


Preserved Metacognition for Undetected Visuomotor Deviations

Michael Pereira,¹ Rafal Skiba,^{2,3,4}  Yann Cojan,² Patrik Vuilleumier,² and Indrit Bègue^{2,5,6}

¹Laboratoire de Psychologie et NeuroCognition, Université Grenoble Alpes and Université Savoie Mont Blanc, Centre National de la Recherche Scientifique, 38000 Grenoble, France, ²Laboratory for Neurology and Imaging of Cognition, Department of Basic Neuroscience, University of Geneva, 1211 Geneva, Switzerland, ³BC Mental Health and Addictions Research Institute, Vancouver, British Columbia V5Z 4H4, Canada, ⁴Department of Psychiatry, University of British Columbia, Vancouver, British Columbia V6T 2A1, Canada, ⁵Adult Psychiatry Division, Department of Psychiatry, University Hospitals of Geneva, 1211 Geneva, Switzerland, and ⁶Laboratory for Clinical and Experimental Psychopathology, Department of Psychiatry, University of Geneva, 1211 Geneva, Switzerland

Humans can successfully correct deviations of movements without conscious detection of such deviations, suggesting limited awareness of movement details. We ask whether such limited awareness impairs confidence (metacognition). We recorded functional magnetic resonance imaging data while 31 human female and male participants detected cursor deviations during a visuomotor reaching task and rated their confidence retrospectively. We show that participants monitor a summary statistic of the unfolding visual feedback (the peak cursor error) to detect visuomotor deviations and adjust their confidence ratings, even when they report being unaware of a deviation. Crucially, confidence ratings were as metacognitively efficient for aware and unaware deviations. At the neural level, activity in the ventral striatum tracks high confidence, whereas a broad network encodes cursor error but not confidence. These findings challenge the notion of limited conscious action monitoring and uncover how humans monitor their movements as they unfold, even when unaware of ongoing deviations.

Key words: awareness; confidence; detection; fMRI; metacognition; visuomotor

Significance Statement

We are unaware of the small corrections we apply to our movements as long as our goals are achieved. Here, although we replicate the finding that participants deny perceiving small deviations they correct, we show that their confidence reliably reflects the presence or absence of a deviation. This observation shows they can metacognitively monitor the presence of a deviation, even when they deny perceiving it. We also describe the hemodynamic correlates of confidence ratings. Our study questions the extent to which humans are unaware of the details of their movements; describes a plausible mechanism for metacognition in a visuomotor task, along with its neural correlates; and has important implications for the construction of the sense of self.

Introduction

Whether reaching for popcorn while viewing a movie or biking while enjoying the scenery, humans rely on reciprocal intricate connections between vision and motor processing to perform efficient behavior. Such visuomotor loops seem to occur mostly in the absence of

awareness. Indeed, seminal work from Fournier and Jeannerod (1998) showed that participants are unaware of their true hand position under imposed visuomotor deviations, although they appropriately correct their movement trajectories. Humans also neglect small spatial incongruences in feedback about their own movements (Farrer et al., 2008) and can reach targets they cannot consciously report (Binsted et al., 2007) or that are displaced without being noticed (Goodale et al., 1986). These findings support the notion that participants show limited awareness of the details of an ongoing movement, as long as their goal is achieved (Blakemore et al., 2002; Custers and Aarts, 2010; Gaveau et al., 2014; Tsay et al., 2020).

Rather than awareness of movement details, other studies tested whether humans can monitor the accuracy of decisions based on movement details (Sinanaj et al., 2015; Arbuzova et al., 2020; Charles et al., 2020; Locke et al., 2020). This ability to monitor (and control) one's internal processes is referred to as metacognition (Flavell, 1979; Koriati, 2006). A standard measure of metacognition

Received Jan. 23, 2023; revised July 26, 2023; accepted July 26, 2023.

Author contributions: Y.C., P.V., and I.B. designed research; Y.C. and I.B. performed research; M.P., R.S., Y.C., and I.B. analyzed data; M.P. and I.B. wrote the paper.

This work was supported by Swiss National Science Foundation Grants P2ELP3_187974 and P400PM_199251 to M.P. and a Projets de Recherche et Développement (PRD 22-2020-I; University Hospital of Geneva) grant to I.B. We thank the technical staff of the imaging platform at the Brain and Behavior Lab at the University of Geneva, Christophe Mermoud for programming of the experimental protocol, and Nathan Favre for comments on the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Indrit Bègue at indrit.begue@unige.ch.

<https://doi.org/10.1523/JNEUROSCI.0133-23.2023>

Copyright © 2023 the authors

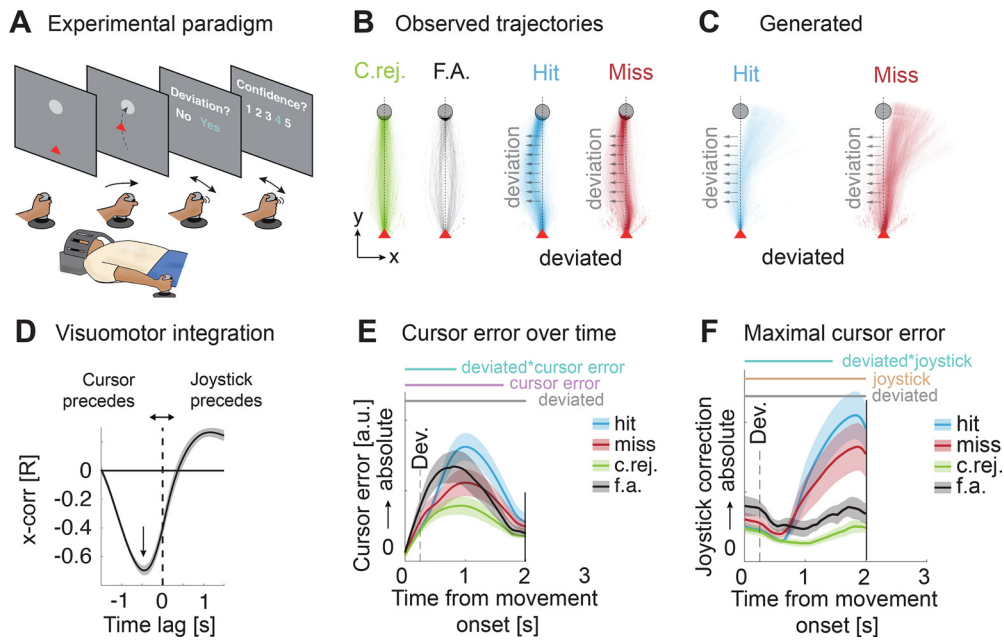


Figure 1. Experimental paradigm and detection responses. **A**, Participants started every trial by using the joystick to bring a visible cursor (red triangle) to a target (circle) on top of the screen. A deviation was applied to the cursor in 79% of the trials, titrated so as to reach 71% of detection accuracy. Titrated detection rates can be found in Extended Data Figure 1-1. After reaching the target, participants reported whether the cursor was deviated or not (detection) and how confident they were about their answer. **B**, Trajectories of the cursor for correct rejections (c. rej.; green), false alarms (f.a.; gray), hits (blue), and misses (red). Note that deviations could be toward the left or the right, but right deviation trials are mirrored and pooled with left trials for display purposes. **C**, Generated trajectories obtained by recomputing the cursor position had there been no deviation for hits (blue) and misses (red). **D**, Cross-correlation between cursor error and joystick lateral position. The vertical arrow indicates the time lag at which cross-correlation is strongest. **E**, Cursor error over time for different signal detection theory categories, hits (blue), misses (red), correct rejections (green), and false alarms (black). The significant ($p < 0.05$; false discovery rate corrected) main effect of cursor error over time is depicted by the purple line and the main effect of deviation by the gray line; and the significant interaction effects between deviated trials and cursor error is depicted by the cyan line. The dashed vertical line represents the onset of the experimental deviation (dev.). Model comparison, peak cursor error distributions and discriminability analyses can be found in Extended Data Figures 1-2, 1-3, 1-4, 1-5. **F**, Same as **E** for joystick correction over time. Shaded areas in **A–F** indicate 95% confidence intervals.

in humans is to ask them to make a decision (about a stimulus or an action) and subsequently rate their confidence in this decision (Rahnev et al., 2020). Together, these two lines of research suggest that although humans have limited awareness of their movement details, they can metacognitively monitor decisions about the details of these movements. As the discrimination tasks used previously (Arbuzova et al., 2020; Charles et al., 2020; Locke et al., 2020) were not specifically designed to contrast trials with and without awareness, it remains unknown whether awareness is necessary for metacognitive monitoring during motor control.

Indeed, during perceptual decision-making, participants are better at using perceptual information to scale their confidence about the presence of a stimulus than about the absence of a stimulus, showing that metacognitive efficiency is impaired when perceptual information is insufficient to reach awareness (Kanai et al., 2010; Meuwese et al., 2014; Mazor et al., 2020). It remains unresolved whether awareness has a similar effect on metacognitive efficiency in visuomotor decisions. In a previous study on metacognition of movement with a visuomotor detection task, participants reported being aware of only some of the experimental deviations in the visual feedback of their movements. Individual differences in metacognitive efficiency were associated with variability in brain gray matter volume in prefrontal cortex and insular and visual areas (Sinanaj et al., 2015). However, the degree to which metacognitive efficiency is related to awareness of visuomotor deviation and to specific movement parameters, as well as their functional neural underpinning, were not examined in our first study.

Here, we sought to explain how confidence is computed in the presence or absence of reported awareness, measured using

introspective subjective reports (Mudrik and Deouell, 2022), of a visuomotor deviation and unveil brain regions mediating this process using functional magnetic resonance imaging (fMRI) data. We could thus determine how visual information about participants' movements was integrated into their confidence ratings while dissociating trials for which they reported being aware of the deviation from those they reported not being aware of it. Finally, we identified brain regions whose activity correlated with confidence, detection response, and monitoring of performance during a visuomotor task.

Materials and Methods

Participants. We recruited 32 healthy right-handed participants based on convenience sampling. One participant did not complete the experimental task; therefore, the final sample included 31 participants of either sex (age, 26 ± 4.7 years). Participants gave written informed consent before the experiment and received 20 Swiss francs per hour as compensation. They had normal or corrected-to-normal vision and reported no neurologic or psychiatric disorder. The study was approved by the Ethics Committee of the University of Geneva and University Hospitals of Geneva (CER:11-214/NAC 11-077). All participants read and signed an informed consent form and were screened for contraindications to MRI with a standard safety questionnaire.

Experimental procedure. We asked participants to perform a visuomotor reaching task (Fournier and Jeannerod, 1998; Farrer et al., 2008). After a short preparation period (white triangle became red, duration 1–2 s, jittered) participants had to reach a centrally located target on the screen with a cursor moving at constant speed in a direction controlled by the joystick (Fig. 1A). The experimental manipulation (79% of the trials) consisted in introducing deviations in the mapping between the joystick and the cursor direction; the cursor direction was deviated by a

certain deviation angle, positive for right (clockwise) deviations and negative for left (counterclockwise) deviations. These deviations were gradually applied (0.3 s ramp), starting after participants reached a fixed distance from the starting point corresponding to 13% of the vertical distance between the initial cursor error and the final target. Participants were informed that these externally originating deviations would not occur all the time; however, when they occurred, participants had to correct for these deviations using the joystick to reach the target. After reaching the target, participants reported whether they noticed any externally originating deviations of their trajectory (yes detection responses) or whether they did not (no responses) and subsequently rated their confidence in their own judgment on a scale ranging from 1, not certain, to 5, completely certain (Fig. 1A). Participants did not receive feedback about the accuracy of either detection or confidence judgments. We encouraged participants to use the whole confidence scale. Participants selected yes or no responses as well as confidence ratings through joystick handle movements on the right and left, respectively, then pressing a button on the joystick.

We asked participants to perform a training run outside the scanner, consisting of 30 nondeviated trajectories to familiarize them with the joystick and experimental environment. For each experimental session, we ran an adaptive staircase procedure that made the task more difficult after two consecutive correct responses by decreasing the next deviation by 2.64° but made it easier after an incorrect response by increasing the next deviation by 1°. After the training session, participants entered the scanner and performed a threshold session of 80 trials to stabilize the staircase procedure. After the threshold session, participants completed two experimental runs. A structural T1 image was acquired between these two runs. Overall, there were 208 trials (21% without experimental deviations). Each trial lasted 11.5 s and was followed by a blank screen with a jittered duration (3–6 s).

Behavioral analyses. We defined a trial with an experimental deviation that was reported as such by participants as a hit and as a miss in case it was not reported. A trial was a correct rejection when there was no deviation, and participants correctly reported no deviation and a false alarm if participants reported a deviation. We grouped hits and false alarms into yes responses and misses and correct rejections into no responses. We computed the sensitivity d' and criterion c using signal detection theory. Cursor and joystick positions were defined every 10 ms. Cursor error was defined as the horizontal distance between the cursor position and the midline between the starting point (triangle at a lower central position on the screen) and the target (top central position). For each trial, we defined peak cursor error as the peak value of this cursor error over time. Onset error (onset err.) was defined as the cursor error at the onset of the deviation. We also defined the average position (avg. pos.) as the absolute value of the average position of the cursor with respect to the sagittal line (the latter can be negative) as well as the average cursor error (avg. err.) as the average of the distance between the cursor and midline (always positive). A trial with a large deviation to the right followed by a large deviation to the left would thus have an average position close to zero but a high average cursor error.

To build two-dimensional histograms of the cursor trajectories (Fig. 1B), we mirrored trajectories when the deviation was toward the right (only for hits and misses). We then computed two-dimensional histograms of the position of the cursor in all trials of a signal detection theory category (hit, miss, correct rejection and false alarms) and normalized the resulting histograms by the number of trials in that category. Finally, we averaged across participants. As the cursor at time k (c_k in the complex plane) can be defined knowing cursor position at time $k-1$ and the angle of the joystick handle at time k (α_k) in the following:

$$c_k = c_{k-1} + \rho \exp(j(\alpha_k + \delta)),$$

with $c_0 = 0 + j*0$, ρ a small real constant defining the speed and δ the angle of the experimental deviation, we could reconstruct the trajectories of the cursor, had there not been any deviation and construct histograms as described above for Figure 1B as follows:

$$\hat{c}_k = \hat{c}_{k-1} + \rho \exp(j(\alpha_k)).$$

For statistics, we defined binomial link mixed-effects models to analyze detection responses and cumulative link mixed-effects models for confidence ratings using the ordinal package in R software (<https://CRAN.R-project.org/package=ordinal>). Inclusion of random effects was guided by regression model selection based on Bayesian information criterion and led to the inclusion of all factors and interactions as random effects. Because most models are multivariate, we ensured that variance inflation factors that measure how much the variance of one coefficient is increased because of collinearity were under three, indicating weak collinearity that does not warrant any correction (Belsley et al., 1980). All statistical tests were two tailed. For Figure 1, E and F , we performed similar analyses but included a factor either for cursor or for joystick position. As these factors are defined at every 10 ms time point, we fitted one mixed-effect model per time point and corrected p values for multiple comparisons using the false-discovery rate. To assess the amount of information in the visual feedback (e.g., the visual trajectory displayed as subjects performed the visuomotor reaching task), we performed receiving operator characteristics analyses computing the true- and false-positive rate for a range of criterion values, from zero to its maximum by steps of 0.01. We then searched for the criterion leading to the same false positive (false alarm) rate as found in the data and compared the corresponding true-positive rate (hit rate) to the one observed in the data.

For Figure 2D, we binned the peak cursor error into five quantiles computed independently for each participant but for all signal detection theory categories together. For Figure 2E, we plotted model predictions of the fixed effects of the cumulative link model for each signal detection theory category and increasing values of peak cursor error. To estimate metacognitive efficiency, or the extent to which the information available to the detection decision was used to scale their confidence, we used the M-ratio between meta- d' (Maniscalco and Lau, 2014) and d' estimated in a response-specific version of the HMeta- d' toolbox (Fleming, 2017). These meta- d' and d' allow us to compute the M-ratio, a ratio between meta- d' and d' , which quantifies metacognitive efficiency or how well information from first-order performance informs the metacognitive process. We used the default parameters of three chains of 10,000 samples with 1000 burn in samples for the Monte-Carlo Markov Chain procedures with no thinning. Visual inspection of chains and \hat{R} values well under 1.1 indicated good convergence. Of note, as the d' does not vary according to the response, differences in the M-ratio are solely driven by differences in the meta- d' .

fMRI data collection, preprocessing, and analyses. We acquired functional MRI images with a 3T whole-body scanner (Trio TIM, Siemens) with a 12-channel head-coil. Functional images were acquired with a susceptibility weighted EPI sequence with the following parameters: TR/TE = 2100/30 ms, flip angle = 80°, Parallel Acquisition Techniques (PAT) factor = 2, 64 × 64 voxel, 3.2 × 3.2 mm, 36 slices, 3.2 mm slice thickness, and 20% slice gap. We acquired structural images using a T1-weighted 3D sequence using the following parameters: MPRAGE; TR/TI/TE = 1900/900/2.32 ms, respectively; flip angle = 9°; voxel dimensions, 0.9 mm isotropic; 256 × 256 × 192 voxels. We presented task stimuli on a back-projection screen inside the scanner bore using an LCD projector (CP-SX1350, Hitachi). We recorded responses via buttons placed on the joystick used for the visuomotor reaching task (HH-JOY-4, Current Designs).

We used SPM8 software (<https://www.fil.ion.ucl.ac.uk/spm/>) for statistical analyses of functional data with a standard pipeline. We first corrected for head movements between scans by an affine registration and realignment to the mean of all images. The anatomic image was spatially normalized on the T1 template. The functional images were also normalized to the EPI template, which were thereby transformed into standard stereotaxic space and resampled with a 3 × 3 × 3 mm voxel size. The normalized images were spatially smoothed using an 8 mm full-width at half-maximum Gaussian kernel. We used the general linear model framework implemented in SPM to analyze our data. We modeled the convolved standard hemodynamic response function with a delta (or stick) function at the onset of the preparatory phase (appearance of white triangle, PREP regressor), at the onset of the joystick movement (start of movement, MOV regressor), and at the onset of the response

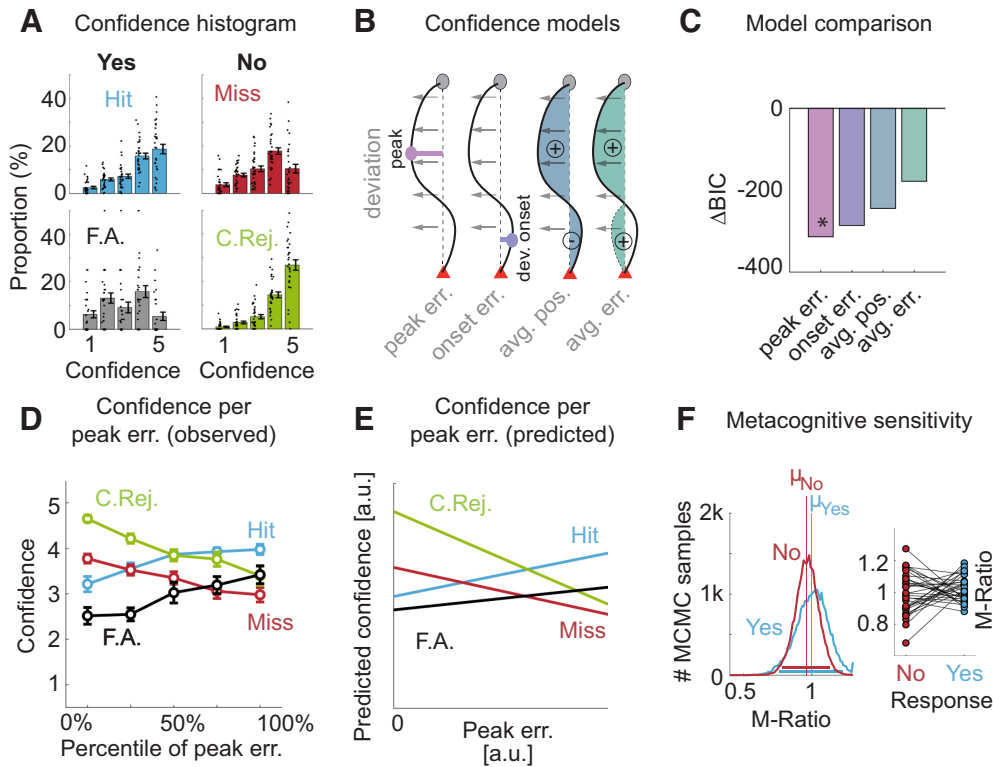


Figure 2. Confidence ratings and metacognition. **A**, Distribution of confidence ratings for hits (blue), misses (red) and correct rejections (green). Each black dot represents the data of one participant. **B**, Schematic depiction of the four regressors used for the four confidence regression models tested. **C**, Improvement in BIC (compared with a model with no cursor information). Note that the peak cursor error (peak err.) model shows the largest improvement (*). Additional model comparisons can be found in Extended Data Figure 2-1. **D**, Confidence for different percentiles of peak cursor error for hits (blue), misses (red), and correct rejections (green). **E**, Fixed effects predictions of confidence for comparable levels of peak error (normalized) for hits (blue), misses (red), correct rejections (green) and false-alarms (black). **F**, Hierarchical Bayesian estimation of response-specific metacognitive efficiency using the M-ratio. Left, Posterior probability for yes (blue) and no (red) responses. Vertical lines show the mean M-ratio, and horizontal bars show the 95% confidence interval. Right, Single participant values of the M-ratio. Shaded areas and whiskers **A–F** indicate 95% confidence intervals.

screen (yes vs no, RESP regressor). To examine brain regions whose activity fluctuated with trial-by-trial confidence, we took a parametric modulation approach (Fleming et al., 2018; Pereira et al., 2020); the MOV regressor event regressors were modulated by four additional parametric factors (Wood et al., 2008) representing in order, the trial-by-trial values of the angle of the deviation, the detection response yes versus no, the peak cursor error, and the confidence rating, leading to the following regression equation: $BOLD \sim deviation_angle + detection_response + peak_cursor_error + confidence$.

However, we also replicated our results using four single regressor models ($BOLD \sim deviation_angle$, $BOLD \sim detection_response$, $BOLD \sim peak_cursor_error$, $BOLD \sim confidence$). To account for head motion-related variance, we included the six differential parameters derived from the realignment process [x , y , and z translations (in millimeters) plus pitch, roll, and yaw rotations] as regressors of no interest. Low-frequency signal drifts were filtered using a cutoff period of 128 s. Global scaling was applied, with each fMRI value rescaled to a percentage value of the average whole-brain signal for that scan.

Contrast images from one-sample t tests corresponding to each event (PREP, MOV, RESP) and their parametric modulators, were fed into a second-level random-effect analysis. All second-level results are reported at a significance-level of $p < 0.05$ using cluster-extent familywise error (FWE) correction with a voxel-height threshold of $p < 0.001$. In Figure 3, activations are displayed at a cluster-size threshold of 30 voxels, using MRICroGL software (<https://www.nitrc.org/projects/mricrogl>). Data and analysis scripts from this study will be made freely available on acceptance.

Data and code availability

Behavioral data and analysis code are available here: <https://osf.io/278t6/>
fMRI activation maps can be found here: <https://neurovault.org/collections/13067/>.

Results

Participants correct for deviations regardless of detection

For both detected (hits) and undetected (misses) deviations, all trajectories correctly ended on the target (Fig. 1B), showing that participants always adjusted to the experimentally induced deviations. Reconstructions of what trajectories would be without deviation for hits and misses confirmed that participants would have clearly missed the target, even for miss trials (Fig. 1C), showing that they corrected for deviations even when not reporting them. We also assessed the relation between the lateral angle of the joystick (i.e., corrective motor command) and the cursor position (i.e., visual feedback). For this, we cross-correlated cursor position and joystick angle vectors across every trial, averaged over trials and participants, and observed that the strongest correlation occurred with a lag of $-0.45 \text{ s} \pm 0.01 \text{ s}$ (mean \pm 95% confidence interval; Fig. 1D). Furthermore, this correlation was significant for detected ($t_{(30)} = -41.30$, $p < 0.001$) and undetected ($t_{(30)} = -36.06$, $p < 0.001$) deviations, showing that participants corrected deviations regardless of whether they reported them or not.

Detection responses rely on cursor error

A one-up/two-down staircase procedure was used throughout the experiment, resulting in an average deviation angle of $20.59^\circ \pm 2.03$ overall. The detection performance was $67.71\% \pm 1.74$ correctly reported deviations ($d' = 1.69 \pm 0.14$), indicating that participants' performance was well above chance level. Participants were conservative in their response ($c = 0.37 \pm 0.08$, $t_{(30)} = 4.63$, $p < 0.001$). For deviated trials, neither small variations in the imposed

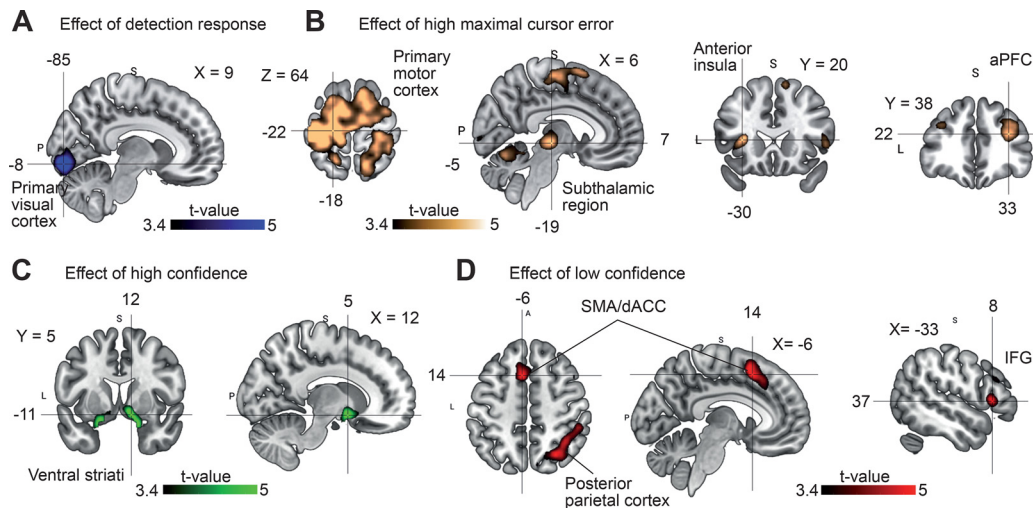


Figure 3. fMRI results. *A–D*, Statistical maps of parametric modulation contrast for explicit detection (yes responses; *A*), high peak cursor error (*B*), high confidence (*C*), and low confidence (*D*). Colors represent different parametric regressors and are independent from Figure 1. Results are displayed at $p < 0.001$, uncorrected, but all clusters displayed were significant after FWE correction ($p < 0.05$). IFG, Inferior frontal gyrus. Other brain activations can be found in Tables 1 and 2. Correlation between regressors can be found in Extended Data Figure 3-1. Extended Data Figure 3-2 shows beta values averaged across participants for each level of confidence for each ventral striatum using MarsBaR toolbox for SPM (Brett et al., 2002).

deviation angle because of the staircase procedure ($t_{(5160)} = 0.91$, $p = 0.36$), nor the index of the trial ($t_{(5160)} = -0.11$, $p = 0.91$) had an effect on detection responses. We therefore considered perceptual difficulty to be constant over time, which is standard in perceptual metacognition (Fleming et al., 2010).

When analyzing cursor position over time, we computed the horizontal distance between the cursor and the midline as an (unsigned) measure of cursor error (Fig. 1*E*). We then regressed detection responses using cursor error and trial type (deviated vs nondeviated) as fixed effects (generalized mixed effect models for every time point). We found a significant interaction effect during the first 0.85 s of the movement ($p < 0.05$, corrected), indicating that during this time, the relation between detection responses and cursor error depended on whether the trial was actually deviated or not. Furthermore, there was a main effect of cursor error without an interaction in a later time window up to 1.63 s after movement onset ($p < 0.05$) and a main effect of trial type during the whole movement ($p < 0.001$), suggesting that during this part of the movement, cursor error had a similar influence on detection responses regardless of whether the trial was deviated, with an additive bias because of trial type. We verified that cursor error also had an effect on detection responses in deviated trials ($p < 0.05$ between 0.49 and 1.94 s) and for undeviated trials ($p < 0.05$ between 0 and 1.61 s), confirming that the effect of cursor error on detection responses was genuine.

Our results are consistent with a threshold operation on the visual feedback, with higher cursor error leading to deviation detection (hit and false alarms; Fig. 1*E*). Therefore, detection responses should be best explained by the peak cursor error over the time course of the movement (i.e., deviations are detected if the cursor error reaches a threshold value at some point during the trial). Indeed, this model had a better Bayesian Information Criterion (BIC) than models with cursor error sampled at single time points (Extended Data Fig. 1-1*A*), as well as other measures of visual feedback (Extended Data Fig. 1-1*B*). Similarly, models based on putative motor or proprioceptive feedback yielded worse BIC (Extended Data Fig. 1-1*C*). The two main effects in this model were significant (peak cursor error, $t_{(6403)} = 5.31$, $p < 0.001$; trial type, $t_{(6403)} = 7.47$, $p < 0.001$), and there was no significant interaction ($t_{(6403)} = -1.32$, $p = 0.19$), confirming that

the occurrence of deviations did not change how visual information (quantified by peak cursor error) is integrated into detection responses but nevertheless influenced conscious detection. (Deviation was more frequently reported on deviated than nondeviated trials overall.) Nonetheless, the information carried by the peak cursor error would not have been sufficient for participants to have such performance (Extended Data Figs. 1-2, 1-3), showing that they must have relied on additional signals, whether those were based on visual, motor, or proprioceptive information.

Confidence scales with peak cursor error

We then examined confidence ratings across signal detection theory categories (Fig. 2*A*). To understand how visual information is integrated into confidence ratings, we regressed confidence using various summary statistics of cursor error over time (Fig. 2*B*). Our baseline model investigated the relation between detection responses and confidence and thus included only a predictor for trial type (deviated vs nondeviated), as well as for detection response (to model the fact that confidence is conditioned to the response; interaction effect, $z = 8.98$, $p < 0.001$). We then determined the best model in terms of BIC, which was obtained by adding a predictor for peak cursor error over time (peak err.). This model showed a lower BIC compared with the baseline model with a single predictor for detection response (Fig. 2*C*; $\Delta\text{BIC} = -313.80$, $p < 0.001$). Alternative models, such as including the cursor error at the onset of the deviation (onset err.), the average (signed) cursor position (avg. pos.), or the average of the cursor error (avg. err.), yielded lower improvements in BIC.

Using the winning model, we found an interaction between peak cursor error and detection responses (Fig. 2*D,E*; $z = 4.69$, $p < 0.001$), showing that confidence ratings were influenced by the peak cursor error conditioned to the detection response (i.e., when the peak cursor error is high, confidence in detected deviations is high but confidence in undetected deviations is low). There was also an interaction between trial type and peak cursor error ($z = 4.20$, $p < 0.001$) and between detection response and trial type ($z = 7.087$, $p < 0.001$) but no triple interaction between detection response, trial type, and peak cursor error ($z = -1.38$, $p = 0.17$). These results show that participants rated their

confidence by conditioning the peak cursor error to their conscious detection, with a confidence bias for nondeviated trials that depended on their response (e.g., more confident for correct rejections vs misses, less confident for false alarms vs hits).

To better understand the relation between confidence and peak cursor error, we fitted the data independently for deviated and nondeviated trials. We confirmed the interaction of response and peak cursor error for both trial types (deviated, $z = 5.77$, $p < 0.001$; nondeviated, $z = 5.76$, $p < 0.001$). Importantly, and contrary to the hypothesis of limited monitoring, we also found that confidence was significantly related to cursor error size for all trial categories ($p < 0.01$ for hits, misses, correct rejections, and false alarms). Together, these results show that confidence increases with peak cursor error when participants report a deviation but decreases with peak cursor error when participants do not report a deviation, independently of whether the deviation was actually because of an experimental manipulation (e.g., during misses) or to their own intrinsic visuomotor variability (e.g., during correct rejections; Fig. 2D,E).

Preserved metacognitive efficiency for unreported deviations

Finally, to confirm that participants reliably monitored their visuomotor actions, we estimated their (metacognitive) efficiency by measuring to what extent the information available to the detection decision was used to scale their confidence. For this, we fitted a response-specific hierarchical Bayesian model based on signal detection theory (Fleming, 2017), which estimates metacognitive efficiency for yes and no responses while controlling for task performance (this model used all yes and no trials in the fitting procedure). Results from this analysis revealed a metacognitive efficiency of 0.96 for the no responses (correct rejections and misses), indicating that participants efficiently used all information available for the decision to adjust their confidence ratings. Metacognitive efficiency for yes responses was similar (0.99; Fig. 2F). These results show that metacognitive efficiency was optimal in the sense that confidence ratings were as informative as what would be expected based on detection performance, and thus confirm that reporting deviations does not lead to better metacognitive efficiency.

Brain correlates of visuomotor performance

To investigate the neural substrate of the cognitive processes described above, we used trial-by-trial measures (confidence, response, peak cursor error, and deviation angle) as parametric regressors to model the fMRI BOLD signal during each movement. These regressors showed only limited correlation among themselves (maximal absolute mean $R = 0.41$; Extended Data Fig. 3-1). First, we found positive relationships between the BOLD signal and detection responses in the right primary visual cortex, meaning that activity increases when participants detected the deviations (Fig. 3A; Table 1). We found no significant negative relationships. We also found that larger peak cursor error yielded widespread activity in visual, motor, and subthalamic regions, as well as in the left insula, cingulate, and lateral prefrontal cortex [right inferior frontal gyrus and anterior prefrontal cortex (aPFC); Fig. 3B; Table 1]. Smaller cursor error yielded no differential activity beyond the statistical threshold, and we did not find any neural correlate of deviation angle.

Brain correlates of confidence

We then turned to confidence and found positive relationships with the BOLD signal, meaning that higher confidence was related to increased bilateral ventral striatum activity, extending to the left

Table 1. Summary of whole brain region activation for response and peak cursor error

Brain region	MNI coordinates (mm; x, y, z)	Cluster extent (voxels)	Significance
Yes responses			
Right primary visual cortex	9 -85 -8	602	$t = 6.64$, $p_{FWEC} < 0.001$
High peak cursor error			
Left primary motor cortex	-18 -22 64	3563	$t = 8.18$, $p_{FWEC} < 0.001$
Right inferior occipital gyrus	45 -67 1	927	$t = 6.17$, $p_{FWEC} < 0.001$
Left anterior insula	-30 20 7	632	$t = 5.33$, $p_{FWEC} < 0.001$
Right subthalamic region	6 -19 -5	285	$t = 5.26$, $p_{FWEC} < 0.001$
Right lateral anterior prefrontal cortex	33 38 22	229	$t = 5.13$, $p_{FWEC} < 0.001$
Right midcingulate gyrus	15 -19 37	106	$t = 4.98$, $p_{FWEC} = 0.028$
Right inferior frontal gyrus	48 29 1	279	$t = 4.47$, $p_{FWEC} < 0.001$
Left calcarine cortex	-12 -85 -8	147	$t = 4.41$, $p_{FWEC} = 0.008$
Left superior occipital gyrus	-21 -79 28	108	$t = 4.16$, $p_{FWEC} = 0.026$

No surviving voxels for the contrasts low peak error, no responses; $p_{FWEC} = p$ corrected for Family-Wise Error at the cluster level.

Table 2. Summary of whole brain region activation for confidence

Brain region	MNI coordinates (mm; x, y, z)	Cluster extent (voxels)	Significance
High confidence			
Right ventral striatum	12 5 -11	103	$t = 5.85$, $p_{FWEC} = 0.040$
Left amygdala	-21 -1 -17	96	$t = 5.80$, $p_{FWEC} = 0.049$
Extending in the ventral striatum	-15 8 -14		$t = 4.69$
Low confidence			
Left dACC/SMA	-6 14 49	221	$t = 5.33$, $p_{FWEC} = 0.002$
Left inferior frontal gyrus	-51 20 4	119	$t = 4.97$, $p_{FWEC} = 0.024$
Left middle frontal gyrus	-33 8 37	134	$t = 4.51$, $p_{FWEC} = 0.016$
Right posterior parietal cortex/angular gyrus	36 -49 40	403	$t = 5.43$, $p_{FWEC} < 0.001$

amygdala (Fig. 3C; Table 2). Conversely, we found negative relationships, meaning that lower confidence was associated with increased activity in the left supplementary motor area (SMA), extending to the dorsal anterior cingulate cortex (dACC), together with the left inferior and middle frontal gyri, as well as the right posterior parietal cortex (Fig. 3D; Table 2).

Discussion

We studied the behavioral and neural correlates of detection and confidence in a visuomotor reaching task with perturbations. We found that although participants often reported not being aware of deviations applied experimentally to the trajectory of their movement, their metacognitive efficiency was not impaired by this reported lack of awareness. Thus, they adjusted confidence judgments to the accuracy of their detection responses even when reporting no deviations. Furthermore, regression model selection revealed that regardless of whether participants were aware of the deviation, their judgment relied on a summary of visual feedback (i.e., peak cursor error), whose magnitude correlated with activity in a specific visuomotor brain network. Finally, we uncovered neural activity patterns associated with confidence, with high confidence engaging subcortical areas in the ventral striatum and low confidence engaging a cortical network in medial and lateral frontal cortex as well as posterior parietal cortex.

In terms of detection responses, we found that participants tended to report an experimentally induced deviation when the visible cursor error was high and conversely did not report one when the cursor error was low. We interpret these detection reports as awareness of the deviation, although as with any measure based on introspection, it is conceivable that participants had partial awareness while reporting being unaware (Mudrik and Deouell, 2022). These findings are consistent with awareness of a deviation occurring when cursor error exceeds a threshold; when the peak cursor error is higher than this threshold, a deviation is reported. Other regression models of detection responses or confidence could not fit the data as well, whether they were based on other measures of visual feedback or on putative proprioceptive signals (indexed by joystick position). Still, a part of the variance of participants' performance was not explained by the peak cursor error, suggesting that participants use other dimensions of the visual feedback not captured by the peak cursor error or other motor or proprioceptive signals. This idea is in line with previous works showing that participants at least partially rely on additional information, possibly stemming from the proprioceptive inputs (Farrer et al., 2008) and/or the comparison of the perceived cursor position relative to a self-generated prediction, based on an efferent copy of their motor command (Wolpert et al., 1995). The later assumption agrees well with past research on motor control and agency (Haggard, 2017). As the contribution of these different signals is still unclear in the context of motor control (e.g., Kasuga et al., 2022), future studies will be required to disentangle their contribution to deviation detection and its associated confidence.

To further assess how well participants monitored their movements, we examined their metacognitive efficiency when they reported being unaware of deviations that they nonetheless successfully corrected. Remarkably, metacognitive efficiency was similar when participants reported deviations and when they did not, suggesting that even if the visuomotor information was not enough to report a deviation, it could still be harnessed to rate confidence. Moreover, we found that confidence increased with peak cursor error when participants reported the deviation and instead decreased with peak cursor error when participants did not report the deviation. This behavior reveals a judicious use of visuomotor signals to guide confidence, regardless of the reported awareness of the deviation. We therefore surmise that participants were able to consciously monitor their unfolding movements by having access to at least a summary of visuomotor cues that guided their explicit detection and confidence responses, similar to what could occur for internally accumulated sensory signals (Pereira et al., 2021). The factors that constitute the evidence used to compute confidence are still unknown. As metacognitive efficiency was as good as expected considering detection performance, it is possible that they directly derive from the decisional signal used for detection responses. Therefore, our findings question the notion that small deviations are corrected mostly in the absence of awareness (Blakemore and Frith, 2003; Custers and Aarts, 2010); people might be unaware of some deviations that they correct, but they still have a calibrated feel of their performance.

Our finding of preserved metacognition for no responses thus sharply contrasts with visual metacognition studies describing lower metacognition for putatively unaware stimuli (Kanai et al., 2010; Mazor et al., 2020; Pereira et al., 2021). Considering that the factors responsible for metacognitive inefficiency are still unclear (Shekhar and Rahnev, 2021), we can only speculate on why metacognitive performance does not decrease for unaware deviations in our task. Confidence for unaware visual stimuli was

proposed to depend on monitoring attention instead of perceptual evidence (Kanai et al., 2010; Mazor et al., 2020). According to this view, confidence for aware and unaware stimuli might involve different mechanisms. In other studies, reduced metacognitive efficiency for unaware stimuli was accounted for by a single mechanism (Kellij et al., 2021; Pereira et al., 2021), based on the fact that the variances for noise and stimuli differ. It is therefore possible that in our study, equal variance between the noise and the signal allowed metacognition to be preserved for no responses. Indeed, the absence of response-specific differences in metacognitive sensitivity suggests that neither unequal variance nor response-specific metacognition are good candidates to explain our data. It is therefore possible that unequal variance occurs only in perceptual metacognition studies with perceptual uncertainty on the stimulus (which is either weak or embedded in noise). In our study, there is no perceptual uncertainty on peak cursor error, it is simply smaller for unaware deviations, therefore not leading to unequal variance nor lower metacognitive sensitivity for unaware deviations.

Our fMRI results showed extended activations associated with increasing peak cursor error involving sensorimotor regions, as well as occipital areas, anterior prefrontal and midcingulate cortex, insula, inferior frontal gyrus, and subthalamic regions. These results suggest the existence of widespread action-monitoring processes (Limanowski et al., 2017), possibly responsible for the integration of visual and proprioceptive feedback signals with motor action planning. Interestingly, the aPFC, including the frontal pole and posterior regions of Brodmann area 46 (Neubert et al., 2014), has been extensively linked to perceptual metacognition in studies using voxel-based morphometry (Fleming et al., 2010) or transcranial magnetic stimulation (Rahnev et al., 2016) or patients with focal lesions (Fleming et al., 2014). In our study, activity in the aPFC was modulated by the peak cursor error and not by the confidence regressor, suggesting that the aPFC might not be involved in confidence per se but rather in monitoring the performance of actions. Further in line with our interpretation, patients with prefrontal lesions reported fewer deviations than healthy controls despite similar corrective behavior in a task similar to ours (Slachevsky et al., 2001). In sum, our finding of increased aPFC activity reflecting cursor errors and not confidence pleads in favor of a role of the aPFC for monitoring action performance rather than the accuracy of decisions.

Finally, we found that low confidence was associated with higher activity in a network of medial frontal cortex, left inferior and middle frontal gyri, and right posterior parietal cortex, providing novel support for a role of frontoparietal regions in metacognitive processes via graded confidence computation (Hebart et al., 2016; Vaccaro and Fleming, 2018), extending here to the motor rather than purely perceptual domain. Conversely, high confidence engaged the ventral striatum, again corroborating fMRI findings in the perceptual domain (Hebart et al., 2016; Guggenmos et al., 2016; Mazor et al., 2020). Apart from its well-known involvement in reward-based learning (Daniel and Pollmann, 2014), the ventral striatum also computes pseudo-reward prediction errors, defined as reward predictions errors related to the subjectively perceived progress in a given task rather than merely driven by external (e.g., monetary) reward (Westbrook and Braver, 2016). Our results therefore dovetail with a broader putative role of the ventral striatum for valuation information predicting reward (Schultz et al., 1992; Daniel and Pollmann, 2014), whereby in the absence of feedback, the valuation information corresponds to confidence (Daniel and Pollmann, 2014). Together, these findings support a key role of

the ventral striatum in monitoring decisional signals for confidence (Daniel and Pollmann, 2014; Hebart et al., 2016; Vaccaro and Fleming, 2018) that can be used to adapt subsequent behavior even in the absence of external feedback (Guggenmos et al., 2016).

To conclude, our study sheds new light on both cognitive and neural mechanisms that underpin the monitoring of visuomotor behavior; participants base their detection and confidence reports on at least some summary statistic of visual feedback. We also map these monitoring and correcting processes to activity in an extended brain network including the aPFC, calling for a revision of its role in tracking confidence only. Importantly, although participants did report not being aware of some deviations, their confidence ratings were still calibrated to their actual performance on a trial-by-trial basis and guided by the same summary statistics as detection responses. These results may offer a plausible explanation for a paradox—that humans perform corrective actions in the absence of awareness but are good at attributing actions to themselves or to an external agent (de Vignemont and Fourneret, 2004). We argue that even if participants are unaware of their corrections, they can still accurately report their confidence through an efficient summary statistic of action parameters. They only become aware that something is wrong when that summary statistic exceeds what could be expected from their own intrinsic motor variability. Elucidating these mechanisms bears important clinical implications as deficits in the awareness of action have been linked to psychiatric diseases (Blakemore and Frith, 2003) such as schizophrenia (Frith et al., 2000; Voss et al., 2010). Our methodology should catalyze future research in the visuomotor domain to uncover metacognitive deficit in schizophrenia, or more generally, psychosis spectrum (Rouy et al., 2021), and what the underlying defective process is. It will be important to examine whether these clinical populations use the same mechanisms to compute confidence as we describe here, and if so, how deficits in such mechanisms relate to specific pathophysiological dimensions (positive, i.e., psychotic; negative, i.e., amotivational symptoms) and to particular brain substrates.

References

- Arbuzaova P, Peters C, Röd L, Koß C, Maurer H, Maurer LK, Müller H, Verrel J, Filevich E (2020) Measuring metacognition of direct and indirect parameters of voluntary movement. *J Exp Psychol Gen* 150:2208–2229.
- Belsley DA, Kuh E, Welsch RE (1980) *Regression diagnostics: identifying influential data and sources of collinearity*. New York: Wiley.
- Brett M, Anton J-L, Valabregue R, Poline J-B (2002) Region of interest analysis using an SPM toolbox. Paper presented at the Eighth International Conference on Functional Mapping of the Human Brain, Sendai, Japan, June.
- Binsted G, Brownell K, Vorontsova Z, Heath M, Saucier D (2007) Visuomotor system uses target features unavailable to conscious awareness. *Proc Natl Acad Sci U S A* 104:12669–12672.
- Blakemore S-J, Frith C (2003) Self-awareness and action. *Curr Opin Neurobiol* 13:219–224.
- Blakemore S-J, Wolpert DM, Frith CD (2002) Abnormalities in the awareness of action. *Trends Cogn Sci* 6:237–242.
- Charles L, Chardin C, Haggard P (2020) Evidence for metacognitive bias in perception of voluntary action. *Cognition* 194:104041.
- Custers R, Aarts H (2010) The unconscious will: how the pursuit of goals operates outside of conscious awareness. *Science* 329:47–50.
- Daniel R, Pollmann S (2014) A universal role of the ventral striatum in reward-based learning: evidence from human studies. *Neurobiol Learn Mem* 114:90–100.
- de Vignemont F, Fourneret P (2004) The sense of agency: a philosophical and empirical review of the “Who” system. *Conscious Cogn* 13:1–19.
- Farrer C, Bouchereau M, Jeannerod M, Franck N (2008) Effect of distorted visual feedback on the sense of agency. *Behav Neurol* 19:53–57.
- Flavell JH (1979) Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist* 34:906–911.
- Fleming SM (2017) HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neurosci Conscious* 2017:nix007.
- Fleming SM, Lau HC (2014) How to measure metacognition. *Front Hum Neurosci* 8:443.
- Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G (2010) Relating introspective accuracy to individual differences in brain structure. *Science* 329:1541–1543.
- Fleming SM, Huijgen J, Dolan RJ (2012) Prefrontal contributions to metacognition in perceptual decision making. *J Neurosci* 32:6117–6125.
- Fleming SM, Ryu J, Golfinos JG, Blackmon KE (2014) Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain* 137:2811–2822.
- Fleming SM, van der Putten EJ, Daw ND (2018) Neural mediators of changes of mind about perceptual decisions. *Nat Neurosci* 21:617–624.
- Fourneret P, Jeannerod M (1998) Limited conscious monitoring of motor performance in normal subjects. *Neuropsychologia* 36:1133–1140.
- Francken JC, Beerendonk L, Molenaar D, Fahrenfort JJ, Kiverstein JD, Seth AK, van Gaal S (2022) An academic survey on theoretical foundations, common assumptions and the current state of consciousness science. *Neurosci Conscious* 2022:niac011.
- Frith CD, Blakemore S, Wolpert DM (2000) Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Res Brain Res Rev* 31:357–363.
- Gaveau V, Pisella L, Priot A-E, Fukui T, Rossetti Y, Pélisson D, Prablanc C (2014) Automatic online control of motor adjustments in reaching and grasping. *Neuropsychologia* 55:25–40.
- Goodale MA, Pelisson D, Prablanc C (1986) Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature* 320:748–750.
- Guggenmos M, Wilbertz G, Hebart MN, Sterzer P (2016) Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *Elife* 5:e13388.
- Haggard P (2017) Sense of agency in the human brain. *Nat Rev Neurosci* 18:196–207.
- Hebart MN, Schriever Y, Donner TH, Haynes J-D (2016) The relationship between perceptual decision variables and confidence in the human brain. *Cereb Cortex* 26:118–130.
- Kanai R, Walsh V, Tseng C (2010) Subjective discriminability of invisibility: a framework for distinguishing perceptual and attentional failures of awareness. *Conscious Cogn* 19:1045–1057.
- Kasuga S, Crevecoeur F, Cross KP, Balalae P, Scott SH (2022) Integration of proprioceptive and visual feedback during online control of reaching. *J Neurophysiol* 127:354–372.
- Kellij S, Fahrenfort J, Lau H, Peters MAK, Odegaard B (2021) An investigation of how relative precision of target encoding influences metacognitive performance. *Atten Percept Psychophys* 83:512–524.
- Koriat A (2006) Metacognition and consciousness. In: *The Cambridge handbook of consciousness*, Vol 3, pp 289–326.
- Limanowski J, Kirilina E, Blankenburg F (2017) Neuronal correlates of continuous manual tracking under varying visual movement feedback in a virtual reality environment. *Neuroimage* 146:81–89.
- Locke SM, Mamassian P, Landy MS (2020) Performance monitoring for sensorimotor confidence: a visuomotor tracking study. *Cognition* 205:104396.
- Maniscalco B, Lau H (2014) Signal detection theory analysis of type 1 and type 2 data: meta-d’, response-specific meta-d’, and the unequal variance SDT model. In: *The cognitive neuroscience of metacognition* (Fleming SM, Frith CD (eds), pp 25–66. Berlin: Springer.
- Mazor M, Friston KJ, Fleming SM (2020) Distinct neural contributions to metacognition for detecting, but not discriminating visual stimuli. *Elife* 9:e53900.
- Meuwese JDI, van Loon AM, Lamme VAF, Fahrenfort JJ (2014) The subjective experience of object recognition: Comparing metacognition for object detection and object categorization. *Atten Percept Psychophys* 76:1057–1068.
- 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.
- Mudrik L, Deouell LY (2022) Neuroscientific evidence for processing without awareness. *Annu Rev Neurosci* 45:403–423.

- Neubert F-X, Mars RB, Thomas AG, Sallet J, Rushworth MFS (2014) Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron* 81:700–713.
- Pereira M, Faivre N, Iturrate I, Wirthlin M, Serafini L, Martin S, Desvachez A, Blanke O, Van De Ville D, Millán J, del R (2020) Disentangling the origins of confidence in speeded perceptual judgments through multimodal imaging. *Proc Natl Acad Sci U S A* 117:8382–8390.
- Pereira M, Megevand P, Tan MX, Chang W, Wang S, Rezaei A, Seeck M, Corniola M, Momjian S, Bernasconi F, Blanke O, Faivre N (2021) Evidence accumulation relates to perceptual consciousness and monitoring. *Nat Commun* 12:3261.
- Rouy M, Saliou P, Nalborczyk L, Pereira M, Roux P, Faivre N (2021) Systematic review and meta-analysis of metacognitive abilities in individuals with schizophrenia spectrum disorders. *Neurosci Biobehav Rev* 126:329–337.
- Rahnev D, et al. (2020) The confidence database. *Nat Hum Behav* 4:317–325.
- Rahnev D, Nee DE, Riddle J, Larson AS, D'Esposito M (2016) Causal evidence for frontal cortex organization for perceptual decision making. *Proc Natl Acad Sci U S A* 113:6059–6064.
- Shekhar M, Rahnev D (2021) Sources of metacognitive inefficiency. *Trends Cogn Sci* 25:12–23.
- Schultz W, Apicella P, Scarmati E, Ljungberg T (1992) Neuronal activity in monkey ventral striatum related to the expectation of reward. *J Neurosci* 12:4595–4610.
- Sinanaj I, Cojan Y, Vuilleumier P (2015) Inter-individual variability in metacognitive ability for visuomotor performance and underlying brain structures. *Conscious Cogn* 36:327–337.
- Slachevsky A, Pillon B, Fournerey P, Pradat-Diehl P, Jeannerod M, Dubois B (2001) Preserved adjustment but impaired awareness in a sensory-motor conflict following prefrontal lesions. *J Cogn Neurosci* 13:332–340.
- Tsay JS, Parvin DE, Ivry RB (2020) Continuous reports of sensed hand position during sensorimotor adaptation. *J Neurophysiol* 124:1122–1130.
- Vaccaro AG, Fleming SM (2018) Thinking about thinking: a coordinate-based meta-analysis of neuroimaging studies of metacognitive judgments. *Brain Neurosci Adv* 2:2398212818810591.
- Voss M, Moore J, Hauser M, Gallinat J, Heinz A, Haggard P (2010) Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain* 133:3104–3112.
- Westbrook A, Braver TS (2016) Dopamine does double duty in motivating cognitive effort. *Neuron* 89:695–710.
- Wolpert D, Ghahramani Z, Jordan M (1995) An internal model for sensorimotor integration. *Science* 269:1880–1882.
- Wood G, Nuerk H-C, Sturm D, Willmes K (2008) Using parametric regressors to disentangle properties of multi-feature processes. *Behav Brain Funct* 4:38.