](#)
- Nelson TO, Narens L (1990) Metamemory: a theoretical framework and new findings. In: The psychology of learning and motivation (Bower GH, ed), pp 1–45. New York: Academic.
- Ochsner KN, Knierim K, Ludlow DH, Hanelin J, Ramachandran T, Glover G, Mackey SC (2004) Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. *J Cogn Neurosci* 16:1746–1772. [CrossRef Medline](#)
- Pannu JK, Kaszniak AW (2005) Metamemory experiments in neurological populations: a review. *Neuropsychol Rev* 15:105–130. [CrossRef Medline](#)
- Pannu JK, Kaszniak AW, Rapcsak SZ (2005) Metamemory for faces following frontal lobe damage. *J Int Neuropsychol Soc* 11:668–676. [CrossRef Medline](#)
- Ramnani N, Owen AM (2004) Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nat Rev Neurosci* 5:184–194. [CrossRef Medline](#)
- Ramnani N, Behrens TE, Penny W, Matthews PM (2004) New approaches for exploring anatomical and functional connectivity in the human brain. *Biol Psychiatry* 56:613–619. [CrossRef Medline](#)
- Rotello CM, Masson ME, Verde MF (2008) Type I error rates and power analyses for single-point sensitivity measures. *Percept Psychophys* 70:389–401. [CrossRef Medline](#)
- Rousseeuw P, Yohai V (1984) Robust and nonlinear time series analysis. *Lecture Notes in Statistics* #26, pp 256–272. Berlin: Springer.
- Schnyer DM, Verfaellie M, Alexander MP, LaFleche G, Nicholls L, Kaszniak AW (2004) A role for right medial prefrontal cortex in accurate feeling-of-knowing judgements: evidence from patients with lesions to frontal cortex. *Neuropsychologia* 42:957–966. [CrossRef Medline](#)
- Schraw G, Nietfeld J (1998) A further test of the general monitoring skill hypothesis. *J Educ Psychol* 90:236–248. [CrossRef](#)
- Schraw G, Dunkle ME, Bendixen LD, Roedel TD (1995) Does a general monitoring skill exist? *J Educ Psychol* 87:433–444. [CrossRef](#)
- Shimamura AP (2008) A neurocognitive approach to metacognitive monitoring and control. In: Handbook of metamemory and memory (Dunlosky J, Bjork RA, eds), pp 373–390. New York: Psychology.
- Song C, Kanai R, Fleming SM, Weil RS, Schwarzkopf DS, Rees G (2011) Relating inter-individual differences in metacognitive performance on different perceptual tasks. *Conscious Cogn* 20:1787–1792. [CrossRef Medline](#)
- StataCorp (2011) Stata structural equation modeling reference manual, Release 12. College Station, TX: StataCorp.
- Tulving E, Kapur S, Markowitsch HJ, Craik FI, Habib R, Houle S (1994) Neuroanatomical correlates of retrieval in episodic memory: auditory sentence recognition. *Proc Natl Acad Sci U S A* 91:2012–2015. [CrossRef Medline](#)
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289. [CrossRef Medline](#)
- Valenstein E, Bowers D, Verfaellie M, Heilman KM, Day A, Watson RT (1987) Retrosplenial amnesia. *Brain* 110:1631–1646. [CrossRef Medline](#)
- Veenman MV, Beishuizen JJ (2004) Intellectual and metacognitive skills of novices while studying texts under conditions of text difficulty and time constraint. *Learn Instruct* 14:619–638.
- Veenman MV, Verheij J (2003) Identifying technical students at risk: Relat-

- ing general versus specific metacognitive skills to study success. *Learn Individ Differ* 13:259–272.
- Veenman MV, Elshout JJ, Meijer J (1997) The generality vs. domain-specificity of metacognitive skills in novice learning across domains. *Learn Instruct* 7:187–209. [CrossRef](#)
- Verardi V, Croux C (2009) MM\_REGRESS: Stata module to compute robust regression estimates. Statistical Software Components S457057. Boston: Boston College Department of Economics.
- Verardi V, Croux C (2008) Robust regression in Stata. FBE Research Report KBI\_0823, pp 1–13. Leuven, Belgium: K.U. Leuven.
- Wagenmakers EJ, Farrell S (2004) AIC model selection using Akaike weights. *Psychon Bull Rev* 11:192–196. [CrossRef Medline](#)
- Wagner AD, Shannon BJ, Kahn I, Buckner RL (2005) Parietal lobe contributions to episodic memory retrieval. *Trends Cogn Sci* 9:445–453. [CrossRef Medline](#)
- Watson AB, Pelli DG (1983) QUEST: a Bayesian adaptive psychometric method. *Percept Psychophys* 33:113–120. [CrossRef Medline](#)
- Weaver CA 3rd, Kelemen WL (2002) Comparing processing-based, stimulus-based, and subject-based factors in metacognition. In: *Metacognition: process, function, and use* (Chambres P, Izaute M, Marescaux P, eds), pp 49–60. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Wilson MD (1988) The MRC psycholinguistic database: machine readable dictionary, Version 2. *Behav Res Methods Instrum Comput* 20:6–11. [CrossRef](#)
- Yokoyama O, Miura N, Watanabe J, Takemoto A, Uchida S, Sugiura M, Horie K, Sato S, Kawashima R, Nakamura K (2010) Right frontopolar cortex activity correlates with reliability of retrospective rating of confidence in short-term recognition memory performance. *Neurosci Res* 68:199–206. [CrossRef Medline](#)
- Yonelinas AP, Otten LJ, Shaw KN, Rugg MD (2005) Separating the brain regions involved in recollection and familiarity in recognition memory. *J Neurosci* 25:3002–3008. [CrossRef Medline](#)
- Young MP (1992) Objective analysis of the topological organization of the primate cortical visual system. *Nature* 358:152–155. [CrossRef Medline](#)