Research

Cite this article: Samaha J, Postle BR. 2017 Correlated individual differences suggest a common mechanism underlying metacognition in visual perception and visual short-term memory. *Proc. R. Soc. B* **284**: 20172035.
<http://dx.doi.org/10.1098/rspb.2017.2035>

Received: 14 September 2017
 Accepted: 20 October 2017

Subject Category:
 Neuroscience and cognition

Subject Areas:
 cognition, behaviour

Keywords:
 metacognition, visual perception, short-term memory, confidence, signal detection theory

Author for correspondence:
 Jason Samaha
 e-mail: jsamaha@wisc.edu

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.3925834>.

Correlated individual differences suggest a common mechanism underlying metacognition in visual perception and visual short-term memory

Jason Samaha¹ and Bradley R. Postle^{1,2}

¹Department of Psychology, and ²Department of Psychiatry, University of Wisconsin–Madison, Madison, WI 53706, USA

JS, 0000-0001-8010-5993

Adaptive behaviour depends on the ability to introspect accurately about one's own performance. Whether this metacognitive ability is supported by the same mechanisms across different tasks is unclear. We investigated the relationship between metacognition of visual perception and metacognition of visual short-term memory (VSTM). Experiments 1 and 2 required subjects to estimate the perceived or remembered orientation of a grating stimulus and rate their confidence. We observed strong positive correlations between individual differences in metacognitive accuracy between the two tasks. This relationship was not accounted for by individual differences in task performance or average confidence, and was present across two different metrics of metacognition and in both experiments. A model-based analysis of data from a third experiment showed that a cross-domain correlation only emerged when both tasks shared the same task-relevant stimulus feature. That is, metacognition for perception and VSTM were correlated when both tasks required orientation judgements, but not when the perceptual task was switched to require contrast judgements. In contrast with previous results comparing perception and long-term memory, which have largely provided evidence for domain-specific metacognitive processes, the current findings suggest that metacognition of visual perception and VSTM is supported by a domain-general metacognitive architecture, but only when both domains share the same task-relevant stimulus feature.

1. Introduction

When humans make decisions, they are capable of estimating the likelihood that their decision was accurate. Intuitively, this metacognitive ability is accurate when an individual's estimate of their decision accuracy (e.g. as expressed by a confidence rating) matches the actual accuracy of their decision [1]. Because decisions can be made on the basis of information from a plethora of sources—for example, deciding on the basis of current sensory input versus deciding on the basis of information culled from long-term memory—an outstanding question is whether metacognitive processes are domain-general or domain-specific [2]. A domain-general metacognitive monitoring process would be expected to evaluate the accuracy of decisions made from both perceptual inputs and those based on memory. By contrast, a domain-specific metacognitive system would use independent neural resources or computations to estimate the quality of memory- versus perception-based judgements, for example.

Recent work on this topic has focused on correlating individual differences in metacognition during perception and long-term memory, and has resulted in mixed findings. Several studies have reported non-significant relationships between individual's metacognitive ability in a perceptual task and their metacognitive ability in a long-term memory task [3–5], suggesting domain-specific metacognition. However, an experiment using similar tasks did find a reliable

positive correlation between metacognitive abilities in both domains [6], and other work has shown correlated metacognitive performance across different perceptual tasks [7]. A number of the above-mentioned studies, however, have also reported that structural and function brain imaging data from distinct regions correlated with metacognitive abilities for the distinct tasks [3,4,6], and a recent experiment showed that patients with anterior prefrontal cortex lesions have a selective deficit in visual perceptual metacognition, but not memory metacognition for a recently studied word list [8], reinforcing domain-specificity at the neural level.

A lack of cross-task correlation in metacognition may sometimes be difficult to interpret because this could result from procedural differences between tasks not necessarily related to the cognitive construct under investigation (e.g. the use of different stimuli in the perception versus memory task). Furthermore, perception and long-term memory are themselves quite distinct cognitive functions (although they can certainly interact in some situations, e.g. [9]), and an under-explored question is whether perceptual metacognition relates to metacognition for other cognitive functions more closely related to perception. Across three experiments, we examined whether metacognition in visual judgements is related to metacognition for visual short-term memory (VSTM) judgements using tasks with the same stimuli that differ only in the requirement for memory storage over a short delay (Experiments 1 and 2), or tasks that differ also in the relevant stimulus feature (Experiment 3). Because perception of and VSTM for a given stimulus feature are hypothesized to rely on shared neural representations [10–13], we might anticipate that metacognition in these domains is also based on some shared resource, leading to positively correlated individual differences in metacognition across tasks, but only when the task-relevant stimulus feature is shared.

2. Material and methods

(a) Experiments 1 and 2

Because of their similarities, the methods pertaining to Experiments 1 and 2 are described together in this section (figures 1–4), followed by the methods for Experiment 3.

(i) Participants

Forty subjects (20 in Experiment 1: mean age = 21 years, s.d. = 1.67, 10 female, and 20 in Experiment 2: mean age = 20.6 years, s.d. = 2.01, 14 female) from the University of Wisconsin–Madison community participated in these experiments and received monetary compensation. All subjects reported normal or corrected-to-normal visual acuity and colour vision, and were naive to the hypothesis of the experiment.

(ii) Perceptual task

To probe each individual's perceptual metacognitive abilities, we employed an orientation estimation task with confidence ratings [14]. On each trial, a target grating of threshold contrast embedded in noise (see electronic supplementary material, Methods) was presented centrally for 33 ms with a randomly determined orientation between 1° and 180°, followed shortly (600 ms) by a highly visible probe grating without noise, whose orientation could be rotated via mouse movement. This short interval between the target and probe was necessary to ensure that the probe had no visual masking effect. Subjects were instructed to match the perceived orientation as closely as possible. Subjects pressed the spacebar

to input their orientation response and then used number keys 1–4 to provide a confidence rating. Because performance in this task varies continuously (as opposed to a binary correct/incorrect outcome), we instructed subjects to use the confidence scale to indicate how close they think they came to the true orientation using the scale labels 1 ('complete guess') to 4 ('very close'). These perceptual task parameters were the same for both experiments. See figure 1a for complete trial timings.

(iii) Visual short-term memory task

To probe metacognitive abilities for VSTM, we introduced a delay period between the target and the response probe. In Experiment 1, the delay period was fixed at 7 s, and in Experiment 2 it was randomly sampled from the set: 3.45, 6.30, 9.15 or 12.00 s. The stimuli and all other task events were identical to the perceptual task.

(iv) Procedure

For Experiment 1, perceptual and working memory tasks were performed in separate blocks. Following an initial one-up, three-down staircase (see electronic supplementary material, Methods), stimulus contrast was fixed and each subject completed one block of 120 trials of the perceptual task, followed by three blocks of 60 trials each of the VSTM task, followed by another block of the perceptual task. This resulted in a total of 240 perceptual trials and 180 VSTM trials per subject, completed in a single 1.5 h session. Experiment 2 differed in that perceptual and VSTM trials were intermixed within blocks and randomly determined with equal probability to be either a perceptual trial or one of the four delay periods (between 3.45 and 12 s) of the VSTM task. Intermixing perception and VSTM trials further minimized procedural differences between tasks by eliminating any task-related expectations (because subjects did not know which type of trial would come next) and by removing temporal delays between task performance. Each subject completed 300 trials, separated into five blocks, resulting in an average ($\pm$ s.d.) of 55.5 (6.4) perceptual trials and 59 (8.0) trials of each delay period of the VSTM task, after removal of trials based on response times (see below). The total task time was approximately 1.5 h.

(v) Quantifying metacognition

To relate continuously varying task performance (the difference between the subject's response and the true orientation, in degrees) to subjective confidence ratings, we computed rank correlations between absolute error and confidence across trials to capture how well confidence tracks performance. Error should decrease with increasing confidence, so a subject with accurate metacognition would have a stronger negative correlation between confidence and error than a subject with poor metacognition. Although intuitive, this metric is potentially influenced by factors beyond metacognitive accuracy *per se*, such as task difficulty and biases in confidence scale use (e.g. under- or overconfidence [1]). Although we used a staircase procedure to match difficulty, there was still considerable variability across subjects in the median absolute error in both Experiment 1 (range: 8°–16.5°) and Experiment 2 (range: 6.9°–23.3°). A recently introduced measure called meta- d' / d' can correct for these influences [15]; however, meta- d' / d' has been developed only for tasks with discrete outcomes amenable to signal detection theory analysis (e.g. hits, misses) and cannot be applied to the continuous estimations tasks we employed in Experiments 1 and 2 (but see Experiment 3). To control for these influences when testing the relationship between perceptual and VSTM metacognition, we ran two additional multiple regression models that included covariates for average and task-specific error and confidence (see Statistics). In the case of models with multiple predictors, the relationship between perceptual and memory metacognition was visualized

Figure 1. Orientation estimation tasks and confidence–error relationships for Experiment 1. (a) On each trial of the perception and VSTM task, subjects moved a computer mouse to match the perceived or remembered orientation and then provided a confidence rating on a 1–4 scale to indicate how close they thought they came to the true orientation. The tasks differed only by the addition of a 7 s delay period for the VSTM task. (b) Distributions of responses relative to the true orientation (i.e. error) show a clear scaling with confidence ratings, suggesting that subjective ratings track objective performance at the group level. (c) Median absolute error scales with confidence and VSTM trials produced overall greater error, indicating that representations became noisier when held in VSTM. ITI, inter-trial interval. Shaded bands and error bars denote ± 1 s.e.m.

(figures 2 and 4) using added variable plots (MATLAB function *plotAdded.m*), which use the Frisch–Waugh–Lovell theorem to partial out the effects of other predictors in the model, revealing the effect of a single predictor while all other predictors are held constant. Predictor R^2 for these models was computed as the sum of squares for the perceptual metacognition predictor divided by the total sum of squares for all other predictors and error.

Additionally, we verified that the results of this analysis were robust to our particular metric of metacognition by repeating all analyses using the non-parametric area under the type 2 receiver operating characteristics curve (A_{ROC} [16–18]) as our measure of metacognitive accuracy. This measure is obtained by taking the area under the curve formed by plotting the type 2 false alarm rate by the type 2 hit rate at different type 2 criteria. A type 2 false alarm is an incorrect but high-confidence trial and type 2 hit is a correct and high-confidence trial, and the number of confidence criteria is the number of ratings on the scale minus 1. At values of 0.5, this metric indicates that confidence ratings do not discriminate between correct and incorrect trials, and values of 1 indicate perfect discriminability. A_{ROC} was computed using the method outlined in [16]. Because this metric requires binarizing the data into correct and incorrect responses, we defined thresholds for each subject based on the 75th percentile of their response error distributions such that a trial with error larger than this threshold was considered incorrect. This analytically set performance at 75% for each subject, equating accuracy for this analysis. Using a common threshold of 25° for each subject did not change the statistical significance of any analyses reported with this metric. Prior to any analysis, trials with response times below 200 ms or above the 95th percentile of the distribution of response times across all subjects were excluded. The same trial exclusion procedure was applied to both experiments.

(vi) Statistics

We used linear regression to predict individual differences in VSTM metacognition from variation in perceptual metacognition scores (figures 2 and 4). In a first, ‘basic model’, we considered only these two variables. Then, to control for individual differences in task performance and confidence ratings, we ran two additional regression models. One included each subject’s mean error and mean confidence as covariates (three predictors in total) and the other included task-specific confidence and error as covariates (i.e. mean perceptual error and confidence and mean VSTM error and confidence; five predictors in total). These three models were run for each metric of metacognition (r values and A_{ROC} ; see above) and for both experiments. To test for linear effects of confidence on error (figures 1c and 3d), we regressed single-trial confidence ratings on absolute error for each subject and task and tested the resulting slopes against zero at the group level using a t -test. To test for performance differences between tasks, we compared median absolute error between the perception and VSTM task with a paired t -test. We additionally tested for a linear effect of delay period duration in Experiment 2 (figure 3b) by fitting slopes to each subject’s single-trial absolute error by delay period data and testing these slopes against zero at the group level with a t -test. All tests were two-tailed.

(b) Experiment 3

This experiment was conducted to test whether the correlation between perceptual and VSTM metacognition depended on both tasks sharing the same task-relevant stimulus feature (e.g. orientation). To this end, we compared metacognition in an orientation perception task and a contrast perception task

Figure 2. Positive relationship between perceptual and VSTM metacognition in Experiment 1. (a) Cross-task regression using confidence–error correlations as the metric of metacognition. Increasingly complex regression models controlling for task performance and confidence shown from left to right (see Methods). (b) The same models as in (a), but using the area under the type 2 ROC curve (A_{ROC}) as a measure of metacognitive performance. Dashed lines denoted 95% confidence intervals on the linear fit. Black points are individual subjects.

(figure 5a) to metacognition in an orientation VSTM task. We used two-choice discrimination tasks in Experiment 3 which are amenable to a recently developed Bayesian hierarchical model-based analysis of metacognition that controls for individual differences in task performance and confidence biases while appropriately accounting for variability in individual subject parameter estimates at the group level [19].

(i) Participants

Twenty subjects (mean age = 21.8 years, s.d. = 3.18, 13 female) from the University of Wisconsin–Madison community were recruited.

(ii) Contrast perception task

Subjects were instructed to indicate whether the left or right stimulus contained a higher contrast grating (see electronic supplementary material, Methods) using the left and right arrow keys, respectively. Subjects then indicated their confidence using number keys 1–4, where 1 denotes a ‘complete guess’ and 4 denotes ‘very confident’. A target and a standard stimulus were presented simultaneously to the left and right of fixation for 50 ms. The location containing the target was randomly determined on each trial. The standard stimulus was created by averaging a 10% Michelson contrast grating with an 80% contrast noise patch and the contrast of the target grating was adapted for each subject with a staircase procedure (see electronic supplementary material, Methods). Each stimulus also had a randomly and independently determined orientation that was task-irrelevant.

(iii) Orientation perception task

This task required subjects to indicate whether the two gratings had the same or different orientation. Both stimuli appeared

simultaneously for 50 ms and then subjects indicated their choice followed by their confidence. On the ‘same’ trials, both stimuli had an identical orientation, which was randomly determined on each trial (between 1° and 180°), whereas on ‘different’ trials, one stimulus was offset by 25° clockwise or counter-clockwise (randomly determined). Whether a trial was the same or different was randomly determined. Difficulty was controlled by a staircase procedure that adapted the contrast of both stimuli (see electronic supplementary material, Methods).

(iv) Orientation visual short-term memory task

This task also required that subjects indicate whether two gratings were the same or different, but here there was a temporal delay of 3 s between the first and second stimulus. Thus, the orientation of the first stimulus must be maintained over the delay in order to perform the task. Each grating was presented for 50 ms, and subjects provided their choice and confidence following the second ‘probe’ stimulus. As in the orientation perception task, both stimuli had an identical randomly determined orientation on same trials, and a difference in orientation of 25° (clockwise or counter-clockwise) on different trials. Trial type was randomly determined, as was the location (left or right of fixation) of both stimuli, although the location of both stimuli was always the same for any given trial.

(v) Procedure

Each subject completed three blocks of 100 trials for each of the three tasks, resulting in 300 trials per task (with the exception of one subject who only completed 100 trials of the contrast perception task). Blocks of the same task were completed sequentially and task order was randomized. Prior to the start of each new

Figure 3. Task and behaviour for Experiment 2. (a) Perceptual trials (delay 0.6 s) and VSTM trials (delay between 3.45 and 12 s) were intermixed within blocks. (b) Error increased with increasing delay length, indicating a loss of information when the orientation needed to be maintained in VSTM. (c) Response error distributions show a clear scaling with confidence. (d) Error decreased as confidence increased in both perceptual and VSTM trials. Shaded bands and error bars indicate ± 1 s.e.m.

task, subjects completed 60 trials of an initial one-up, three-down staircase corresponding to that task (see electronic supplementary material, Methods). These 60 trials were not included in any analysis. The staircase was then continuously run throughout each task. The total task time was approximately 1.5 h.

(vi) Model-based analysis of metacognition

We used an estimate of metacognitive efficiency (meta- d' / d' , or M -ratio) that quantifies the extent to which confidence ratings discriminate between correct and incorrect decisions (i.e. type 2 performance), given the underlying difficulty of the discrimination itself (i.e. type 1 performance), thereby optimally controlling for task difficulty and confidence biases [1]. M -ratio is the ratio between the d' estimated from the confidence data according to a metacognitively ideal observer and the actual d' computed from task performance. Because both metrics are in the same units, M -ratio will approach 1 if all the information used for the type 1 decision was also available to the type 2 decision, indicating optimal metacognition. Values below 1 reflect suboptimal metacognition.

We used a recently introduced hierarchical modelling approach to estimate the cross-task correlation between individual differences in M -ratio, as is implemented in the freely available toolbox HMeta-d [19] for MATLAB. This toolbox is a hierarchical Bayesian extension of Maniscalco & Lau's [15] meta- d' model. The advantage of a Bayesian model in this context is that the estimation of group-level parameters of interest (i.e. M -ratio correlation coefficient across tasks) takes into account parameter uncertainty at the single subject level. This means that a subject whose M -ratio is estimated with high uncertainty will contribute less to the group-level parameter estimate than a subject whose M -ratio is estimated more precisely. In typical maximum-likelihood or sum-of-squares fitting [18], this knowledge of parameter uncertainty is discarded. Simulations suggest this approach produces more accurate parameter

recovery and lower false-positive rates than non-Bayesian alternatives [19]. Cross-task M -ratio correlations were estimated using the HMeta-d function `fit_meta_d_mcmc_groupCorr.m`.

Posterior distributions of parameters were sampled using Markov chain Monte Carlo methods (MCMC) implemented in JAGS (<http://mcmc-jags.sourceforge.net>). We ran three chains of 20 000 samples each with 5000 burn-in samples. Each subject's log(M -ratio) for each domain are specified as draws from a bivariate Gaussian. We used a weakly informative normal prior on log(M -ratio) which encompasses estimates from 167 previous subjects [19], and a uniform prior between -1 and 1 for the correlation coefficient. To assess convergence, we ensured that all MCMC chains were well mixed and that the Gelman and Rubins R -hat convergence statistics were between 1 and 1.1. Statistical significance for each correlation was assessed by computing the proportion of MCMC samples that fell below zero, multiplied by 2 (akin to a two-tailed non-parametric frequentist test) and by computing 95% high-density intervals (HDI) on the correlation posterior distributions.

3. Results

(a) Experiment 1

Distributions of response error as a function of confidence are shown in figure 1b. Absolute error significantly decreased with increasing confidence for both the perceptual task ($t_{19} = -13.48$, $p < 0.0001$) and the VSTM task ($t_{19} = -14.88$, $p < 0.0001$), indicating that subjects' confidence reasonably reflected their task performance at the group level (figure 1c). Error was also significantly greater in the VSTM task when compared with the perceptual task ($t_{19} = -2.10$, $p = 0.049$), reflecting an expected degradation of

Figure 4. Replication of the positive relationship between perceptual and VSTM metacognition in Experiment 2. (a) The same regression models as in figure 2, indicating the cross-task relationship using confidence–error correlations as the metric of metacognition. (b) The same as in (a), but with A_{ROC} as the metric of metacognition. Dashed lines denoted 95% confidence intervals on the linear fit.

orientation information when the task required short-term memory maintenance. Confidence ratings were distributed similarly for perception and VSTM tasks (electronic supplementary material, figure S1), and average confidence ratings did not significantly differ between tasks ($p = 0.15$). See electronic supplementary material, figure S2 for a breakdown of accuracy and response time by block.

Central to our hypothesis, we found a positive relationship across individuals between perceptual metacognition and VSTM metacognition (figure 2). This relationship was observed when using confidence–error correlations as the measure of metacognition (slope = 0.51, $t = 2.63$, predictor $R^2 = 0.28$, $p = 0.016$; figure 2a) and, importantly, was still present after controlling for average confidence and error (slope = 0.54, $t = 2.55$, predictor $R^2 = 0.28$, $p = 0.021$) and in the model controlling for task-specific confidence and error (slope = 0.62, $t = 2.66$, predictor $R^2 = 0.29$, $p = 0.018$). All covariate predictors in both control models were not statistically significant ($ps > 0.30$). These results indicate that, although the confidence–error correlation may be influenced by task performance and confidence biases, these factors did not account for the across-subject correlation between perceptual and VSTM metacognition.

The same relationship was observed when using A_{ROC} as the metric of metacognition (figure 2b). With the basic model, perceptual metacognition significantly predicted VSTM metacognition (slope = 0.77, $t = 3.96$, predictor $R^2 = 0.46$, $p = 0.0009$). This relationship held when controlling for average confidence and error (slope = 0.82, $t = 3.78$, predictor $R^2 = 0.46$, $p = 0.0016$) and when controlling for task-specific confidence and error (slope = 0.88, $t = 3.85$, predictor $R^2 =$

0.45, $p = 0.0017$). As before, all other covariate predictors across both control models were non-significant ($p > 0.26$). We examined the correlation between all predictor variables in all of our models (table 1) and found that there was collinearity between several (quite expected) covariate predictors (e.g. average confidence predicted average error, perceptual confidence predicted VSTM confidence). Importantly, however, there were no significant correlations between our predictors of interest (both perceptual metacognitive scores, A_{ROC} or ρ) and any other covariate predictors, indicating that task performance and confidence are unlikely to be driving the cross-task correlation in metacognition. These results indicate that the relationship observed between perceptual and VSTM metacognition was independent of the particular metric used and was not accounted for by correlated individual differences in task performance or average confidence.

(b) Experiment 2

This experiment served to replicate the cross-task correlation observed in Experiment 1 while further minimizing procedural differences between tasks by intermixing perceptual and VSTM trials of differing delays (figure 3a). Error increased monotonically with delay duration ($t_{19} = 2.85$, $p = 0.010$; figure 3b), and perception trials had lower error than VSTM trials, collapsing across delays ($t_{19} = 3.33$, $p = 0.003$), indicating the expected loss of information in VSTM relative to perception. As in Experiment 1, error decreased with increasing confidence during both perception ($t_{19} = -7.56$, $p < 0.0001$) and VSTM trials ($t_{19} = -8.99$, $p < 0.0001$), indicating that confidence reliably tracked performance at the group

Figure 5. Tasks, behaviour and metacognitive correlations from Experiment 3. (a) To compare metacognition across discrimination tasks while varying the task-relevant stimulus feature, subjects performed (i) a contrast perception task, judging which stimulus contained a higher contrast grating, (ii) an orientation perception task, judging whether the two stimuli had the same orientation, and (iii) an orientation VSTM task, judging whether a memorized target grating had the same orientation as a probe grating that appeared 3 s later. (b) Individual subject accuracies (proportion correct) and estimates of metacognitive efficiency (M -ratio) for each task. Note that because M -ratio for each subject is estimated in the same model, estimates are not fully independent [19]. (c) Posterior distributions of cross-task correlations in metacognitive efficiency. This reveals a strong positive correlation between orientation perception and orientation VSTM metacognition, but not between other tasks.

level (figure 3c,d). Average confidence was lower on VSTM (mean = 2.59, s.e.m. = 0.133) than on perception trials (mean = 2.71, s.e.m. = 0.132; $t_{19} = 2.97$, $p = 0.0077$; electronic supplementary material, figure S1).

Importantly, we replicated the positive relationship between perceptual and VSTM metacognition with quantitatively better model fits in a new set of subjects. Using confidence–error correlations (figure 4a), perceptual metacognition robustly predicted VSTM metacognition in the one-predictor basic model (slope = 0.60, $t = 5.21$ predictor $R^2 = 0.60$, $p < 0.0001$), the three-predictor model controlling for average confidence and error (slope = 0.60, $t = 4.91$, predictor $R^2 = 0.59$, $p = 0.0001$) and in the five-predictor model controlling for task-specific confidence and error (slope = 0.59, $t = 4.44$, predictor $R^2 = 0.58$, $p = 0.0005$). All covariate predictors in both control models were non-significant ($p > 0.64$). As in Experiment 1, some covariate predictors were significantly correlated (table 1; italicized region), but no covariates were significantly correlated with perceptual metacognition, the

predictor of interest. This effect was also observed when using A_{ROC} as the metric of metacognition for the basic model (slope = 0.53, $t = 3.88$, predictor $R^2 = 0.45$, $p = 0.0011$), the three-predictor model (slope = 0.52, $t = 3.58$, predictor $R^2 = 0.44$, $p = 0.002$) and the five-predictor model (slope = 0.54, $t = 3.43$, predictor $R^2 = 0.44$, $p = 0.004$). All covariates in both control models were non-significant ($ps > 0.52$).

(c) Experiment 3

Task accuracy (% correct) and metacognitive efficiency (M -ratio) for each task and subject are shown in figure 5b. Accuracy was comparable between the contrast perception task (mean = 78.6%, s.e.m. = 0.012) and the orientation perception task (mean = 80.4%, s.e.m. = 0.009; $p = 0.33$), as well as between the contrast perception task and the orientation VSTM task (mean = 76.5%, s.e.m. = 0.018%, $p = 0.39$), but differed significantly between the orientation perception and the orientation VSTM tasks ($t_{19} = 2.68$,

Table 1. Correlation matrix for every predictor in each model in Experiments 1 and 2 (italicized regions). X's denote predictor combinations that were not used in any model. Significant correlations ($p < 0.05$) are noted in bold and with asterisks.

predictor—Exp. 1, Exp. 2	1	2	3	4	5	6	7	8
1. perceptual metacognition (ρ)	—	−0.13	−0.03	−0.16	−0.04	−0.10	−0.01	X
2. average confidence	0.33	—	−0.41	X	X	X	X	0.15
3. average error	0.01	−0.54*	—	X	X	X	X	0.01
4. VSTM confidence	0.33	X	X	—	−0.42	0.95*	−0.38	0.17
5. VSTM error	0.08	X	X	−0.54*	—	−0.42	0.91*	−0.01
6. perceptual confidence	0.32	X	X	0.94*	−0.44*	—	−0.38	0.13
7. perceptual error	−0.05	X	X	−0.53*	0.83	−0.53*	—	0.03
8. perceptual metacognition (A_{ROC})	X	−0.29	−0.07	−0.28	−0.01	−0.14	−0.29	—

$p = 0.015$). This is partly because four subjects reached the maximum contrast allowable by the VSTM staircase (25%), so the VSTM task never got easier for them. In general, higher target contrasts were needed in the VSTM task (mean = 15.5%, s.e.m. = 0.012) compared with both the orientation perception task (mean = 8.5%, s.e.m. = 0.005; $t_{19} = 6.50$, $p < 0.0001$) and the contrast perception task (mean = 6.2%, s.e.m. = 0.005; $t_{19} = 6.49$, $p < 0.0001$), indicating that more signal was needed in the VSTM task to achieve threshold performance. Metacognitive efficiency, on the other hand, did not significantly differ between tasks ($ps > 0.16$; see electronic supplementary material, figure S3), and was well below the optimal M -ratio of 1 for all tasks (contrast perception: mean M -ratio = 0.4, HDI = (0.27, 0.54), orientation perception: mean = 0.44, HDI = (0.30, 0.59), orientation VSTM: mean = 0.53, HDI = (0.38, 0.68)). Mean confidence (see electronic supplementary material, figure S1) also did not differ between either of the perceptual tasks and the VSTM task ($p > 0.08$), but was significantly lower in the contrast perception task when compared with the orientation perception task ($t_{19} = 3.06$, $p = 0.006$).

To our primary question of correlated individual differences in metacognition, we observed a large positive correlation between metacognitive efficiency in the orientation perception task and orientation VSTM task (figure 5c; $\rho = 0.90$, HDI = (0.58, 0.99), $p < 0.0001$), consistent with the results of Experiments 1 and 2. However, we did not observe a significant correlation between metacognition in the contrast perception task and the orientation VSTM task ($\rho = 0.32$, HDI = (−0.32, 0.85), $p = 0.32$), nor between the two perceptual tasks ($\rho = 0.24$, HDI = (−0.35, 0.76), $p = 0.42$), suggesting that a cross-task correlation depends on both tasks sharing the same relevant stimulus feature. We also computed the difference between these correlation distributions and found that the correlation between orientation perception and orientation VSTM memory was significantly larger than the correlation between contrast perception and orientation perception (HDI = (0.07, 1.27), $p = 0.027$), and trending larger than the correlation between contrast perception and orientation VSTM (HDI = (−0.03, 1.24), $p = 0.063$).

4. Discussion

Metacognition is an important aspect of decision-making [20,21], learning [22], development [23] and perhaps certain

aspects of conscious experience [24,25], and can be compromised in psychiatric disorders [26–28]. It is currently unclear whether an individual with accurate metacognitive abilities in one domain also has accurate metacognition in other domains. In Experiments 1 and 2, we found that individuals with more accurate metacognition in perceptual judgements also showed more accurate metacognition in a VSTM task requiring stimulus maintenance over a brief delay period. A medium to high correlation was present when using two different measures of metacognitive performance, across both experiments, and when controlling for task performance and mean confidence. In Experiment 3, we compared an orientation VSTM task to an orientation perception task and a contrast perception task. We again found a large positive correlation ($R^2 = 0.81$) between metacognition in the orientation VSTM and orientation perception task, but not between the orientation VSTM and contrast perception task, nor between contrast perception and orientation perception tasks, highlighting the importance of both tasks sharing the same relevant stimulus feature. Importantly, given known biases in VSTM metacognitive judgements [29], metacognition in Experiment 3 was quantified within a signal detection theory model [15,19] that controls for confidence biases and task performance. Taken together, these results provide the first evidence in humans for a medium to high positive correlation between an individual's metacognitive abilities in perception and VSTM, when both domains share a common stimulus feature.

The present results contrast with recent experiments examining the relationship between metacognition of visual perception and long-term memory, which have typically observed no correlation [3–5] (but see [6]). We reason that, in contrast with long-term memory, VSTM for a given stimulus feature is thought to rely on the same neural representations that support perception of that stimulus feature [10–13], and this may underlie the cross-task correlation in metacognitive performance. This explanation follows naturally from 'first-order' models of metacognition according to which confidence and task performance are driven by the same internal representation of stimulus evidence [30–33]. For example, in signal detection theoretic models, the absolute distance of the decision variable from the decision criterion is a proxy for confidence [34,35]. Thus, if perception and VSTM were supported by the same internal representation of the stimulus, then the computation of confidence across the two tasks would also be based on the same representations, leading to correlated

behaviour. ‘Second-order’ models of metacognition, by contrast, posit an architecture with a secondary confidence read-out process, which may be influenced by additional sources of noise [36] or other signals not directly related to the stimulus, such as action-related states [37,38], cortical excitability [39] or arousal [40,41].

The results of Experiments 1 and 2 are also compatible with second-order models of metacognition, although several possible relationships between first- and second-order processes could explain our findings. Shared first-order (sensory) representations across tasks might be enough to produce a behavioural correlation despite separate second-order read-out mechanisms. Alternatively, both first- and second-order processes may be shared across tasks, or only the second-order process shared, though this latter possibility is unlikely, given existing neural evidence for shared representations in visual regions across perception and VSTM [13,42,43]. The results of Experiment 3, however, provide support for first-order models because they suggest that shared sensory representations underlie the cross-task correlation in metacognition. Because metacognition was not reliably correlated when tasks differed in their relevant stimulus feature, even when both tasks were perceptual, this points towards a first-order model of metacognition. Yet another alternative is that the correlation was dependent on the task structure, for example, because both orientation tasks involved same/different judgements. This account may also explain why a previous report comparing metacognition for contrast and orientation judgements in the context of a visual search paradigm did find correlated individual differences [7], but recent work comparing a variety of perceptual paradigms with different task structures and stimuli did not find a correlation [44].

Although the present findings are consistent with some domain-general processing in metacognition for perception and VSTM, correlations at the behavioural level raise further questions about what specific aspects of metacognitive processing are shared. For example, if one’s ability to learn stable confidence criteria over time improves metacognitive accuracy [33], then metacognitive abilities may be high across domains for an individual with superior learning abilities, perhaps related to recent work implicating hippocampal myelination in perceptual metacognition [45]. However, this

need not imply that the underlying neural substrate responsible for computing the appropriate levels of confidence is itself domain-general. Similarly, recent work has highlighted specific factors beyond stimulus evidence that modulate confidence, leading to dissociations of confidence and task performance within an individual [46–48]. For example, spontaneous trial-to-trial fluctuations in oscillatory neural activity in the α -band (8–13 Hz), which is thought to reflect visual cortical excitability [49,50], have been shown to bias confidence ratings, but not objective performance in a visual discrimination task [39]. Perhaps a subject who is less susceptible to such influences from sources not directly related to the difficulty of stimulus discrimination would show better metacognition across different domains. Future work examining neural correlates of metacognitive performance across different domains may contribute in a substantive way to this issue. As an example, McCurdy *et al.* [6] observed a positive correlation between metacognition of perception and recollection memory at the behavioural level, but found distinct (as well as overlapping) neural structures whose grey matter volume related to metacognitive performance in the different tasks. This suggests that only a portion of the processing stages or computations involved in generating confidence need be shared across tasks in order to produce a behavioural correlation. Nevertheless, the experiments reported here provide an important first step for future work by demonstrating a clear correlation between metacognitive behaviour in perception and VSTM.

Ethics. All subjects provided written consent and all experiments were approved by the University of Wisconsin–Madison Institutional Review Board.

Data accessibility. In accordance with the practices of open science and reproducibility, all raw data and code used in the present analyses are freely available through the Open Science Framework (<https://osf.io/py38c/>) [51].

Authors’ contributions. J.S. and B.R.P. designed the experiments. J.S. collected and analysed data. J.S. and B.R.P. wrote the manuscript.

Competing interests. We declare we have no competing interests.

Funding. This study was supported by National Institute of Mental Health (MH095984) to B.R.P.

Acknowledgements. We thank John B. Barrett, Missy Switzky and Sawyer Cimaroli for help with data collection, and two anonymous reviewers for their helpful feedback.

References

1. Fleming SM, Lau HC. 2014 How to measure metacognition. *Front. Hum. Neurosci.* **8**, 16657. (doi:10.3389/fnhum.2014.00443)
2. Kelemen WL, Frost PJ, Weaver CA. 2000 Individual differences in metacognition: evidence against a general metacognitive ability. *Mem. Cognit.* **28**, 92–107. (doi:10.3758/BF03211579)
3. Baird B, Cieslak M, Smallwood J, Grafton ST, Schooler JW. 2015 Regional white matter variation associated with domain-specific metacognitive accuracy. *J. Cogn. Neurosci.* **27**, 440–452. (doi:10.1162/jocn_a_00741)
4. Baird B, Smallwood J, Gorgolewski KJ, Margulies DS. 2013 Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *J. Neurosci.* **33**, 16 657–16 665. (doi:10.1523/JNEUROSCI.0786-13.2013)
5. Fitzgerald LM, Arvanah M, Dockree PM. 2017 Domain-specific and domain-general processes underlying metacognitive judgments. *Conscious Cogn.* **49**, 264–277. (doi:10.1016/j.concog.2017.01.011)
6. McCurdy LY, Maniscalco B, Metcalfe J, Liu KY, de Lange FP, Lau H. 2013 Anatomical coupling between distinct metacognitive systems for memory and visual perception. *J. Neurosci.* **33**, 1897–1906. (doi:10.1523/JNEUROSCI.1890-12.2013)
7. 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. (doi:10.1016/j.concog.2010.12.011)
8. Fleming SM, Ryu J, Golfinos JG, Blackmon KE. 2014 Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain* **137**, 2811–2822. (doi:10.1093/brain/awu221)
9. Fan JE, Hutchinson JB, Turk-Browne NB. 2016 When past is present: substitutions of long-term memory for sensory evidence in perceptual judgments. *J. Vis.* **16**, 1. (doi:10.1167/16.8.1)
10. D’Esposito M, Postle BR. 2015 The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* **66**, 115–142. (doi:10.1146/annurev-psych-010814-015031)

11. Harrison SA, Tong F. 2009 Decoding reveals the contents of visual working memory in early visual areas. *Nature* **458**, 632–635. (doi:10.1038/nature07832)
12. Postle BR. 2006 Working memory as an emergent property of the mind and brain. *Neuroscience* **139**, 23–38. (doi:10.1016/j.neuroscience.2005.06.005)
13. Serences JT, Ester EF, Vogel EK, Awh E. 2009 Stimulus-specific delay activity in human primary visual cortex. *Psychol. Sci.* **20**, 207–214. (doi:10.1111/j.1467-9280.2009.02276.x)
14. Rademaker RL, Tredway CH, Tong F. 2012 Introspective judgments predict the precision and likelihood of successful maintenance of visual working memory. *J. Vis.* **12**, 21.
15. Maniscalco B, Lau H. 2012 A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Conscious Cogn.* **21**, 422–430. (doi:10.1016/j.concog.2011.09.021)
16. Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G. 2010 Relating introspective accuracy to individual differences in brain structure. *Science* **329**, 1541–1543. (doi:10.1126/science.1191883)
17. 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. (doi:10.3758/BF03196546)
18. Kornbrot DE. 2006 Signal detection theory, the approach of choice: model-based and distribution-free measures and evaluation. *Percept. Psychophys.* **68**, 393–414. (doi:10.3758/BF03193685)
19. Fleming SM. 2017 HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neurosci. Conscious* **3**, nix007. (doi:10.1093/nc/nix007)
20. Kiani R, Corthell L, Shadlen MN. 2014 Choice certainty is informed by both evidence and decision time. *Neuron* **84**, 1329–1342. (doi:10.1016/j.neuron.2014.12.015)
21. van den Berg R, Zylberberg A, Kiani R, Shadlen MN, Wolpert DM. 2016 Confidence is the bridge between multi-stage decisions. *Curr. Biol.* **26**, 3157–3168. (doi:10.1016/j.cub.2016.10.021)
22. Butler AC, Karpicke JD, Roediger III HL. 2008 Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. *J. Exp. Psychol. Learn. Mem. Cogn.* **34**, 918–928. (doi:10.1037/0278-7393.34.4.918)
23. Weil LG, Fleming SM, Dumontheil I, Kilford EJ, Weil RS, Rees G, Dolan RJ, Blakemore S-J. 2013 The development of metacognitive ability in adolescence. *Conscious Cogn.* **22**, 264–271. (doi:10.1016/j.concog.2013.01.004)
24. Lau H, Rosenthal D. 2011 Empirical support for higher-order theories of conscious awareness. *Trends Cogn. Sci.* **15**, 365–373. (doi:10.1016/j.tics.2011.05.009)
25. Samaha J. 2015 How best to study the function of consciousness? *Front. Psychol.* **6**, 604. (doi:10.3389/fpsyg.2015.00604)
26. David AS, Bedford N, Wiffen B, Gilleen J. 2012 Failures of metacognition and lack of insight in neuropsychiatric disorders. *Phil. Trans. R. Soc. B* **367**, 1379–1390. (doi:10.1098/rstb.2012.0002)
27. Hauser TU, Allen M, Rees G, Dolan RJ. 2017 Metacognitive impairments extend perceptual decision making weaknesses in compulsivity. *Sci. Rep.* **7**, 906. (doi:10.1038/s41598-017-06116-z)
28. Bliksted V, Samuelsen E, Sandberg K, Bibby BM, Overgaard MS. 2017 Discriminating between first- and second-order cognition in first-episode paranoid schizophrenia. *Cognit. Neuropsychiatry* **22**, 95–107. (doi:10.1080/13546805.2016.1268954)
29. Adam KCS, Vogel EK. 2017 Confident failures: lapses of working memory reveal a metacognitive blind spot. *Atten. Percept. Psychophys.* **79**, 1506–1523. (doi:10.3758/s13414-017-1331-8)
30. Kepecs A, Uchida N, Zariwala HA, Mainen ZF. 2008 Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231. (doi:10.1038/nature07200)
31. Kiani R, Shadlen MN. 2009 Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* **324**, 759–764. (doi:10.1126/science.1169405)
32. Sanders JL, Hangya B, Kepecs A. 2016 Signatures of a statistical computation in the human sense of confidence. *Neuron* **90**, 499–506. (doi:10.1016/j.neuron.2016.03.025)
33. Treisman M, Faulkner A. 1984 The setting and maintenance of criteria representing levels of confidence. *J. Exp. Psychol. Hum. Percept. Perform.* **10**, 119–139. (doi:10.1037/0096-1523.10.1.119)
34. Macmillan NA, Creelman CD. 2004 *Detection theory: a user's guide*, 2nd edn. Mahwah, NJ: Lawrence Erlbaum Associates.
35. Cartwright D, Festinger L. 1943 A quantitative theory of decision. *Psychol. Rev.* **50**, 595–621. (doi:10.1037/h0056982)
36. Maniscalco B, Lau H. 2016 The signal processing architecture underlying subjective reports of sensory awareness. *Neurosci. Conscious* **2016**, niw002. (doi:10.1093/nc/niw002)
37. Fleming SM, Maniscalco B, Ko Y, Amendi N, Ro T, Lau H. 2015 Action-specific disruption of perceptual confidence. *Psychol. Sci.* **26**, 89–98. (doi:10.1177/0956797614557697)
38. Fleming SM, Daw ND. 2017 Self-evaluation of decision-making: a general bayesian framework for metacognitive computation. *Psychol. Rev.* **124**, 91–114. (doi:10.1037/rev0000045)
39. Samaha J, Iemi L, Postle BR. 2017 Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy. *Conscious Cogn.* **54**, 47–55. (doi:10.1016/j.concog.2017.02.005)
40. Allen M, Frank D, Schwarzkopf DS, Fardo F, Winston JS, Hauser TU, Rees G. 2016 Unexpected arousal modulates the influence of sensory noise on confidence. *eLife* **5**, e18103.
41. Hauser TU, Allen M, Purg N, Moutoussis M, Rees G, Dolan RJ. 2017 Noradrenaline blockade specifically enhances metacognitive performance. *eLife* **6**, e24901. (doi:10.7554/eLife.24901)
42. Riggall AC, Postle BR. 2012 the relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J. Neurosci.* **32**, 12 990–12 998. (doi:10.1523/JNEUROSCI.1892-12.2012)
43. Emrich SM, Riggall AC, LaRocque JJ, Postle BR. 2013 Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *J. Neurosci.* **33**, 6516–6523. (doi:10.1523/JNEUROSCI.5732-12.2013)
44. Ais J, Zylberberg A, Barttfeld P, Sigman M. 2016 Individual consistency in the accuracy and distribution of confidence judgments. *Cognition* **146**, 377–386. (doi:10.1016/j.cognition.2015.10.006)
45. Allen M, Glen JC, Müllensiefen D, Schwarzkopf DS, Fardo F, Frank D, Callaghan MF, Rees G. 2017 Metacognitive ability correlates with hippocampal and prefrontal microstructure. *Neuroimage* **149**, 415–423. (doi:10.1016/j.neuroimage.2017.02.008)
46. Koizumi A, Maniscalco B, Lau H. 2015 Does perceptual confidence facilitate cognitive control? *Atten. Percept. Psychophys.* **77**, 1295–1306. (doi:10.3758/s13414-015-0843-3)
47. Samaha J, Barrett JJ, Sheldon AD, LaRocque JJ, Postle BR. 2016 Dissociating perceptual confidence from discrimination accuracy reveals no influence of metacognitive awareness on working memory. *Front. Psychol.* **7**, 851. (doi:10.3389/fpsyg.2016.00851)
48. Zylberberg A, Barttfeld P, Sigman M. 2012 The construction of confidence in a perceptual decision. *Front. Integr. Neurosci.* **6**, 79. (doi:10.3389/fnint.2012.00079)
49. Romei V, Brodbeck V, Michel C, Amedi A, Pascual-Leone A, Thut G. 2008 Spontaneous fluctuations in posterior α -band EEG activity reflect variability in excitability of human visual areas. *Cereb. Cortex* **18**, 2010–2018. (doi:10.1093/cercor/bhm229)
50. Samaha J, Gosseseries O, Postle BR. 2017 Distinct oscillatory frequencies underlie excitability of human occipital and parietal cortex. *J. Neurosci.* **37**, 2824–2833. (doi:10.1523/JNEUROSCI.3413-16.2017)
51. Samaha J, Postle BR. 2017 Data from: Correlated individual differences suggest a common mechanism underlying metacognition in visual perception and visual short-term memory. Open Science Framework. (<https://osf.io/py38c>)