# Overlapping and unique neural circuits support perceptual decision making and confidence

Jiwon Yeon and Dobromir Rahnev

School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332

**Author contributions**

The two authors designed the study. J. Yeon collected and analyzed data, and both authors worked together to interpret the results. J. Yeon drafted the manuscript and D. Rahnev revised and commented on it. Both authors approved the final version of the manuscript for submission.

**Acknowledgements**

We thank Farshad Rafiei for help with the data collection. This work was funded by a startup grant to D.R. from the Georgia Institute of Technology.

**Keywords:** perceptual decision making, confidence, metacognition, prefrontal cortex, fMRI

**Corresponding author**

Jiwon Yeon

Georgia Institute of Technology

654 Cherry Str. NW

Atlanta, GA 30332

E-mail: wiseriver531@gmail.com

**Abstract**

Perceptual decision making is the process of making a judgment about the identity of a stimulus based on the available sensory information. Perceptual decisions are naturally accompanied by a sense of confidence in the accuracy of the decision. Although the neural correlates of perceptual decision making and confidence have been the subject of substantial research, there is disagreement on whether they are supported by the same or different neural circuits. To address this question, we designed a functional MRI experiment that temporally separated the processes of perceptual decision making and confidence by presenting the confidence scale only after the offset of the decision. Additionally, we decorrelated the regressors for decision and confidence by slowing down participants' responses and presenting the confidence question only on a fraction of the trials. We found that perceptual decision making and confidence were supported by large and mostly overlapping brain circuits including frontal, parietal, posterior, and cingulate regions. Further, confidence recruited a number of unique regions, whereas there was no evidence for perceptual decision making recruiting unique regions not involved in confidence. These results demonstrate that the neural circuits supporting perceptual decision making and confidence have a very high degree of overlap, and suggest a largely shared computational mechanisms with unique components for confidence but not for perceptual judgments.

## Introduction

Perceptual decision making is the process of making a judgment about the identity of a stimulus based on the available sensory information (Hanks & Summerfield, 2017). Such process of decision-making is engaged, for example, each time we judge the color of a traffic signal or the identity of a person down the hallway. Once our perceptual decision is formed, we are able to evaluate the likely accuracy our this decision using ratings of confidence (Fleming & Dolan, 2012; Metcalfe & Shimamura, 1994). Confidence judgments are often referred to as "metacognitive" because they represent a second-order decision about the accuracy of a first-order decision (Fleming & Daw, 2017; Shimamura, 2000; Yeung & Summerfield, 2012)

Perceptual decision making and confidence are strongly related to each other. In most computational frameworks, they are conceptualized as two separate judgments made on the exact same underlying information (Fetsch, Kiani, Newsome, & Shadlen, 2014; Galvin, Podd, Drga, & Whitmore, 2003; Hangya, Sanders, & Kepecs, 2016; Pouget, Drugowitsch, & Kepecs, 2016; Rahnev, Bahdo, de Lange, & Lau, 2012; Sanders, Hangya, & Kepecs, 2016). Further, the quality of decision and confidence judgments tends to be highly correlated such that higher stimulus sensitivity leads to higher metacognitive sensitivity (Fleming & Lau, 2014; Maniscalco & Lau, 2012).

This strong relationship between perceptual decision making and confidence suggests that these processes may be supported by similar brain circuits. This possibility has received support from animal studies, which have demonstrated that often the same neurons that encode the decision

appear to also encode confidence (Fetsch et al., 2014; Kepecs, Uchida, Zariwala, & Mainen, 2008; Kiani & Shadlen, 2009; Middlebrooks & Sommer, 2012; So & Stuphorn, 2015). These findings suggest that computations related to the primary perceptual decision and the confidence judgment are performed in the same neural populations. In this view, perceptual decision making and confidence are two sides of the same coin and are supported by the exact same brain areas.

At the same time, many other experiments have demonstrated that perceptual decision making and confidence can be dissociated, thus casting doubt on whether they are really supported by the same brain mechanisms. Behaviorally, there is mounting evidence that confidence judgments can be dissociated from the accuracy of the perceptual decision across a variety of different paradigms in both humans (Boldt, de Gardelle, & Yeung, 2017; Desender, Boldt, & Yeung, 2018; Lau & Passingham, 2006; Rahnev, Koizumi, McCurdy, D'Esposito, & Lau, 2015; Rahnev, Maniscalco, et al., 2011; Samaha, Barrett, Sheldon, LaRocque, & Postle, 2016; Song, Koizumi, & Lau, 2015; Spence, Dux, & Arnold, 2015; Vlassova, Donkin, & Pearson, 2014; Zylberberg, Fetsch, & Shadlen, 2016) and monkeys (Ferrigno, Kornell, & Cantlon, 2017). Typically, these studies construct two conditions that are matched on accuracy but participants express a different level of confidence in each condition. Such dissociations suggest that confidence judgments are not based on the exact same information as the perceptual decision. In fact, a number of researchers now include metacognitive noise – that is, noise that affects confidence ratings but not the perceptual decision – in their models of confidence (De Martino, Fleming, Garrett, & Dolan, 2013; Jang, Wallsten, & Huber, 2012; Maniscalco & Lau, 2016; Mueller & Weidemann, 2008; Rahnev, Nee, Riddle, Larson, & D'Esposito, 2016; Shekhar & Rahnev, 2018; van den Berg, Yoo, & Ma, 2017).

The existence of such noise, which has also been used to make novel behavioral predictions that were subsequently experimentally confirmed (J. W. Bang, Shekhar, & Rahnev, 2018) suggesting that the neural processes related to confidence may be fully or partially separated from neural processes related to perceptual decision making.

Even more direct evidence for a separation between the neural circuits supporting perceptual decision making and confidence judgments comes from studies employing brain stimulation. Specifically, a number of studies have reported that transcranial magnetic stimulation (TMS) delivered to the prefrontal cortex can alter participants' confidence ratings, while leaving their perceptual decisions unaffected (Fleming et al., 2015; Rahnev et al., 2016; Rounis, Maniscalco, Rothwell, Passingham, & Lau, 2010; Shekhar & Rahnev, 2018). Similar dissociations between the primary decision and confidence have been observed in studies of memory (Ryals, Rogers, Gross, Polnaszek, & Voss, 2016; Yazar, Bergström, & Simons, 2014; Ye, Zou, Lau, Hu, & Kwok, 2018). Such findings have compelled many researchers to hypothesize that perceptual decision making and confidence are based on partially separate neural mechanisms and circuits.

It thus remains unclear whether most of the neural circuits supporting perceptual decision making and confidence are shared or separate. This question is difficult to answer in animal models because only a subset of brain regions is typically investigated in such studies. Instead, the overlap between these two processes needs to be addressed by assessing the involvement of areas across the whole brain in each process.

Here we examined the neural circuits supporting perceptual decision making and the confidence using functional MRI (fMRI). We separated the decision and confidence processes temporally by asking participants to first report their perceptual decision and then report their confidence rating. Importantly, in order to delay the onset of confidence computations, we only provided the scale on which confidence should be rated after the perceptual decision was completed. Further, we decorrelated the regressors for the decision and confidence periods by extending both in time and additionally requiring confidence ratings to be made on only a fraction of all trials. To anticipate, we observed a large degree of overlap in the activated regions between perceptual decision making and confidence across the whole brain. Nevertheless, the overlap was incomplete and a number of brain areas were preferentially involved in confidence. These findings demonstrate that perceptual decision making and confidence are supported by largely the same neural circuits but that confidence does recruit additional brain areas compared to the perceptual decision.

## Methods

### Participants

Twenty-five participants completed the experiment (12 females, average age = 21.4 years, range = 18-32 years). All participants were right-handed, had no history of neurological disorders, and had normal or corrected-to-normal vision. The study was approved by the Georgia Tech Institutional Review Board. All participants were screened for MRI safety and provided informed consent.

### Stimulus

Participants judged the direction of motion of white dots (density: 2.4/degree$^2$; speed: 5°/s) presented in a black circle (2° radius). The black circle positioned either left or right of fixation (its center was 4° from the center of the screen). Presenting the stimuli in the periphery led to slower perceptual decisions thus allowing us to decorrelate the decision and confidence periods. In each trial, either 4 (low-coherence) or 8% (high-coherence) of the white dots moved coherently. The direction of the coherently moving dots was either left or right. The rest of the dots moved in random directions. Each dot had lifetime between three to five frames (refresh rate of the projector: 60 Hz) and the coherent motion was carried by a random subset of dots on each frame. The screen had gray background color. All stimuli were created in MATLAB, using the Psychtoolbox 3 (Pelli, 1997).

### Task

Participants indicated the direction of motion (left or right) and provided a confidence rating (Figure 1). Each trial began with a white fixation cross at the center of the screen. The fixation cross was presented randomly for two or three seconds. Following the fixation cross, we presented a cue indicating the likely side of the screen where the stimulus would be presented. The cue was invalid on 10% of all trials, which we considered as catch trials. The duration of the cue was randomly chosen to be .5, 1, 2, or 4 seconds. Participants were then presented with the moving dots stimulus and asked to decide on the direction of coherently moving dots. The stimulus presentation lasted until the participants indicated their response via a button press. Participants gave their responses using the index and middle fingers to indicate left and right direction, respectively. After participants provided their response, a prompt to report their confidence was presented on 55% of all trials. The confidence period was never presented on catch trials. The prompt indicated the confidence scale on which confidence should be rated on each trial. We randomly alternated between 2-point, 3-point, and 4-point scales. The scale to be used was signaled by presenting the number 2, 3, or 4 as the confidence prompt. The lowest level of confidence was always indicated by pressing a button with the index finger. There was no time pressure for either the decision or confidence responses.

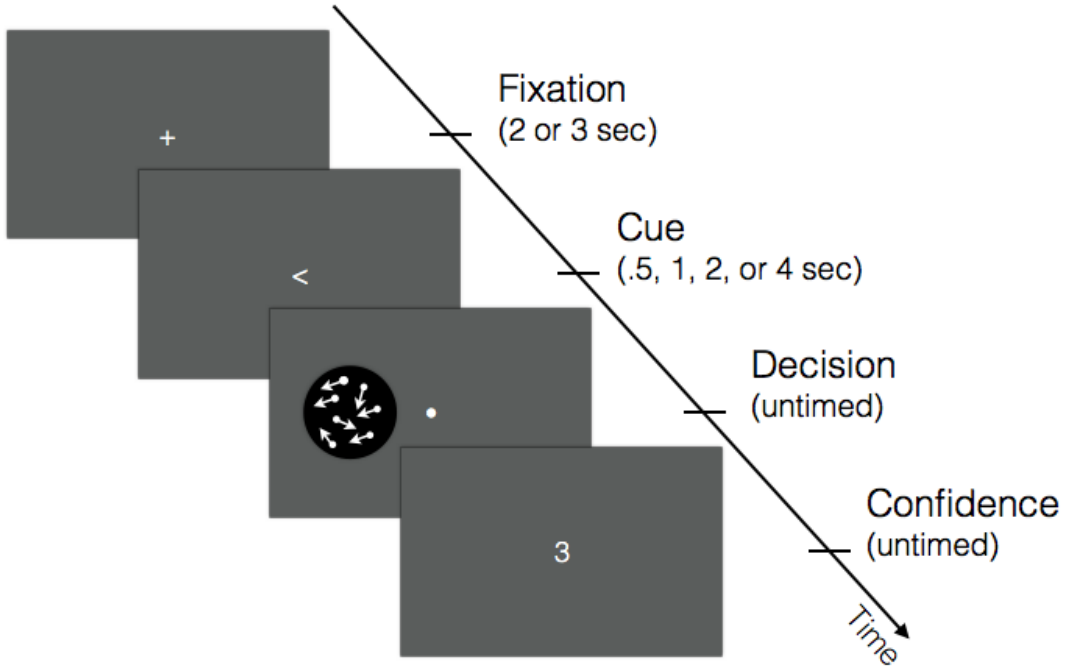


*Figure 1. Task. Each trial started with a white fixation cross. A cue then indicated the likely location of the following stimulus (either left or right of fixation). In the decision period, participants were asked to judge the direction of coherently moving dots. A confidence prompt was presented on 55% of all trials. Three rating scales (i.e., 2-, 3-, and 4-point scales) were used and the scale was indicated by the number presented on the screen (3-point scale in this figure).*

After coming to the lab, participants first underwent a short training outside the scanner to get acquainted with the task. During the training, participants received instructions on how to perform the task and completed 40 example trials. Following the training, participants were positioned in the scanner where we first collected a structural scan. During the structural image scanning, 23 participants practiced the task more (37.9 trials on average). Two participants did not have additional practice due to technical issues. The main experiment consisted of four runs, each consisting of four 20-trial blocks (for a total of 320 trials). Participants had unlimited resting time between the runs. The between-block intervals lasted 15 seconds and were signaled by a black fixation cross at the center of the screen. Responses were given with the right hand via an

MRI-compatible button box. Two participants completed only three runs and one participant completed three runs and a single block from the fourth run. The other 22 participants completed the full four runs.

Behavioral analyses

In order to delay the onset of the confidence computation, we interleaved three different scales for confidence evaluation. In order to be able to analyze confidence responses across the different scales, we normalized each confidence rating such that a confidence rating of $k$ given on an n-point scale was transformed to $\frac{k-1}{n-1}$. This transformation ensured that the ratings for all confidence scales were in the $[0,1]$ interval.

fMRI acquisition and preprocessing

The MRI data were collected 3T Trio MRI system (Siemens) using a 12-channel head coil. Anatomical images were acquired using a T1-weighted MPRAGE sequence (FoV = 256 mm; TR = 2530 ms; TE = 1.74 ms; 176 slices; flip angle = 7˚; voxel size = 1.0 × 1.0 × 1.0 $mm^3$). Functional images were acquired using a T2-wieghted gradient echo-planar imaging sequence (FoV = 220 mm; TR = 1780 ms; TE = 24 ms; 37 descending slices; flip angle = 70˚; voxel size = 3.0 × 3.0 × 3.5 $mm^3$).

We used SPM12 (Wellcome Department of Imaging Neuroscience, London, UK) to analyze the MRI data. The first two volumes of each run were removed to allow for scanner equilibration. Functional images were first converted from DICOM to NIFTI and were then preprocessed using

following steps: de-spiking, slice-timing correction, realignment, coregistration, segmentation, normalization, and smoothing with a 4-mm full-width-half-maximum (FWHM) Gaussian kernel.

fMRI analyses

The main goal of our fMRI analyses was to reveal the neural substrates of perceptual decision making and confidence. In a first set of analyses, we developed a general linear model (GLM) with 16 regressors. The first five regressors modeled the blood-oxygen level-dependent (BOLD) responses related to the cue period (spanning the period from onset to offset of the cue), decision period (spanning the period from onset of the stimulus until the decision response), catch trials decision period (the same decision period as in the "decision period" regressor but only in catch trials), confidence period (spanning the period from decision response to confidence response), and rest periods between blocks (spanning the period of rest in-between blocks). In addition, we included six regressors related to head movement (three translation and three rotation regressors), four tissue regressors (i.e., white matter, cerebrospinal fluid and bone, soft tissues, and air and background), and a constant term.

In order to determine the brain regions involved in perceptual decision making and confidence, we investigated the brain regions that showed higher activity in the decision and the confidence periods compared to the cue period (that is, Decision > Cue and Confidence > Cue). The cue period was specifically chosen as an active control: unlike break periods, it required participants to remain engaged in attentional processes and be alert to the forthcoming stimulus. The jittering in the cue duration ensured that the cue period regressor showed slightly negative correlations

with both the decision (average r = -.128, SD = .054) and the confidence (average r = -.211, SD = .052) periods. In addition, in order to find brain regions that were preferentially active for perceptual decision making or confidence, we directly compared the decision and the confidence periods (Decision > Confidence and Confidence > Decision). We modeled catch trials separately since they may invoke additional processes unrelated to perceptual decision making. The results of individual participants were submitted to a group-level t-test. Unless otherwise indicated, the statistical results reported were based on $p < 0.05$ threshold corrected for false discovery rate (FDR) and cluster size larger than 80 voxels.

To isolate the brain activations related to perceptual decision making and confidence, we attempted to decorrelate the decision and confidence regressors. To achieve this, we used moving dots as stimuli and placed them in the periphery in order to slow down the perceptual decision-making process. In addition, we slowed down the confidence process by varying the confidence scale across trials. Finally, we only asked for confidence responses on a little more than half of all trials. These design characteristics resulted in a low correlation between the decision and the confidence period regressors but the correlation remained positive (average r = .201, SD = .089). Therefore, to ensure that we fully isolate the processes related to perceptual decision making and confidence, we performed two control analyses in which we first regressed out the decision or confidence period regressors. Specifically, in the first analysis, we regressed out the decision period regressor from the BOLD time course and performed the Confidence > Cue contrast on the residuals. The GLM used for the Confidence > Cue contrast was equivalent to the main 16-regressor GLM, except that we did not include the decision regressor, which was

already removed. In the second analysis, we regressed out the confidence period regressor and extracted the activations for the Decision > Cue contrast in an equivalent way.

To check the quality of the collected data, we created a final GLM to compare activations for stimuli presented to the left or right of fixation. The model contained separate regressors for left-side and right-side stimuli instead of the single regressor for the decision period. The other 15 regressors were kept the same as in our main GLM analyses.

Data and code

All data and codes for the behavioral analyses are freely available at https://github.com/wiseriver531/Decision_Confidence_fMRI. In addition, unthresholded fMRI maps have been are uploaded at https://neurovault.org/collections/4323.

## Results

We investigated whether perceptual decision making and confidence are supported by the same or different brain regions. To this end, we designed an fMRI study in which we decorrelated the decision and confidence regressors, and then compared each of them to a cue period that served as an active control.

### Behavioral results

Participants' accuracy was 67.6% correct (SD = 9.91) and the average response time was 1.58 seconds (SD = 0.415). Performance was higher for high-coherence (accuracy = 69.2%, SD = 12.2) than for low-coherence stimuli (accuracy = 65.7%, SD = 9.40; paired t-test: $t(24) = 2.11$, $p = .045$).

To compare the confidence ratings across the different scales, we mapped each scale on the [0, 1] interval (see Methods). The average transformed confidence value was .649 (SD = .191). Participants were more confident for correct trials (mean confidence = .672, SD = .191) than incorrect trials (mean confidence = .592, SD = .193; paired t-test: $t(24) = 5.217$, $p < .001$). Confidence was also higher for the high-coherence stimuli (mean confidence = .660, SD = .188) than for the low-coherence stimuli (mean confidence = .638, SD = .197; paired t-test: $t(24) = 2.142$, $p = .043$). These results suggest that participants were able to perform the task as intended and provide appropriate confidence ratings.

### Visual cortex activity for stimuli presented on the left vs. right of fixation

To confirm the reliability of our data, we compared the activations produced by random motion dot stimuli presented on the left and right side of the screen. As expected, we found that stimuli presented on the left of fixation produced strong activations in the right visual cortex (maximum activation observed at [18, -86, -10], $t(24) = 7.01$, $p_{FWE_corrected} = .026$, cluster size = 1,077; Figure 2A), and stimuli presented on the right of fixation produced strong activations in the left visual cortex (maximum activation observed at [-8, -88, -8], $t(24) = 9.18$, $p_{FWE_corrected} < .001$, cluster size = 2,389, Figure 2B).

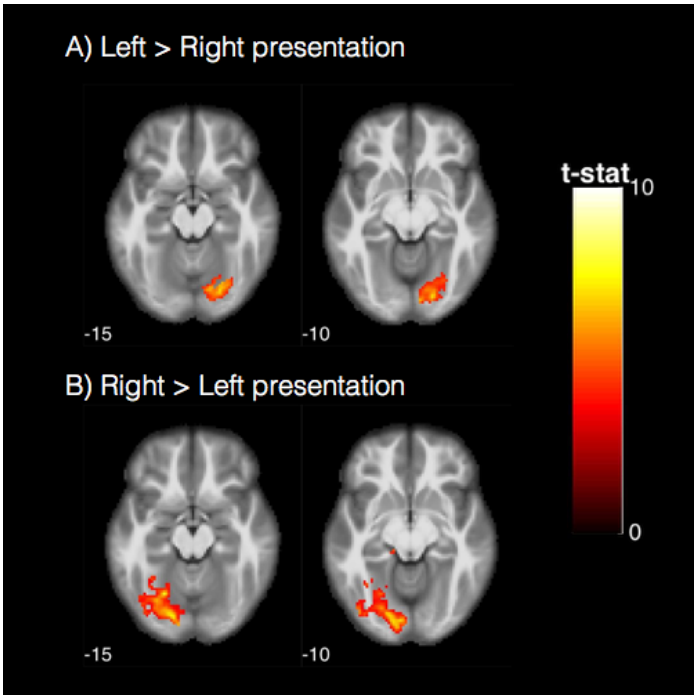


*Figure 2. Visual cortex activity for stimuli presented on the left vs. right of fixation. (A) Moving dots presented left of fixation elicited responses in right visual cortex. (B) Moving dots presented right of fixation elicited responses in left visual cortex. Colors indicate t-values.*

<u>Shared activity between perceptual decision making and confidence</u>

To explore the possible overlap between processes related to perceptual decision making and confidence, we first examined the activations for the decision and confidence periods separately. To do so, we compared the periods of decision and confidence deliberation with the pre-stimulus period when the spatial cue was presented. Our design was successful in decorrelating the cue regressor from both the decision regressor (average r = -.128, SD = .054) and the confidence regressor (average r = -.211, SD = .052) thus allowing us to use the cue period as an active control baseline.

We found that perceptual decision making (Figure 3A) and confidence (Figure 3B) activated a similar set of fronto-parieto-posterior brain regions. In order to examine the amount of overlap, we created a map of the intersection of the decision- and confidence-related activations (*Figure 3*C). This map showed extensive bilateral activations in a number of fronto-parietal areas including the middle frontal gyrus (MFG), frontal eye fields (FEF; in the superior precentral sulcus, sPCS), inferior precentral sulcus (iPCS), inferior frontal sulcus (IFS), and the intraparietal sulcus (IPS). In addition, strong bilateral activations were observed in the dorsal anterior cingulate cortex (dACC), anterior insula, and motion complex area (MT+). We also observed bilateral motor cortex activity, which was likely related to the fact that both the decision and confidence periods featured a button press. Coordinates of peak activity for all regions and corresponding *t*-values are shown in Table 1.

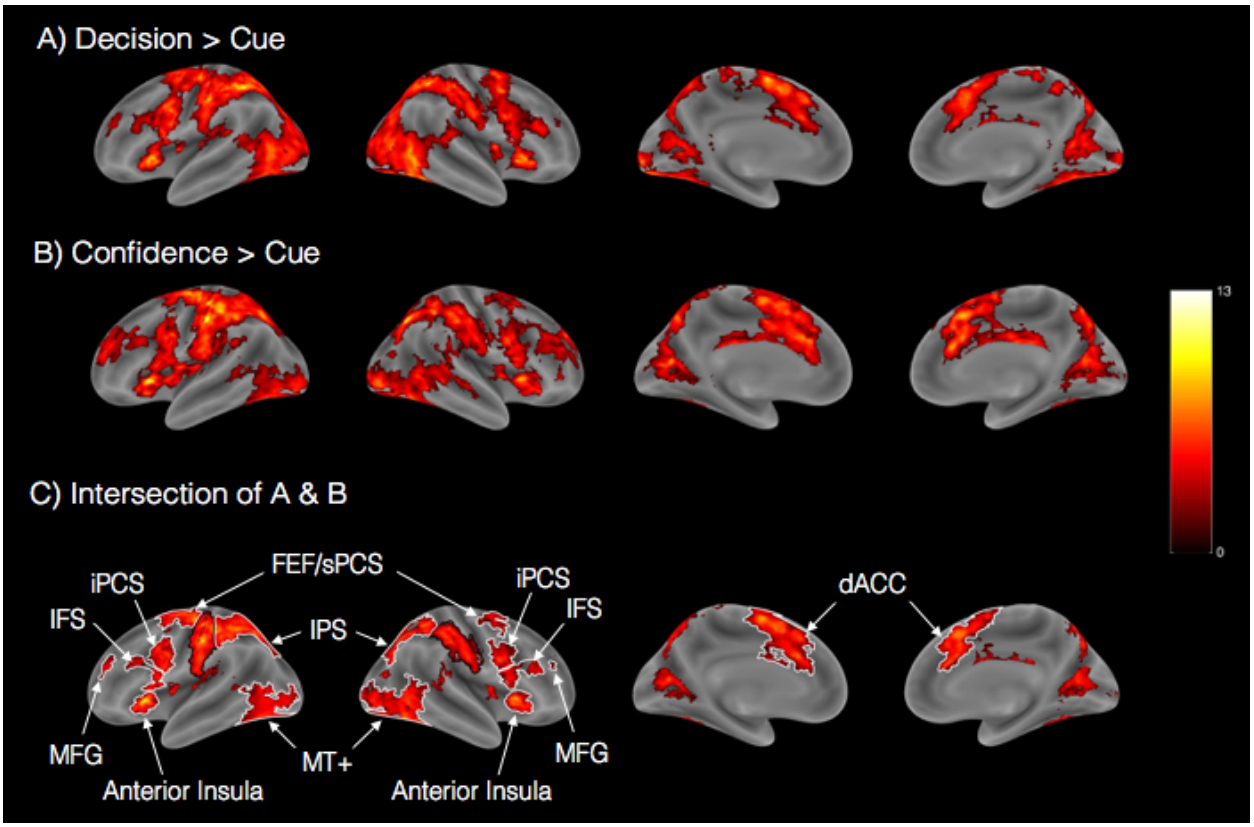


*Figure 3. Shared activity between perceptual decision making and confidence. (A) Decision-related brain activity obtained from the contrast Decision > Cue. (B) Confidence-related brain activity obtained from the contrast Confidence > Cue. (C) Intersection between the activation maps for perceptual decision making and confidence showing the areas of activation overlap. Colors indicate t-values. The t-values in (C) are the average of the Decision > Cue and Confidence > Cue t-values. The black borders delineate the motor cortex. dACC, dorsal anterior cingulate cortex; FEF, frontal eye field; IFS, inferior frontal sulcus; IPS, intraparietal sulcus; iPCS, inferior precentral sulcus; MFG, middle frontal gyrus; MT+, motion complex area; sPCS, superior precentral sulcus.*

| Anatomical region | Side | MNI Coordinates | | | t-value | |
|---|---|---|---|---|---|---|
| | | x | y | z | Decision > Cue | Confidence > Cue |
| Frontal eye field (FEF) | L | -26 | -4 | 54 | 12.575 | 6.827 |
| | R | 28 | -2 | 58 | 7.567 | 6.566 |
| Intraparietal sulcus (IPS) | L | -40 | -34 | 40 | 10.059 | 10.722 |
| | R | 38 | -42 | 48 | 7.532 | 10.998 |
| Anterior insula | L | -28 | 20 | 4 | 8.657 | 11.246 |
| | R | 34 | 18 | 8 | 8.926 | 9.623 |
| Inferior precentral sulcus (iPCS) | L | -48 | 2 | 38 | 9.711 | 6.443 |
| | R | 54 | 10 | 32 | 6.845 | 7.812 |
| Dorsal anterior cingulate cortex (dACC) | L | -8 | 10 | 46 | 6.427 | 7.740 |
| | R | 10 | 18 | 38 | 7.776 | 9.208 |
| Inferior frontal sulcus (IFS) | L | -50 | 6 | 8 | 6.795 | 4.795 |
| | R | 52 | 10 | 8 | 4.143 | 6.818 |
| Middle frontal gyrus (MFG) | L | -32 | 48 | 26 | 4.571 | 6.148 |
| | R | 30 | 46 | 32 | 3.902 | 5.243 |

*Table 1. Coordinates and t-values of peak voxels for the intersection of the Decision > Cue and the Confidence > Cue contrasts.*

These results point to the presence of substantial overlap between the neural processes related to perceptual decision making and confidence. However, an alternative explanation for this overlap is that the decision and confidence periods may not have been perfectly separated in our GLM analyses. Indeed, the decision and confidence regressors showed a small but positive correlation (r = .201, SD = .089). Therefore, we performed control analyses in which we first regressed out the confidence-related activity from the BOLD time course and then created a new GLM on the residuals where we assessed the activations for Decision > Cue. This procedure ensured that confidence-related processes were completely removed from the decision regressor. This analysis revealed a very similar set of activations for the Decision > Cue contrast (Figure 4A). Similarly, in a different set of analyses, we regressed out the decision-related activity and created a map for Confidence > Cue contrast using the residuals. Again, we found a very similar pattern of activations (Figure 4B). Finally, we examined the intersection of the two sets of

activations and found activations in each of the regions found in our previous analysis, such as MFG, FEF/sPCS, iPCS, IFS, IPS, dACC, anterior insula, MT+, and the motor cortex (Figure 4C).

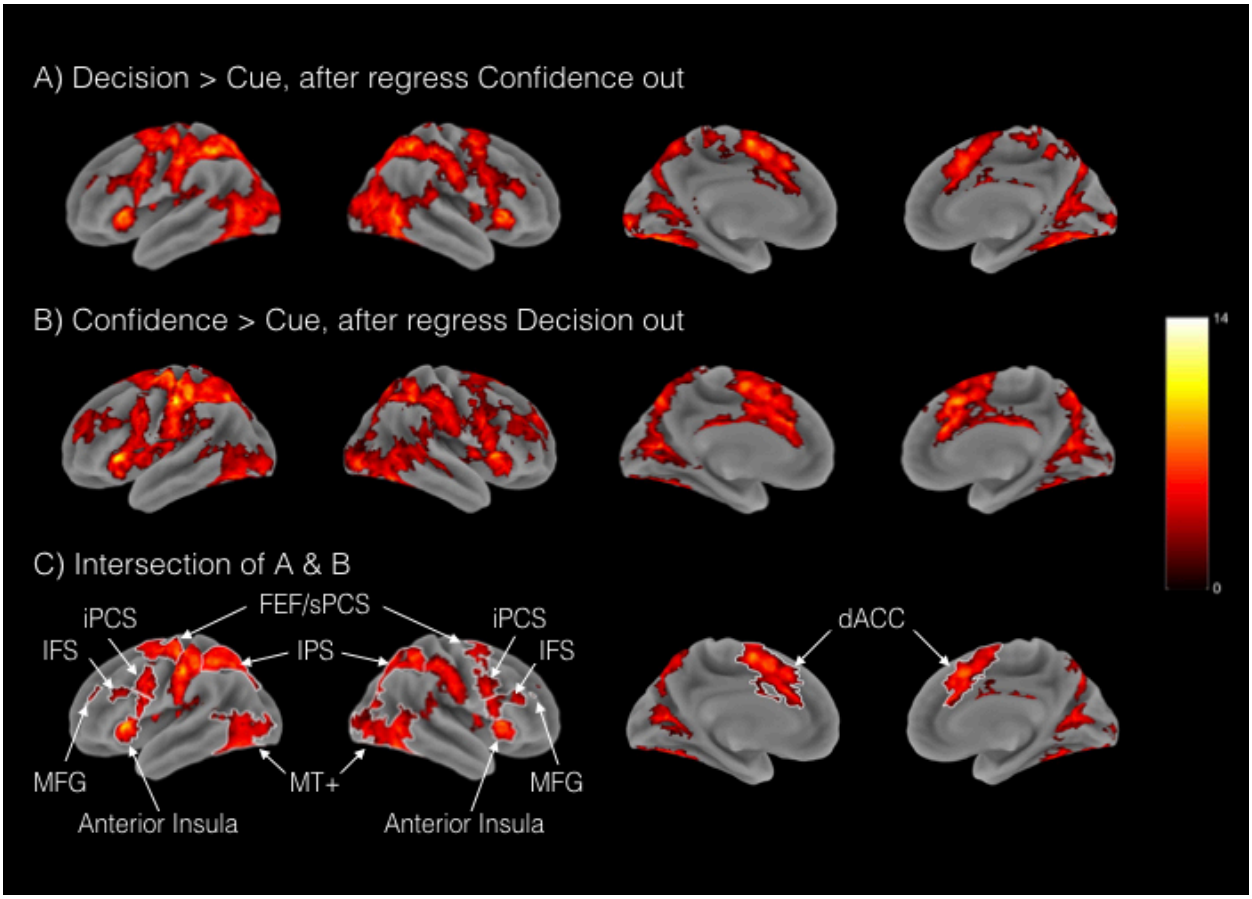


*Figure 4. Control analyses on the shared activity between perceptual decision making and confidence. (A) Decision-related brain activity obtained from the contrast Decision > Cue after confidence-related activations had been removed in a previous regression. (B) Confidence-related brain activity obtained from the contrast Confidence > Cue after decision-related activations had been removed in a previous regression. (C) Intersection between the activation maps in panels A and B. All figure conventions are the same as in Figure 3.*

<u>Unique activity for decision and confidence processes</u>

Despite the presence of substantial overlap, it is possible that perceptual decision making and confidence also produce dissociable activations in a subset of brain regions. To check for the

presence of such unique activations, we examined the set of brain regions activated more strongly for decision-related processes (using the contrast Decision > Confidence). We found activity in bilateral MT+, right FEF/sPCS, and right IPS (Figure 5A and Table 2). Notably, the large bilateral activation in MT+ is likely driven by the presence of moving stimuli during the decision but not the confidence period and therefore may not reflect decision processes that are independent of the moving dots stimulus. Similarly, right FEF and right IPS are part of the dorsal attention network (Corbetta & Shulman, 2002), which may have been engaged in sustaining attention to the peripheral location where the stimulus was presented. Therefore, the observed results do not necessarily suggest the existence of any regions with higher activity for perceptual decision making itself compared to the confidence itself (independent of attentional processes that may have occurred during the decision period).

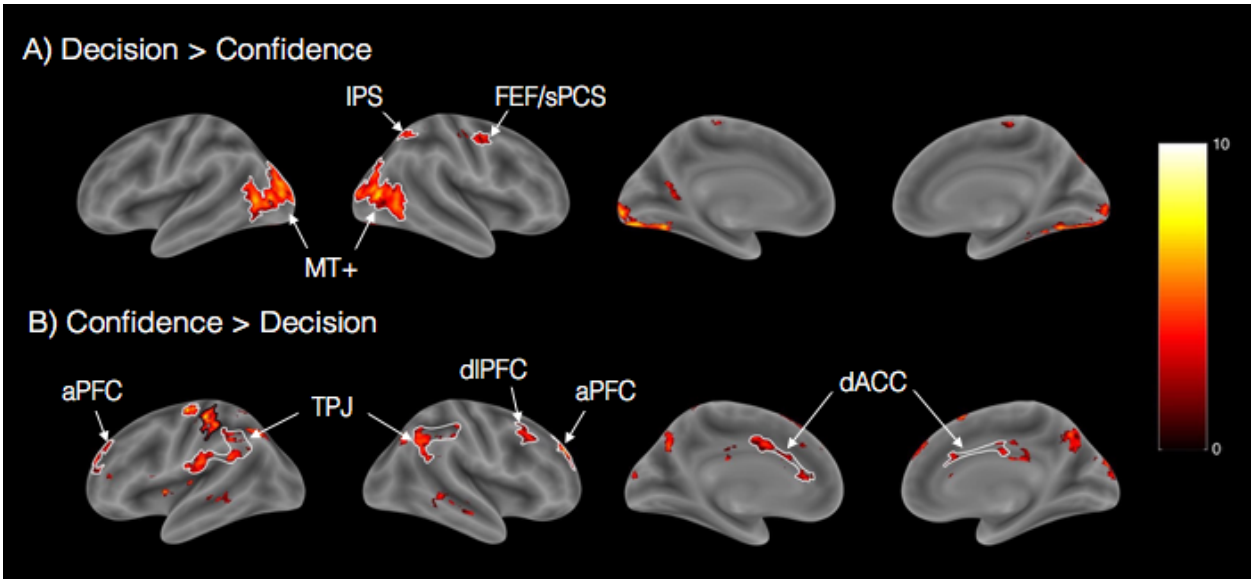


*Figure 5. Unique activity for perceptual decision making and confidence. (A) Regions showing higher activations for decision- than confidence-related processes. (B) Regions showing higher activations for confidence- than decision-related processes. Colors indicate t-values. aPFC, anterior prefrontal cortex; dACC, dorsal anterior cingulate cortex; dlPFC, dorsolateral prefrontal*

*cortex; FEF, frontal eye field; IPS, intraparietal sulcus; MT+, motion complex area; sPCS, superior precentral sulcus; TPJ, temporo-parietal junction.*

| | | MNI Coordinates | | | |
|---|---|---|---|---|---|
| Anatomical region | Side | x | y | z | t-value |
| Motion complex area (MT+) | R | 44 | -80 | 10 | 8.255 |
| | L | -44 | -66 | 2 | 8.004 |
| Intraparietal sulcus (IPS) | R | 26 | -52 | 52 | 7.418 |
| Frontal eye field (FEF) | R | 38 | -4 | 46 | 5.627 |

*Table 2. Coordinates and t-values of activated regions for the Decision > Confidence contrast.*

While we could not find any regions that clearly showed higher activity for the decision itself, a number of brain areas were activated more for confidence compared to perceptual decision making. Indeed, the Confidence > Decision contrast produced bilateral activations in bilateral anterior prefrontal cortex (aPFC), right dorsolateral prefrontal cortex (dlPFC), dACC, and a collection of regions near the temporo-parietal junction (TPJ) (Figure 5B and Table 3). These results suggest the presence of an extensive network of regions that are more activated by confidence compared to decision processes.

| | | MNI Coordinates | | | |
|---|---|---|---|---|---|
| Anatomical region | Side | x | y | z | t-value |
| Anterior prefrontal cortex (aPFC) | R | 22 | 48 | 32 | 6.859 |
| | L | -30 | 42 | 34 | 6.559 |
| Dorsolateral prefrontal cortex (dlPFC) | R | 40 | 22 | 40 | 5.024 |
| Dorsal anterior cingulate cortex (dACC) | | 0 | 42 | 14 | 5.228 |
| Temporoparietal junction (TPJ) | R | 62 | -46 | 28 | 6.097 |
| | L | -52 | -24 | 18 | 6.445 |

*Table 3. Coordinates and t-values of activated regions in the Confidence > Decision contrast.*

## Discussion

We sought to uncover whether perceptual decision making and confidence are supported by the same or different brain regions. To do so, we were able to largely decorrelate the BOLD regressors for the decision and the confidence periods in our task by including the confidence question only on a fraction of the trials, showing the confidence scale only after the decision has been made, and slowing down both the processes related to perceptual decision making and confidence. We found very extensive overlap between the brain circuits involved in perceptual decision making and confidence. At the same time, only areas related to top-down endogenous attention or stimulus processing were more active during the decision period, whereas a network of frontal, temporo-parietal, and cingulate regions were more active during the confidence period. These findings suggest that computations related to perceptual decision making and confidence are performed by largely overlapping brain circuits though the confidence judgment recruits additional areas.

### Shared brain circuits for perceptual decision making and confidence

Our most striking finding was the extensive overlap between regions that were involved in both perceptual decision making and confidence. Shared activations emerged in frontal (MFG, FEF/SPC, IFS, and anterior insula), parietal (IPS), posterior (MT+), and cingulate (dACC) regions. These findings suggest that the computational and neural bases of perceptual decision making and confidence likely share a large degree of overlap.

Overlap between perceptual decision making and confidence has been demonstrated for a number individual regions in animal studies (Fetsch et al., 2014; Kepecs et al., 2008; Middlebrooks & Sommer, 2012; Resulaj, Kiani, Wolpert, & Shadlen, 2009; So & Stuphorn, 2015). These studies have shown that the same populations of neurons predict both the decision and the confidence in the decision. Nevertheless, studies in animals typically do not separate the decision and confidence computations temporally. For example, in the common opt-out task, animals are allowed to choose a safe option that results in certain but small reward (Hampton, 2001). This common paradigm thus features a complete temporal overlap between perceptual decision making and confidence. However, without a temporal separation between decision and confidence processes, it is difficult to determine whether neurons are explicitly involved in the decision process, the confidence process, or both. Indeed, each one of these possibilities may allow for both the decision and the confidence to be read out since the activity of neurons that code for either is likely to correlate with signal strength and thus predict the other process. By separating decision and confidence processes temporally, our results confirm that many brain regions are explicitly recruited during both decision and confidence computations.

The large degree of overlap also suggests that perceptual decisions and confidence judgments may not be as fundamentally different from each other as sometimes assumed. For example, it is common to refer to these judgments as Type 1 vs. Type 2 because the perceptual decision is about the stimulus (Type 1 judgment), whereas the confidence judgment is about the accuracy of the decision (Type 2 judgment) (Clarke, Birdsall, & Tanner, 1959). This terminology may suggest that confidence judgments are qualitatively different than perceptual decision making. However,

another view that comes from the early days of signal detection theory is that confidence judgments can be conceptualized as simply another perceptual decision that uses different decision criteria on the same underlying signal (Green & Swets, 1966). This conceptualization implies that perceptual decision making and confidence are in fact quite similar. Although our brain results address this question only indirectly, they can be seen as implying that the computations related to perceptual decision making and confidence are indeed more similar than different from each other.

Nevertheless, it must be appreciated that even if perceptual decision making and confidence are computationally similar and rely on mostly the same brain regions, they are not equivalent as demonstrated by many dissociations between them. Such dissociations can be induced by specific stimulus properties (Boldt et al., 2017; Desender et al., 2018; Ferrigno et al., 2017; Lau & Passingham, 2006; Rahnev et al., 2015; Rahnev, Maniscalco, et al., 2011; Samaha et al., 2016; Song et al., 2015; Spence et al., 2015; Vlassova et al., 2014; Zylberberg et al., 2016), damage to early visual cortex such as in blindsight (Weiskrantz, 1986), or transcranial magnetic stimulation (TMS) to the early visual cortex (Hurme, Koivisto, Revonsuo, & Railo, 2017; Rahnev et al., 2013; Rahnev, Maniscalco, Luber, Lau, & Lisanby, 2012). The fact that altering the stimuli or processing at early stages of the visual hierarchy can lead to confidence-accuracy dissociations implies the existence of at least partially separate mechanisms for perceptual decision making and confidence.

Brain areas more active in confidence judgments

Despite the extensive overlap between perceptual decision making and confidence, we also found a number of brain regions that were more active during the confidence compared to the decision period. These regions were located in the prefrontal cortex (aPFC and dlPFC), the dorsal anterior cingulate cortex (dACC), and near the temporoparietal junction (TPJ). Many of these areas have been linked to confidence computations in previous studies (D. Bang & Fleming, 2018; Cortese, Amano, Koizumi, Kawato, & Lau, 2016; Desender, van Opstal, Hughes, & van den Bussche, 2016; Fleming & Dolan, 2012; Fleming, van der Putten, & Daw, 2018; Morales, Lau, & Fleming, 2018; Rahnev et al., 2016; Shekhar & Rahnev, 2018; Wokke, Cleeremans, & Ridderinkhof, 2017) but it has remained unclear whether such activations are stronger than activations that may be caused by the perceptual decision itself. For example, dlPFC has been implicated in both the perceptual decision (Georgiev et al., 2016; Hunt, Behrens, Hosokawa, Wallis, & Kennerley, 2015; Philiastides, Auksztulewicz, Heekeren, & Blankenburg, 2011; Rahnev, Lau, & de Lange, 2011) and confidence (Chiang, Lu, Hsieh, Chang, & Yang, 2014; Lau & Passingham, 2006; Rounis et al., 2010; Shekhar & Rahnev, 2018). Our results suggest that at least one subregion of dlPFC is in fact more active during confidence compared to decision processes.

Why are some regions more active during periods of confidence computation? One possibility is that judgments of confidence serve not only as a guide to the external environment but also reflect our own internal states. For example, repeatedly having low confidence may indicate not only that the stimulus is difficult but that one is losing alertness. Such process of self-evaluation could be supported by regions in medial prefrontal cortex and TPJ that are known to be involved in processing social information (Gusnard, Akbudak, Shulman, & Raichle, 2001; Saxe & Kanwisher,

2003). A related possibility is that confidence evaluations are used to alter one's strategy for doing the task on subsequent trials (Desender et al., 2018) and are thus linked to control processes in the frontal and cingulate regions.

Are there any regions more active for perceptual decision making than confidence?

The only regions that were more active during the decision period than the confidence period were bilateral MT+, right FEF, and right IPS. Although our data do not allow us to pinpoint precisely the role of these areas, it appears likely that none of them were directly involved in the perceptual decision-making process. Most obviously, activity in MT+ was likely driven by the moving dot stimuli. On the other hand, right FEF and right IPS are part of the dorsal attention network (Corbetta & Shulman, 2002), which also typically includes MT+. The right side of this network is known to be dominant (Ruff et al., 2006), which may also explain why we only found activity in right FEF and right IPS. These considerations do not preclude the possibility that some of these regions were involved in the decision itself. This possibility should be tested in studies where stimuli are presented at fixation thus removing the need to engage top-down spatial attention.

The lack of regions that are clearly selective for the decision itself could be interpreted as suggesting that perceptual decision making does not involve unique computational components not present in confidence judgments. Instead, it could be argued that confidence processes completely subsume all decision processes (and include additional components). Further research is needed in order to support (or falsify) this possibility.

## Conclusion

We found that perceptual decision making and confidence show surprisingly high degree of overlap in the brain. In addition, unique brain activations were found for confidence judgments but not for the perceptual decisions. Our results point to the possibility that perceptual decision making and confidence judgments are computed in largely the same neural circuits, with confidence judgments recruiting a number of additional brain areas.

**References**

Bang, D., & Fleming, S. M. (2018). Distinct encoding of decision confidence in human medial prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(23), 6082–6087. https://doi.org/10.1073/pnas.1800795115

Bang, J. W., Shekhar, M., & Rahnev, D. (2018). Sensory noise increases metacognitive efficiency. *Journal of Experimental Psychology: General*.

Boldt, A., de Gardelle, V., & Yeung, N. (2017). The impact of evidence reliability on sensitivity and bias in decision confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(8), 1520–1531. https://doi.org/10.1037/xhp0000404

Chiang, T.-C., Lu, R.-B., Hsieh, S., Chang, Y.-H., & Yang, Y.-K. (2014). Stimulation in the dorsolateral prefrontal cortex changes subjective evaluation of percepts. *PLoS ONE*, *9*(9), e106943. https://doi.org/10.1371/journal.pone.0106943

Clarke, F. R., Birdsall, T. G., & Tanner, W. P. (1959). Two types of ROC curves and definitions of parameters. *The Journal of the Acoustical Society of America*, *31*, 629. https://doi.org/10.1121/1.1907764

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Cortese, A., Amano, K., Koizumi, A., Kawato, M., & Lau, H. (2016). Multivoxel neurofeedback selectively modulates confidence without changing perceptual performance. *Nature Communications*, *7*, 13669. https://doi.org/10.1038/ncomms13669

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), 105–110. https://doi.org/10.1038/nn.3279

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science*, *29*(5), 761–778. https://doi.org/10.1177/0956797617744771

Desender, K., van Opstal, F., Hughes, G., & van den Bussche, E. (2016). The temporal dynamics of metacognition: Dissociating task-related activity from later metacognitive processes. *Neuropsychologia*, *82*, 54–64. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2016.01.003

Ferrigno, S., Kornell, N., & Cantlon, J. F. (2017). A metacognitive illusion in monkeys. *Proceedings of Biological Sciences*, *284*(1862), 20171541. https://doi.org/10.1098/rspb.2017.1541

Fetsch, C. R., Kiani, R., Newsome, W. T., & Shadlen, M. N. (2014). Effects of cortical microstimulation on confidence in a perceptual decision. *Neuron*, *83*(4), 797–804. https://doi.org/10.1016/J.NEURON.2014.07.011

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

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1594), 1338–1349. https://doi.org/10.1098/rstb.2011.0417

Fleming, S. M., & Lau, H. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*, 443. https://doi.org/10.3389/fnhum.2014.00443

Fleming, S. M., Maniscalco, B., Ko, Y., Amendi, N., Ro, T., & Lau, H. (2015). Action-specific

disruption of perceptual confidence. *Psychological Science*, *26*(1), 89–98. https://doi.org/10.1177/0956797614557697

Fleming, S. M., van der Putten, E. J., & Daw, N. D. (2018). Neural mediators of changes of mind about perceptual decisions. *Nature Neuroscience*, 1. https://doi.org/10.1038/s41593-018-0104-6

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*, *10*(4), 843–876. https://doi.org/10.3758/BF03196546

Georgiev, D., Rocchi, L., Tocco, P., Speekenbrink, M., Rothwell, J. C., & Jahanshahi, M. (2016). Continuous theta burst stimulation over the dorsolateral prefrontal cortex and the pre-SMA alter drift rate and response thresholds respectively during perceptual decision-making. *Brain Stimulation*, *9*(4), 601–608. https://doi.org/10.1016/J.BRS.2016.04.004

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and Psychophysics*. New York: John Wiley & Sons Ltd.

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(7), 4259–4264. https://doi.org/10.1073/pnas.071043098

Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(9), 5359–5362. https://doi.org/10.1073/pnas.071600998

Hangya, B., Sanders, J. I., & Kepecs, A. (2016). A mathematical framework for statistical decision

confidence. *Neural Computation*, *28*(9), 1840–1858. https://doi.org/10.1162/NECO_a_00864

Hanks, T. D., & Summerfield, C. (2017). Perceptual decision making in rodents, monkeys and humans. *Neuron*, *93*(1), 15–31. https://doi.org/10.1016/J.NEURON.2016.12.003

Hunt, L. T., Behrens, T. E. J., Hosokawa, T., Wallis, J. D., & Kennerley, S. W. (2015). Capturing the temporal evolution of choice across prefrontal cortex. *ELife*, *4*, e11945. https://doi.org/10.7554/eLife.11945.001

Hurme, M., Koivisto, M., Revonsuo, A., & Railo, H. (2017). Early processing in primary visual cortex is necessary for conscious and unconscious vision while late processing is necessary only for conscious vision in neurologically healthy humans. *NeuroImage*, *150*, 230–238. https://doi.org/10.1016/J.NEUROIMAGE.2017.02.060

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

Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227–231. https://doi.org/10.1038/nature07200

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, *324*(5928), 759–764. https://doi.org/10.1126/science.1169405

Lau, H., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences*,

*103*(49), 18763–18768. https://doi.org/10.1073/pnas.0607716103

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

Maniscalco, B., & Lau, H. (2016). The signal processing architecture underlying subjective reports of sensory awareness. *Neuroscience of Consciousness*, *2016*(1). https://doi.org/10.1093/nc/niw002

Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). *Metacognition: Knowing about knowing*. Cambridge, MA, US: The MIT Press.

Middlebrooks, P. G., & Sommer, M. A. (2012). Neuronal correlates of metacognition in primate frontal cortex. *Neuron*, *75*(3), 517–530. https://doi.org/10.1016/J.NEURON.2012.05.028

Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2360-17.2018

Mueller, S. T., & Weidemann, C. T. (2008). Decision noise: An explanation for observed violations of signal detection theory. *Psychonomic Bulletin & Review*, *15*(3), 465–494. https://doi.org/10.3758/PBR.15.3.465

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Philiastides, M. G., Auksztulewicz, R., Heekeren, H. R., & Blankenburg, F. (2011). Causal role of dorsolateral prefrontal cortex in human perceptual decision making. *Current Biology*, *21*(11), 980–983. https://doi.org/10.1016/J.CUB.2011.04.034

Pouget, A., Drugowitsch, J. D., & Kepecs, A. (2016). Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, *19*(3), 366–374. https://doi.org/10.1038/nn.4240

Rahnev, D., Bahdo, L., de Lange, F. P., & Lau, H. (2012). Prestimulus hemodynamic activity in dorsal attention network is negatively associated with decision confidence in visual perception. *Journal of Neurophysiology*, *108*(5), 1529–1536. https://doi.org/10.1152/jn.00184.2012

Rahnev, D., Koizumi, A., McCurdy, L. Y., D'Esposito, M., & Lau, H. (2015). Confidence leak in perceptual decision making. *Psychological Science*, *26*(11), 1664–1680. https://doi.org/10.1177/0956797615595037

Rahnev, D., Kok, P., Munneke, M., Bahdo, L., de Lange, F. P., & Lau, H. (2013). Continuous theta burst transcranial magnetic stimulation reduces resting state connectivity between visual areas. *Journal of Neurophysiology*, *110*(8), 1811–1821. https://doi.org/10.1152/jn.00209.2013

Rahnev, D., Lau, H., & de Lange, F. P. (2011). Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain. *Journal of Neuroscience*, *31*(29), 10741–10748. https://doi.org/10.1523/JNEUROSCI.1478-11.2011

Rahnev, D., Maniscalco, B., Graves, T., Huang, E., de Lange, F. P., & Lau, H. (2011). Attention induces conservative subjective biases in visual perception. *Nature Neuroscience*, *14*(12), 1513–1515. https://doi.org/10.1038/nn.2948

Rahnev, D., Maniscalco, B., Luber, B., Lau, H., & Lisanby, S. H. (2012). Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence. *Journal of*

*Neurophysiology*, *107*(6), 1556–1563. https://doi.org/10.1152/jn.00985.2011

Rahnev, D., Nee, D. E., Riddle, J., Larson, A. S., & D'Esposito, M. (2016). Causal evidence for frontal cortex organization for perceptual decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(21), 6059–6064. https://doi.org/10.1073/pnas.1522551113

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*, *461*(7261), 263–266. https://doi.org/10.1038/nature08275

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience*, *1*(3), 165–175. https://doi.org/10.1080/17588921003632529

Ruff, C. C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Freeman, E., Haynes, J.-D., … Driver, J. (2006). Concurrent TMS-fMRI and psychophysics reveal frontal influences onhHuman retinotopic visual cortex. *Current Biology*, *16*(15), 1479–1488. https://doi.org/10.1016/J.CUB.2006.06.057

Ryals, A. J., Rogers, L. M., Gross, E. Z., Polnaszek, K. L., & Voss, J. L. (2016). Associative recognition memory awareness improved by theta-burst stimulation of frontopolar cortex. *Cerebral Cortex*, *26*(3), 1200–1210. https://doi.org/10.1093/cercor/bhu311

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

Sanders, J. I., Hangya, B., & Kepecs, A. (2016). Signatures of a statistical computation in the human sense of confidence. *Neuron*, *90*(3), 499–506. https://doi.org/10.1016/J.NEURON.2016.03.025

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Shekhar, M., & Rahnev, D. (2018). Distinguishing the roles of dorsolateral and anterior PFC in visual metacognition. *Journal of Neuroscience*, *38*(22), 5078–5087. https://doi.org/10.1523/JNEUROSCI.3484-17.2018

Shimamura, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology*, *28*(2), 207–218. https://doi.org/10.3758/bf03331979

So, N., & Stuphorn, V. (2015). Supplementary eye field encodes confidence in decisions under risk. *Cerebral Cortex*, *26*(2), 764–782. https://doi.org/10.1093/cercor/bhv025

Song, A., Koizumi, A., & Lau, H. (2015). Awareness, a behavioral method to manipulate metacognitive awareness independent of stimulus. In M. Overgaard (Ed.), *Behavioral Consciousness Research*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199688890.003.0005

Spence, M. L., Dux, P. E., & Arnold, D. H. (2015). Computations underlying confidence in visual perception. *Journal of Experimental Psychology: Human Perception & Performance*, *42*(5), 671–682. https://doi.org/10.1037/xhp0000179

van den Berg, R., Yoo, A. H., & Ma, W. J. (2017). Fechner's law in metacognition: A quantitative model of visual working memory confidence. *Psychological Review*, *124*(2), 197–214.

Retrieved from http://psycnet.apa.org/buy/2017-07341-003

Vlassova, A., Donkin, C., & Pearson, J. (2014). Unconscious information changes decision accuracy but not confidence. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(45), 16214–16218. https://doi.org/10.1073/pnas.1403619111

Weiskrantz, L. (1986). *Blindsight: A case study and its implications*. *Oxford University Press*. Oxford: Clarendon Press.

Wokke, M. E., Cleeremans, A., & Ridderinkhof, K. R. (2017). Sure I'm sure: Prefrontal oscillations support metacognitive monitoring of decision making. *Journal of Neuroscience*, *37*(4), 781–789. https://doi.org/10.1523/JNEUROSCI.1612-16.2016

Yazar, Y., Bergström, Z. M., & Simons, J. S. (2014). Continuous theta burst stimulation of angular gyrus reduces subjective recollection. *PLoS ONE*, *9*(10), e110414. https://doi.org/10.1371/journal.pone.0110414

Ye, Q., Zou, F., Lau, H., Hu, Y., & Kwok, S. C. (2018). Causal evidence for mnemonic metacognition in human precuneus Running title : Mnemonic metacognition in precuneus. *Journal of Neuroscience*, *0600-18*. https://doi.org/10.1523/JNEUROSCI.0660-18.2018

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1594), 1310–1321. https://doi.org/10.1098/rstb.2011.0416

Zylberberg, A., Fetsch, C. R., & Shadlen, M. N. (2016). The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. *ELife*, e17688. https://doi.org/10.7554/eLife.17688