Overlapping and unique neural circuits support perceptual decision making and confidence

Jiwon Yeon, Medha Shekhar, and Dobromir Rahnev

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

Author contributions

All authors designed the studies. J. Yeon and M. Shekhar collected the data. J. Yeon analyzed the data, and J. Yeon and D. Rahnev worked together to interpret the results. J. Yeon drafted the manuscript; D. Rahnev and M. Shekhar revised and commented on it. All authors approved the final version of the manuscript for submission.

Acknowledgements: We thank Farshad Rafiei for help with the data collection. This work was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number R56MH119189 to D.R.

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: j.yeon@gatech.edu

Abstract (Word count: 178)

Perceptual decisions are naturally accompanied by a sense of confidence in the accuracy of the decision. However, it remains unclear whether perceptual decision making and confidence are supported by the same or different neural circuits. To address this question, we conducted two functional MRI (fMRI) experiments in which we dissociated the periods related to perceptual decision making and confidence by either decorrelating their respective regressors or asking for confidence ratings only in the second half of the experiment. We found that perceptual decision making and confidence were supported by large and mostly overlapping brain circuits including frontal, parietal, posterior, and cingulate regions with the results being remarkably consistent across the two experiments. Further, the confidence period recruited a number of unique regions, whereas there was no evidence for the decision period recruiting unique regions not involved in the confidence period. These results suggest that the neural circuits supporting perceptual decision making and confidence have a very high degree of overlap, which suggests largely shared computational mechanisms with unique components for the confidence but not for the 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 and Summerfield, 2017). This process 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 (Metcalfe and Shimamura, 1996; Fleming and Dolan, 2012). Confidence judgments are often referred to as 'metacognitive' because they represent a second-order decision about the accuracy of a first-order decision (Shimamura, 2000; Yeung and Summerfield, 2012; Fleming and Daw, 2017)

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 (Galvin et al., 2003; Rahnev et al., 2012a; Fetsch et al., 2014; Hangya et al., 2016; Pouget et al., 2016; Sanders et al., 2016). The computational similarities suggest that these processes may be supported by similar brain circuits. Support for this notion has come from animal studies demonstrating that the two judgments can be decoded from the same neurons (Kepecs et al., 2008; Kiani and Shadlen, 2009; Middlebrooks and Sommer, 2012; Fetsch et al., 2014; So and Stuphorn, 2015).

At the same time, many other experiments imply the presence of dissociable neural circuits for perceptual decision making and confidence. Behaviorally, confidence judgments can be dissociated from the accuracy of the perceptual decision in both humans (Lau and Passingham,

2006; Rahnev et al., 2011b, 2015; Zylberberg et al., 2012, 2016; Vlassova et al., 2014; Spence et al., 2015; Koizumi et al., 2015; Song et al., 2015; Samaha et al., 2016, 2017; Boldt et al., 2017; Peters et al., 2017; Desender et al., 2018) and monkeys (Ferrigno et al., 2017), whereas other studies suggest that confidence judgments but not perceptual decisions are subject to late metacognitive noise (Mueller and Weidemann, 2008; Jang et al., 2012; De Martino et al., 2013; Maniscalco and Lau, 2016; Rahnev et al., 2016; van den Berg et al., 2017; Shekhar and Rahnev, 2018; Bang et al., 2019). Neurally, studies employing transcranial magnetic stimulation (TMS) delivered to the prefrontal cortex have been able to alter subjects' confidence ratings, while leaving their perceptual decisions unaffected (Rounis et al., 2010; Fleming et al., 2015; Rahnev et al., 2016; Shekhar and Rahnev, 2018). Similar dissociations between the primary decision and confidence have been observed in studies of memory (Yazar et al., 2014; Ryals et al., 2016; Ye et al., 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. To address this question, here we examined the neural circuits supporting perceptual decision making and the confidence using functional MRI (fMRI). In Experiment 1, we dissociated the decision and confidence periods by decorrelating the regressors for each period. In Experiment 2, we dissociated perceptual decision making and confidence further by only including confidence judgments in the second half of the experiment. The two experiments produced remarkably similar results showing a large degree

of overlap in the brain regions supporting perceptual decision making and confidence. Nevertheless, the overlap between the decision and confidence periods was incomplete: both experiments showed several brain areas that were preferentially involved in the confidence period. These findings suggest that perceptual decision making and confidence are supported by largely the same neural circuits but that confidence recruits additional brain areas compared to the perceptual decision.

Methods

Subjects

Twenty-five subjects completed Experiment 1 (12 females, average age = 21.4 years, range = 18-32 years, compensated $35 for participation). Forty-six subjects initially participated in Experiment 2. Seven subjects were excluded because of substantial head movement or very low stimulus sensitivity (d' < 0.2) in the perceptual decision-making task. Hence, 39 subjects were included in the data analyses (23 females, average age = 21.5 years, range = 18-28 years, compensated $50 for participation). Except for one subject, all subjects were right-handed. Subjects 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 subjects were screened for MRI safety and provided informed consent.

Stimulus

In both experiments, subjects judged the direction of motion of white dots (density: 2.4/degree^2; speed: 5°/s) presented in a black circle (Experiment 1: 2° radius; Experiment 2: 3° radius). In Experiment 1, the black circle was positioned either left or right of fixation (its center was 4° from the center of the screen). On each trial, either 4 (low-coherence) or 8% (high-coherence) of the white dots moved coherently and subjects had to determine the direction of motion (either left or right). We used two different coherence levels to ensure that subjects would use both the lower and higher ends of the confidence scales. In Experiment 2, the black circle was positioned at the center of the screen, the coherence level was individually determined for each subject, and the task was to detect whether there was coherent motion

(which was always downward) or not. Each dot had a lifetime between three and 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 (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007).

Task

In Experiment 1, the subjects indicated the direction of motion (left or right) and provided a confidence rating (Figure 1A). 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. The subjects 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 subjects indicated their response via a button press. The subjects gave their responses using the index and middle fingers to indicate left and right direction, respectively. After the subjects 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.

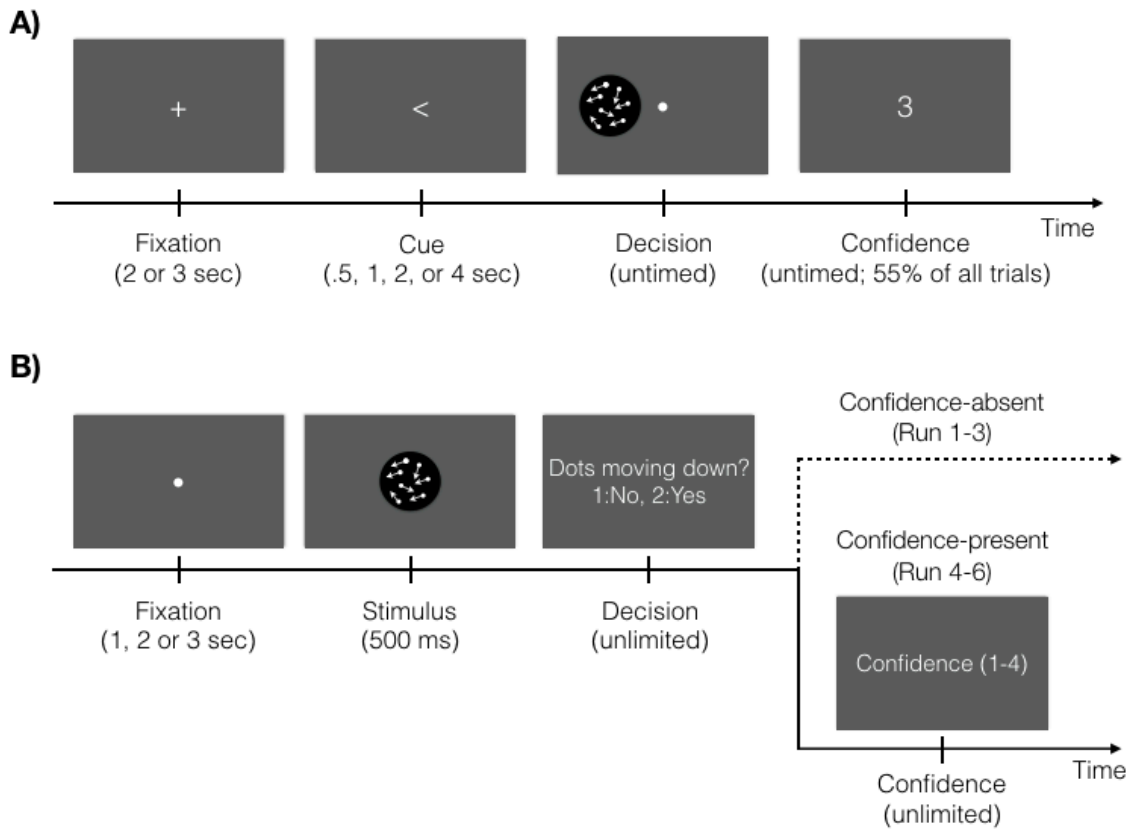

Figure 1. Tasks. (A) Each trial started with a white fixation cross. A cue then indicated the likely location of the following stimulus (either left or right of fixation). In the decision period, the subjects were asked to judge the direction of coherently moving dots. A confidence prompt was presented on 55% of all trials. Three rating scales (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). (B) Each trial started with a white fixation mark, followed by a moving-dots stimulus presented for 500 ms. Subjects indicated whether the dots moved coherently in downward direction or not. In the first half of the experiment (Runs 1-3), subjects were not asked to evaluate their confidence level (Confidence-absent condition). In the second half of the experiment (Runs 4-6), subjects had to rate their confidence level on a 4-point scale (Confidence-present condition). Subjects were first instructed about giving confidence after completing the Confidence-absent condition.

In Experiment 2, subjects indicated whether the moving dot stimuli had an overall coherent motion (always in downward direction) or not (Figure 1B). Each trial began with a fixation mark presented randomly for 1, 2, or 3 seconds. A moving-dot stimulus followed and was presented for 500 ms. After the stimulus offset, the subjects indicated whether they detected downward motion by using their right index and middle fingers to respond 'No' and 'Yes,' respectively. The experiment was separated in two halves. In the first half (Runs 1-3), subjects simply performed the task and were never told anything about evaluating their confidence level. Therefore, we call this the Confidence-absent condition. In the second half of the experiment (Runs 4-6), subjects made their perceptual decision and immediately after were asked to indicate their confidence level. We call this the Confidence-present condition. Importantly, subjects were first informed about the confidence judgments during a short practice run after the end of Run 3 (that is, after completing the Confidence-absent condition). Confidence was always given on a 4-point scale, where the index and the little fingers indicated the lowest and the highest confidence level, respectively. Subjects had unlimited time for both the perceptual decision and confidence judgments.

Procedure

Both experiments started with a short training outside the scanner. During the training, the subjects received instructions on how to perform the task and completed 40 example trials. For Experiment 2, we then determined subjects' threshold coherence level using a 2-down-1-up staircase procedure and presented 60 additional practice trials with the target coherence level. Following the training in both experiments, subjects were positioned in the scanner where we

first collected a structural scan. While collecting the structural scan, subjects were given additional practice trials (excepts for two subjects in Experiment 1 due to technical issues). In Experiment 2, we measured subjects' coherence threshold again using the same 2-down-1-up staircase procedure and then used the threshold in an additional practice block. Based on the performance in this block, the coherence level was adjusted manually for some subjects to ensure between 70 and 75% task accuracy (mean coherence level = 17%, SD = 5%; mean accuracy = 74.1%, SD = 8%). All responses were given with the right hand via an MRI-compatible button box.

Experiment 1 had four runs, each consisting of four 20-trial blocks (for a total of 320 trials). Blocks were separated by 15-second breaks that were indicated by a black fixation cross at the center of the screen. Subjects were given untimed rest periods between runs. Two subjects completed only three runs and one subject completed three runs and a single block from the fourth run. The remaining 22 subjects completed the full four runs.

Experiment 2 consisted of six runs (Run 1-3: Confidence-absent condition; Run 4-6: Confidence-present condition), with each run containing nine 16-trial blocks (a total of 864 trials). In Experiment 2, we manipulated subjects' response bias by presenting them with blocks in which a certain type of stimulus appeared more frequently. The experiment consisted of three types of blocks – a random motion block (randomly moving dots were more frequent), a downward motion block (downward moving dots were more frequent), and a neutral block (both stimuli presented with equal chance). The identity of each block was indicated by a cue that appeared

before each block. Within a run, the order of the blocks was pseudorandomized such that the three types of blocks always appeared together in the same random order. However, since the current study is not focused on the effects of cuing, we only included the neutral blocks in our analyses (288 trials in total). Subjects were given untimed rest period between runs. Two subjects completed only five runs and the rest 37 subjects completed all six runs.

Decorrelating the regressors for perceptual decision making and confidence in Experiment 1

Experiment 1 was initially designed to investigate the shared and unique brain activations produced by the processes of attention, decision, and confidence (Rahnev, 2017; Rahnev et al., 2016). Hence, our design was optimized to decorrelate the regressors for the cue, decision, and confidence periods. This decorrelation was achieved via several design features. First, we presented the cue for .5, 1, 2, or 4 seconds (duration was randomly chosen on each trial), thus decorrelating the regressor for the cue period from the regressors for the decision and confidence periods. Second, we used moving dots as stimuli and placed them in the periphery in order to slow down the perceptual decision-making process and ensure that the decision period was as long as possible. Third, we varied the confidence scale across trials, which slowed down the confidence process thus making that period longer too. Elongating the decision and confidence periods essentially worked to decorrelate the corresponding fMRI regressors (Friston et al., 2007; Ashby, 2011). Finally, we only asked for confidence responses on just over half of all trials. Overall, these design characteristics resulted in low correlations between the cue, decision, and confidence period regressors (average cue-decision correlation: r = -.128, SD

= .054; average cue-confidence correlation: r = -.211, SD = .052; average decision-confidence correlation: r = .201, SD = .089).

However, our attentional manipulation (i.e., the spatial cue indicating the likely location of the stimulus) did not appear to be effective. First, we did not find differences in either accuracy ($t(24) = .60$, $p = .56$) or reaction time ($t(24) = -.17$, $p = .87$) between valid and invalid cues. Second, the fMRI activations associated with the cue period did not reveal any of the known brain areas related to spatial attention (e.g., frontal eye fields and intraparietal sulcus). Instead, the cue period appeared to elicit activity in the default mode network (Supplementary Figure 1A). These results suggest that subjects did not engage spatial attention during the cue period, and therefore this period could not be used to reveal attention-related processes. We suspect that the reason why subjects employed little to no spatial attention during the cue period is that stimulus presentation was already quite long (average = 1.58 seconds) and hence deploying attention in advance was not beneficial for performance. In fact, because the stimulus was presented peripherally, spatial attention appeared to be engaged the most during the decision period. Because of these considerations, we focused our analyses on the overlapping and unique brain activations produced by the decision and confidence periods.

fMRI acquisition and preprocessing

The MRI data were collected on 3T MRI systems (Experiment 1: Trio MRI system, Experiment 2: Prisma-Fit MRI system; Siemens) using 12-channel (Experiment 1) and 32-channel (Experiment 2) head coils. Anatomical images were acquired using T1-weighted sequences (Experiment 1:

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; Experiment 2: MEMPRAGE sequence, FoV = 256 mm; TR = 2530 ms; TE = 1.69 ms; 176 slices; flip angle = 7˚; voxel size = 1.0 × 1.0 × 1.0 mm^3). Functional images were acquired using T2*-weighted gradient echo-planar imaging sequences (Experiment 1: 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; Experiment 2: FoV = 220 mm; TR = 1200 ms; TE = 30 ms; 51 slices; flip angle = 65˚; voxel size = 2.5 × 2.5 × 2.5 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 then preprocessed with the following steps: de-spiking, slice-timing correction, realignment, coregistration, segmentation (segmentation was done before coregistration in Experiment 2), and normalization. The functional images were smoothened with full-width-half-maximum (FWHM) Gaussian kernels (Kernel size: Experiment 1 = 4 mm; Experiment 2 = 6 mm).

Analyses

For all contrast tests across the two experiments, the results of individual subjects were submitted to a group-level t-test. Unless otherwise indicated, statistical tests were based on $p < .05$ corrected for false discovery rate (FDR) and cluster size greater than 80 voxels.

Experiment 1

In a first set of analyses for Experiment 1, we constructed a general linear model (GLM) with 17 regressors for each run. The first five regressors modeled the blood-oxygen level-dependent (BOLD) responses related to the fixation period (spanning the period from onset to offset of the fixation mark), cue period, decision period, catch trials decision period (the same decision period as in the "decision period" regressor but only in catch trials), confidence period, and rest periods between blocks (spanning the period of rest in-between blocks). We modeled catch trials separately since they may invoke additional processes unrelated to perceptual decision making. In addition, we included six regressors related to head movement (three translation and three rotation regressors), four tissue regressors (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 compared the decision and the confidence periods to the cue period using the contrasts Decision > Cue and Confidence > Cue. The cue period was chosen as an active control baseline, but the results remained the same if the fixation period was used as a baseline instead (Supplementary Figures 1B and 2). 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 further examined whether the results obtained from the analyses above depended on the coherence level or the confidence rating scales. To do so, in a separate GLM analysis, we created separate regressors for the decision period that corresponded to the low- and high-

coherence stimuli while keeping all the other regressors the same as in the GLM for the first set of analyses (18 regressors in total). In a separate GLM analysis, we created separate regressors for the confidence period that corresponded to the three different rating scales presented (19 regressors in total). We then ran the same analyses as before for the decision period associated with each coherence level and for the confidence period associated with each confidence scale.

To ensure that the small correlation between the decision and confidence regressors did not drive our results, 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 confidence period regressor from the BOLD time course and performed the Decision > Cue contrast on the residuals. In this way, we could remove the unique variability explained exclusively by the confidence regressor from the BOLD signal (for more details, see Supplementary Figure 3). Then, we built a GLM for the Decision > Cue contrast, which was equivalent to the main 17-regressor GLM, except that we did not include the confidence period regressor. In the second analysis, we regressed out the decision period regressor and extracted the activations for the Confidence > Cue contrast in an equivalent way.

Furthermore, we wanted to confirm that our results were not due to an early start of the confidence processes. By definition, decision processes could only occur up to the time of the button press indicating a subject's choice; any processes taking place afterwards would serve to evaluate the decision. Therefore, although decision processes are unlikely to have taken place during the confidence period, it is possible that confidence processes could start before the

decision-related button press. In a separate analysis, we defined the decision period as ending 150 ms before the decision button press, and the confidence period as starting immediately afterwards. We chose 150 ms as simple auditory reaction times are known to take on average about 160 ms, whereas simple visual reaction times take on average about 190 ms (Welford, 1980). Given the conduction delays in auditory and visual cortex, 150 ms is thus a conservative estimate of how long people take to make a button press after an internal decision has been made.

The design of Experiment 1 was optimized for separating decision- and confidence-related fMRI regressors. This design did not allow us to compute metacognitive performance or other confidence-related variables because of the inclusion of three different scales for the confidence response (and the corresponding small number of trials related to each scale). Nevertheless, for the purposes of providing summary statistics of subjects' performance, 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. We then averaged the transformed confidence responses to report summary statistics.

Experiment 2

We generated two different GLMs in Experiment 2 for the Confidence-absent (Runs 1-3) and Confidence-present (Runs 4-6) conditions. For the Confidence-absent condition, we included regressors related to the fixation period, decision period, and between-block periods. For the

Confidence-present condition, we included regressors of the fixation, decision, confidence, and between-block periods. In addition, we included six head-motion related regressors (three translation and three rotation regressors), four tissue regressors (white matter, cerebrospinal fluid and bone, soft tissues, and air and background), and a constant term for all sessions.

We first compared the brain responses related to the decision periods in Confidence-absent ($\text{Decision}_{\text{conf-abs}}$) and Confidence-present ($\text{Decision}_{\text{conf-pres}}$) conditions to the fixation period ($\text{Decision}_{\text{conf-abs}}$ > Fixation and $\text{Decision}_{\text{conf-pres}}$ > Fixation). Also, to check whether we would observe similar activation patterns for the confidence period, we contrasted the regressors related to the confidence period in the Confidence-present condition to the fixation period (Confidence > Fixation). Furthermore, we examined the unique neural responses related to each decision and confidence periods. Similar to the contrast tests in Experiment 1, we compared the decision period of the two conditions separately with the confidence period from the Confidence-present condition ($\text{Decision}_{\text{conf-abs}}$ vs. Confidence, and $\text{Decision}_{\text{conf-pres}}$ vs. Confidence).

Data and code

All data and codes for the behavioral analyses are freely available at https://osf.io/pn283 and have also been uploaded to the Confidence Database (Rahnev et al., 2020). In addition, unthresholded fMRI maps have been are uploaded in NeuroVault (Gorgolewski et al., 2015) and can be accessed at https://neurovault.org/collections/6897.

Results

We investigated whether perceptual decision making and confidence are supported by the same or different brain regions using fMRI. In Experiment 1, we decorrelated regressors for the decision and confidence periods in order to be able to mathematically determine their unique contributions to BOLD activity. In Experiment 2, subjects only made perceptual decisions in the first half of the experiment, but provided both perceptual and confidence judgments in the second half. This design allowed us to ensure that almost no confidence processes took part during the period of decision making in the first half of Experiment 2. We then compared the brain activations related to perceptual decision making and confidence in both experiments.

Experiment 1

Behavioral results

Task accuracy of the subjects was 67.6% correct (SD = 9.91) and the average response time was 1.58 seconds (SD = 0.418). Performance was higher for high-coherence (accuracy = 69.5%, SD = 12.1) than for low-coherence stimuli (accuracy = 65.6%, SD = 9.39; paired t-test: $t(24) = 2.28$, $p = .032$). Similarly, reaction time was faster for the high-coherence task (RT = 1.54 sec, SD = .399) compared to the low-coherence task (RT = 1.63 sec, SD = 0.444; paired t-test: $t(24) = -3.81$, $p = 8.47 \times 10^{-4}$). To compare the confidence ratings across the three different scales (we interleaved 2-, 3-, and 4-point scales), we mapped each scale on the [0, 1] interval such that the lowest confidence rating was always changed to 0 and the highest confidence rating was always changed to 1 (see Methods). The average transformed confidence value across all rating scales was .649 (SD = .191). Subjects were more confident on correct trials (mean confidence = .672,

SD = .191) than on incorrect trials (mean confidence = .592, SD = .193; paired t-test: t(24) = 5.217, $p = 2.40 \times 10^{-5}$). These results suggest that the subjects were able to perform the task as intended and provide appropriate confidence ratings.

Shared activations between perceptual decision making and confidence

To explore the possible overlap between the 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 the decision and confidence periods activated a similar set of fronto-parieto-posterior brain regions (Figure 2A, B). In order to examine the amount of overlap, we created a map of the intersection of the decision- and confidence-period activations (Figure 2C). This map showed extensive bilateral activations in a number of areas including the dorsolateral prefrontal cortex (dlPFC), frontal eye fields (FEF), intraparietal sulcus (IPS), and right temporoparietal junction (TPJ). In addition, strong bilateral activations were observed in the dorsal anterior cingulate cortex (dACC), anterior insula, and occipitotemporal cortex (OTC) which includes the motion complex area (MT+). We also observed bilateral somatosensory and motor cortices activity, which were 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 Supplementary Table 1.

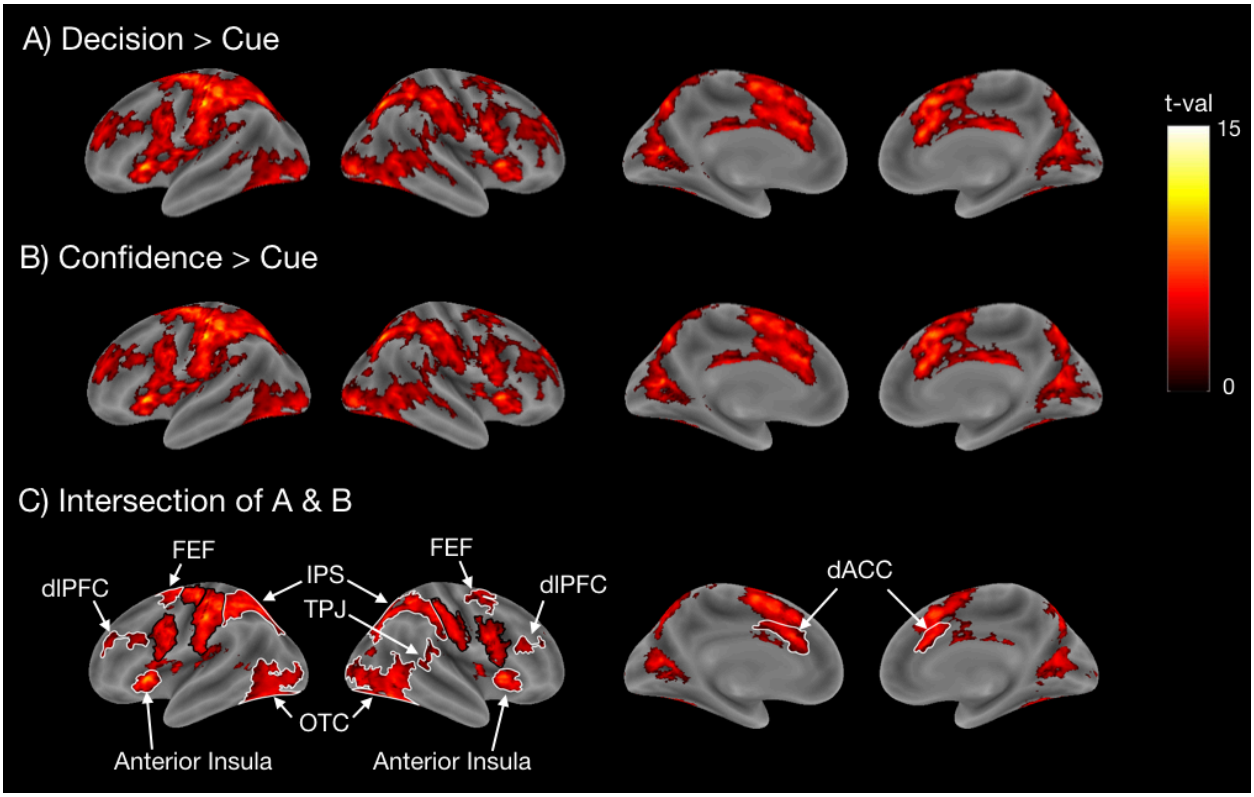

Figure 2. Shared activity between perceptual decision making and confidence. (A) Decision-related brain activity obtained from the contrast Decision > Cue. (B) Confidence-related brain activity obtained from the contrast Confidence > Cue. (C) Intersection between the activation maps for perceptual decision 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 somatosensory and motor cortices. dACC, dorsal anterior cingulate cortex; dlPFC, dorsolateral prefrontal cortex; FEF, frontal eye field; IPS, intraparietal sulcus; OTC, occipitotemporal cortex; TPJ, temporoparietal junction.

These results suggest 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 period's 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 the confidence period's effects were completely removed from the decision regressor (see Supplementary Figure 3). Despite this removal, we found a very similar set of activations for the Decision > Cue contrast (Figure 3A). Similarly, in a different set of analyses, we regressed out the decision period's activity from the BOLD time course and created a map for the Confidence > Cue contrast using the residuals. Again, we found a very similar pattern of activations (Figure 3B). Finally, we examined the intersection of the two sets of activations in these new analyses and found activations in each of the regions found in our previous analysis, such as dlPFC, FEF, IPS, dACC, anterior insula, OTC, right TPJ, and the somatosensory and motor cortices (Figure 3C).

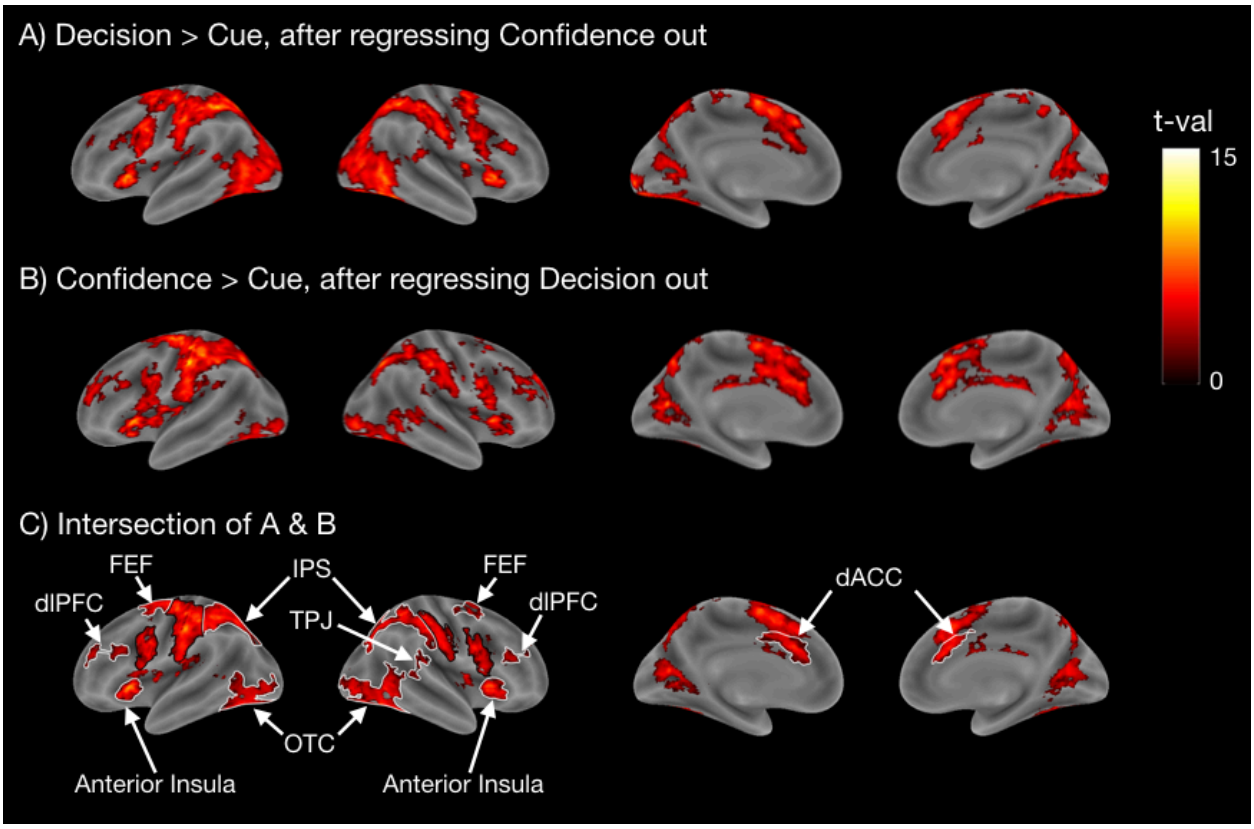

Figure 3. Control analyses on the shared activity between perceptual decision making and confidence. (A) Decision-related brain activity obtained from the contrast Decision > Cue after confidence-related activations had been removed in a previous regression. (B) Confidence-related brain activity obtained 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 2.

We further performed a number of control analyses to ensure the robustness of our finding. First, we confirmed that our results were not due to the chosen baseline (i.e., the cue period) by running the same contrasts but using the fixation period (spanning the period from onset of the fixation cross until the onset of the cue) as the baseline. We observed similar activation patterns for both the decision and the confidence periods, as well as the intersection of the two sets of activations (Supplementary Figure 2). Second, we confirmed that our results were independent of the coherence level used and that similar results were obtained when only the

low or only the high coherence levels were analyzed (Supplementary Figure 4). Third, we verified that our results did not depend on the confidence scale used and that the same results were obtained for each of the three different rating scales (Supplementary Figure 5). Finally, a potential concern in all of these analyses is that confidence processes may have started before the button press with which subjects indicated their choice. In particular, there is a delay between the time when an internal decision is made and when a button press occurs. To account for this type of delay, we performed another control analysis in which the end of the decision period and the start of the confidence period were moved to 150 ms before the choice button press. Based on previous research on simple visual and auditory reaction times (Welford, 1980), 150 ms was chosen as a conservative estimate of how long people take to make a button press after an internal decision has been made. The results again revealed similar activation patterns for the decision and the confidence periods, as well as the intersection of the two activations (Supplementary Figure 6).

Unique activations for decision and confidence processes

The analyses so far point to a considerable overlap between the brain regions involved in perceptual decision making and confidence. However, despite this overlap, it is possible that decision 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 the decision period (using the contrast Decision > Confidence). We found activity in bilateral OTC, right FEF, and right superior parietal lobule (SPL) (Figure 4A and Supplementary Table 2). Notably, the large bilateral activation in OTC that includes MT+ is likely

driven by the presence of moving-dot stimuli during the decision but not during the confidence period and therefore may not reflect decision processes that are independent of the moving-dots stimulus. Similarly, right FEF and right SPL are part of the dorsal attention network (Corbetta and 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).

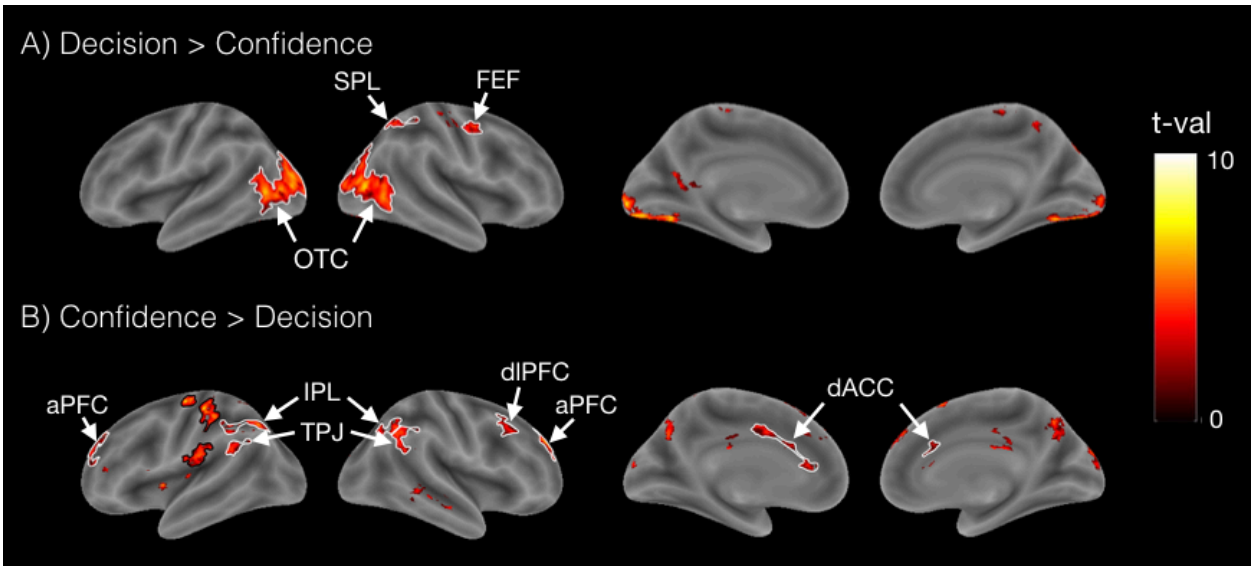

Figure 4. Unique activity for perceptual decision making and confidence. (A) Regions showing higher activations for the decision than the confidence period. (B) Regions showing higher activations for the confidence than the decision period. Colors indicate t-values. aPFC, anterior prefrontal cortex; dACC, dorsal anterior cingulate cortex; dlPFC, dorsolateral prefrontal cortex; FEF, frontal eye field; IPL, Inferior parietal lobule; OTC, occipitotemporal cortex; SPL, Superior parietal lobule; TPJ, Temporoparietal junction.

Although we could not find any regions that clearly showed higher activity for the decision process itself, a number of brain areas were activated more for the confidence compared to the decision period. Indeed, the Confidence > Decision contrast produced activations in bilateral aPFC, right dlPFC, dACC, and a collection of regions near the temporo-parietal junction (TPJ) (Figure 4B and Supplementary Table 3). These results suggest the presence of an extensive network of regions that are more activated during the confidence period.

Finally, the patterns of results above remained the same when we analyzed separately the different coherence levels and confidence rating scales. Specifically, regardless of whether low- or high-coherence trials were analyzed, the decision period resulted in higher activations than the confidence period in bilateral OTC, right FEF, and right SPL (Supplementary Figure 7). Similarly, when we analyzed the 2-, 3-, and 4-point scale separately, the confidence period resulted in higher activations than the decision period in bilateral aPFC, right dlPFC, dACC, and a collection of regions near TPJ (Supplementary Figure 8). Thus, our results are robust to manipulations of both the stimulus strength and the confidence scale.

Experiment 2

In Experiment 1, we decorrelated the regressors associated with the periods of decision making and confidence. However, in order to ensure that the obtained activations reflect processes that are specific to decision and confidence, it is critical that decision and confidence processes are constrained to the decision and confidence periods, respectively. By definition, decision processes can only occur up to the time of the button press indicating a subject's choice; any

processes taking place afterwards would serve to evaluate the decision that was already made. Therefore, decision processes are unlikely to have taken place during the confidence period. However, it is possible that confidence processes could start before the decision-related button press and that activity during the decision period reflects both processes related to decision-making and confidence. To address this issue, we conducted Experiment 2 in which subjects completed the decision-making task without confidence for the first half of the experiment, and only rated confidence in the second part of the experiment (Figure 1B). We call the first half of the experiment the "Confidence-absent" condition and the second half of the experiment the "Confidence-present" condition.

Behavioral results

The overall task accuracy was 74.2% (SD = .079) and the average response time was 728 ms (SD = 249). Accuracy was lower in the Confidence-absent condition (accuracy = 73.2%, SD = .078) than in the Confidence-present condition (accuracy = 75.3%, SD = .088; paired t-test: t(38) = -2.52, p = .016), potentially suggesting the presence of a small learning effect. Further, response time in the Confidence-absent condition was faster (RT = 664 ms, SD = 282) than in the Confidence-present condition (RT = 794 ms, SD = 254; paired t-test: t(38) = -3.91, $p = 3.65 \times 10^{-4}$), which is likely due to slowing down associated with the presence of confidence ratings in the Confidence-present condition. The average confidence rating across all trials was 2.866 (SD = .459) and average confidence RT was 516 ms (SD = 168). The confidence level was higher in correct trials (2.953, SD = .454) than in incorrect trials (2.546, SD = .491; paired t-test: t(38) = 9.80, $p = 6 \times 10^{-12}$). Response time for confidence judgment was slightly faster in correct trials

(509 ms, SD = 166) than in incorrect trials (541 ms, SD = 176; paired t-test: t(38) = -2.98, *p* = .005). These results suggest that the subjects performed the task well and evaluated their confidence level appropriately.

Shared activations between perceptual decision making and confidence

As in Experiment 1, we first examined the activity related to the decision and confidence periods separately. We first examined the brain activity for the decision period from the Confidence-absent condition ($\text{Decision}_{\text{conf-abs}}$, which does not contain any confidence-related signals), the decision period from the Confidence-present condition ($\text{Decision}_{\text{conf-pres}}$, which, like in Experiment 1, may contain some confidence-related signals), and the confidence period from the Confidence-present condition (Confidence).

We observed largely overlapping patterns of activation in fronto-parieto-posterior regions for all three of these periods (Figure 5A-C) even though $\text{Decision}_{\text{conf-abs}}$ was free of confidence-related processes. Further, as in Experiment 1, we created intersection maps between $\text{Decision}_{\text{conf-abs}}$ and Confidence (Figure 5D and Table 4), as well as between $\text{Decision}_{\text{conf-pres}}$ and Confidence (Figure 5E and Supplementary Table 5). Critically, the intersection map for $\text{Decision}_{\text{conf-abs}}$ and Confidence was extremely similar to what we observed in Experiment 1: there were shared activations in bilateral dACC, dlPFC, FEF, IPS, OTC, anterior insula, and right TPJ. The intersection map for $\text{Decision}_{\text{conf-pres}}$ and Confidence, which is conceptually identical to the analyses from Experiment 1, also showed activations in the same set of brain regions. These

results strongly suggest that the large overlap observed in Experiment 1 is not simply due to confidence processes occurring in both the decision and confidence periods.

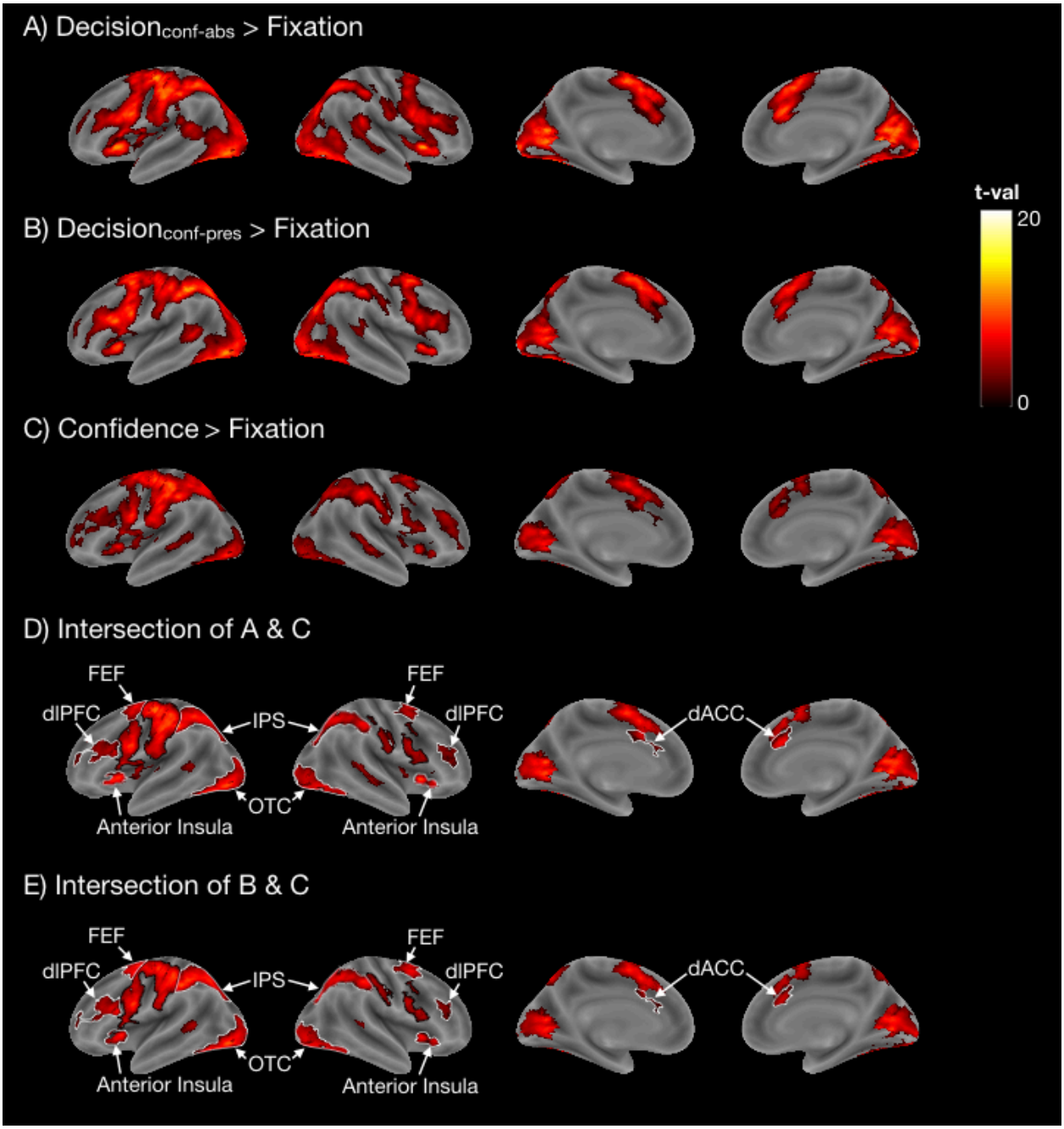

Figure 5. Overlapping brain activations for perceptual decision making and confidence in Experiment 2. (A) Decision-related activity in the Confidence-absent condition (Decision$_{conf\text{-}abs}$ > Fixation). Note that the decision period in the Confidence-absent condition could not be contaminated by confidence-related signals because subjects never rated confidence in that

condition. (B) Decision-related activity in the Confidence-absent condition ($Decision_{conf\text{-}pres}$ > Fixation). (C) Confidence-related activity (Confidence > Fixation). (D,E) Intersection maps for the decision-related and confidence-related activations. Colors indicate t-values. The t-values in (D) and (E) are the average t-values from the corresponding maps. The black borders delineate the somatosensory and motor cortices. dACC, dorsolateral anterior cingulate cortex; dlPFC, dorsolateral prefrontal cortex; FEF, frontal eye field; IPS, intraparietal sulcus; OTC, occipitotemporal cortex.

Unique activations for decision and confidence processes

As in Experiment 1, we examined the brain regions that were preferentially activated by either decision- or confidence-related processes. We first explored the regions that were activated more strongly for decision-related processes by performing the contrasts $Decision_{conf\text{-}abs}$ > Confidence and $Decision_{conf\text{-}pres}$ > Confidence. As in Experiment 1, the strongest activation was in bilateral OTC (Figure 6A, B and Supplementary Table 6). As we already discussed in Experiment 1, this activation is likely generated by the moving dot stimuli and not by the decision process itself. It should be noted that no other activation for the Decision > Confidence contrasts was consistent between Experiments 1 and 2, suggesting that perceptual decision making may not consistently drive any brain area more than confidence.

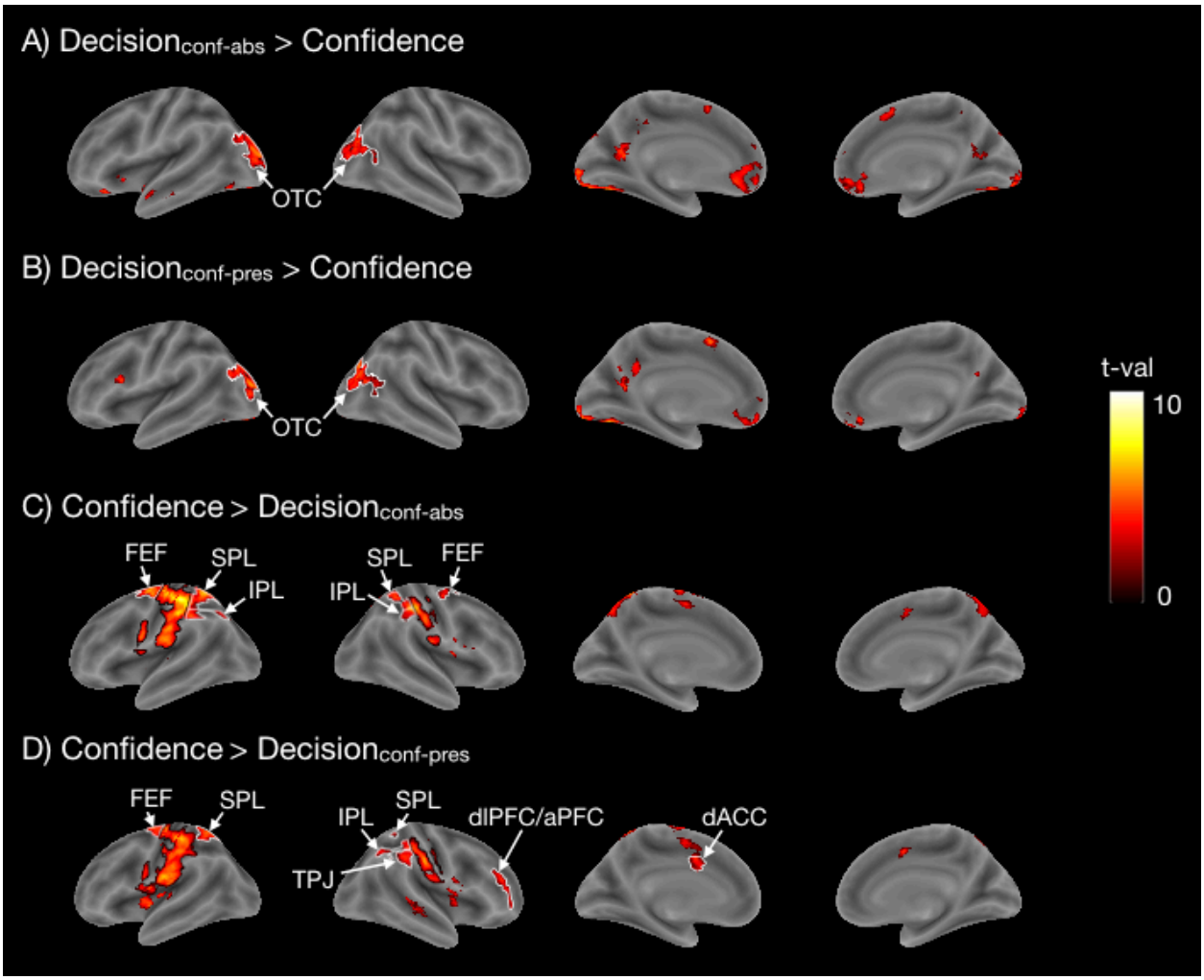

Figure 6. Unique brain regions activated for perceptual decision making and confidence judgment processes in Experiment 2. (A,B) Regions showing higher activations for Decision$_{conf-abs}$ and Decision$_{conf-pres}$ than Confidence$_{conf-pres}$. (C,D) Regions showing higher activations for Confidence$_{conf-pres}$ than Decision$_{conf-abs}$ and Decision$_{conf-pres}$. Black borders delineate the pre- and post-central gyri where the somatosensory and motor cortices are located. Colors indicate t-values. aPFC, anterior prefrontal cortex; dACC, dorsal anterior cingulate cortex; dlPFC, dorsolateral prefrontal cortex; FEF, frontal eye field; IPL, Inferior parietal lobule; IPS, intraparietal sulcus; OTC, occipitotemporal cortex; SPL, Superior parietal lobule; TPJ, Temporoparietal junction.

Finally, as in Experiment 1, we examined what brain regions are more activated for confidence compared to perceptual decision making. To do so, we contrasted the confidence period with the decision period in both the Confidence-absent (Confidence > Decision$_{conf-abs}$) and

Confidence-present conditions (Confidence > $\text{Decision}_{\text{conf-pres}}$). Confidence > $\text{Decision}_{\text{conf-abs}}$ (Figure 6C and Supplementary Table 7) revealed activations in bilateral IPL, SPL, and FEF. Confidence > $\text{Decision}_{\text{conf-pres}}$ (Figure 6D and Supplementary Table 8) showed mostly similar activations in bilateral SPL and FEF; left dACC; and right aPFC/dlPFC, TPJ, and IPL. In addition, we observed activations in motor and somatosensory cortex that are likely related to the button presses and not the confidence process itself. In control analyses, we performed the same contrasts but not only for the Neutral blocks (where neutral cues were given) but for all trials (including blocks with predictive cues). This analysis tripled the number of trials and thus provided more power. Consequently, we observed more extensive activations, including extensive activations in bilateral aPFC/dlPFC and anterior insula (Supplementary Figure 9). Overall, there were a number of similarities between the areas that were more activated for confidence in Experiments 1 and 2, including activations in aPFC/dlPFC, ACC, IPL, and TPJ.

Discussion

We sought to uncover whether perceptual decision making and confidence are supported by the same or different brain regions. In two experiments, we found extensive overlap between the brain circuits involved in the decision and the confidence periods. At the same time, a network of frontal, temporo-parietal, and cingulate regions were more active during the confidence period but no regions were consistently more active for perceptual decision making than confidence. These findings suggest that 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

We found extensive overlap between regions that were active during the decision and the confidence periods. In the results of the both experiments, shared activations emerged in frontal (FEF, dlPFC, and anterior insula), parietal (IPL, SPL, and TPJ), posterior (OTC), and cingulate (dACC) regions. These findings suggest that the neural bases of perceptual decision making and confidence share a large degree of overlap.

Overlap between perceptual decision making and confidence has been demonstrated for a number individual regions in animal studies (Kepecs et al., 2008; Resulaj et al., 2009; Middlebrooks and Sommer, 2012; Fetsch et al., 2014; So and 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 features a complete temporal overlap between perceptual decision making and confidence, making it challenging to ensure that neural activity is specifically related to decision or confidence.

The large degree of overlap observed in the current studies 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 et al., 1959). This terminology may suggest that confidence judgments are completely 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 and 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 though we observed similar activation patterns for decision and confidence, the computations that underlie these two processes may still be

very different. Such differences are implied by accuracy-confidence dissociations (Lau and Passingham, 2006; Rahnev et al., 2011b, 2015; Vlassova et al., 2014; Song et al., 2015; Spence et al., 2015; Zylberberg et al., 2016; Samaha et al., 2016; Boldt et al., 2017; Ferrigno et al., 2017; Desender et al., 2018), neurological phenomena such as in blindsight (Weiskrantz, 1986), and studies that delivered transcranial magnetic stimulation (TMS) to the early visual cortex and found separable effects on accuracy and confidence (Rahnev et al., 2012b, 2013; Hurme et al., 2017). Moreover, a recent study provided direct evidence that sensory neural signals are indeed used in different ways in making perceptual decisions and confidence judgments (Peters et al., 2017). The dissociation between the decision and confidence processes that previous studies have been shown implies the existence of at least partially separate neural computations for perceptual decision making and confidence.

Brain areas more active in confidence judgments

Despite the extensive overlap between the decision and the confidence periods, we also found a number of brain regions that were more active during the confidence compared to the decision period. These regions appeared in both experiments and were located in the prefrontal cortex (aPFC and dlPFC), the dorsal anterior cingulate cortex (dACC), and near the temporoparietal junction (TPJ and IPL). Many of these areas have been linked to confidence computations in previous studies (Fleming and Dolan, 2012; Cortese et al., 2016; Desender et al., 2016; Rahnev et al., 2016; Wokke et al., 2017; Bang and Fleming, 2018; Fleming et al., 2018; Morales et al., 2018; Shekhar and Rahnev, 2018) 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 (Philiastides et al., 2011; Rahnev et al., 2011a; Hunt et al., 2015; Georgiev et al., 2016) and confidence (Lau and Passingham, 2006; Rounis et al., 2010; Chiang et al., 2014; Shekhar and Rahnev, 2018). Our results suggest that at least a 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 areas in and around TPJ that are known to be involved in processing social information (Gusnard et al., 2001; Saxe and Kanwisher, 2003). Further, 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.

<u>Are there any regions more active for perceptual decision making than confidence?</u>

The only region, that was more active during the decision period than the confidence period across the two experiments was bilateral occipitotemporal cortex (OTC). However, activity in OTC is likely to be driven by the moving dot stimuli and not by the decision-making process itself. We also observed several activations that were specific to either Experiment 1 or 2 and thus are unlikely to be critically involved in perceptual decision making independent of the

specifics of each experimental paradigm. For example, right FEF and right SPL, which were activated in Experiment 1, are a part of the dorsal attention network (Corbetta and Shulman, 2002) and may reflect the need for deploying spatial attention in Experiment 1 (but not in Experiment 2 where all stimuli were at fixation). The right side of the dorsal attention network is known to be dominant (Ruff et al., 2006), which may also explain why we only found activity in right FEF and right SPL in Experiment 1.

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 process 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.

Separating perceptual decision making and confidence

In Experiment 1, we examined the separate brain activations for perceptual decision making and confidence by decorrelating the fMRI regressors associated with the decision and confidence periods. Further, we performed several control analyses such as a stepwise regression designed to reveal the unique contribution of the decision and confidence periods (Supplementary Figure 3) and still found the same results (Figure 3). However, a more challenging issue in Experiment 1 was to ensure that decision and confidence processes were temporally confined to the decision period (i.e., the period from stimulus onset to the decision button press) and confidence period (i.e., the period from decision button press to confidence

button press), respectively. In particular, it is possible that the confidence process might have started before the decision button press with “early” confidence computations occurring during the decision period and “late” confidence computations occurring during the confidence period. We tried to minimize such possibility by presenting the confidence scale, based on which confidence responses had to be evaluated, only after the decision button press. Besides this design consideration, there are several reasons to suspect that putative early confidence computations taking place during the decision period must have been relatively insubstantial. First, the Confidence > Decision contrast revealed a set of brain regions that has typically been found to be engaged in confidence computations (Fleming and Dolan, 2012; Cortese et al., 2016; Desender et al., 2016; Rahnev et al., 2016; Wokke et al., 2017; Bang and Fleming, 2018; Fleming et al., 2018; Morales et al., 2018; Shekhar and Rahnev, 2018). If the effects of early confidence computations were larger than the effect of late confidence computations, the Confidence > Decision contrast result would reveal a much smaller network of regions. Similarly, if early confidence signals were substantial, one may also expect that the Decision > Confidence contrast would reveal activations related to these putative early confidence signals but no such activations were found. Finally, the activations found in the Decision > Cue contrast revealed a similar set of brain regions that has previously been reported for perceptual decision making in the absence of confidence ratings anywhere in the experiment (Huettel et al., 2005; Kayser et al., 2009). Taken together, these results suggest that putative early confidence signals that could have occurred during the decision period must have been relatively insubstantial and are unlikely to meaningfully contribute to our results.

Nevertheless, the possibility of confidence computations partially occurring during the decision period in Experiment 1 motivated us to use an alternative design in Experiment 2 that can provide a greater separation between the processes related to perceptual decision making and confidence. We achieved such separation by never asking for confidence in the first part of the experiment. Nevertheless, we observed similar results across both experiments, thus strongly suggesting that perceptual decision making and confidence are indeed supported by largely overlapping brain regions. The inclusion of two experiments with different designs also helped establish which activations are robust across non-trivial changes in the experimental paradigm and which activations appear to be specific to the experimental design used in each experiment. We suggest that future research should adopt a similar approach where the overlap between the brain areas supporting various processes should be examined across several different contexts.

Conclusion

In two different experiments, we found a surprisingly high degree of overlap in the brain regions that support perceptual decision making and confidence. In addition, unique brain activations were found for the confidence judgments but not for the perceptual decisions. Our results suggest the possibility that perceptual decision making and confidence judgments are computed in largely the same neural circuits, with the confidence judgments recruiting a number of additional brain areas.

References

Ashby FG (2011) Statistical analysis of fMRI data. MIT Press.

Badre D, D'Esposito M (2009) Is the rostro-caudal axis of the frontal lobe hierarchical? Nat Rev Neurosci 10:659–669.

Bang D, Fleming SM (2018) Distinct encoding of decision confidence in human medial prefrontal cortex. Proc Natl Acad Sci U S A 115:6082–6087 Available at: http://www.ncbi.nlm.nih.gov/pubmed/29784814.

Bang JW, Shekhar M, Rahnev D (2019) Sensory noise increases metacognitive efficiency. J Exp Psychol Gen 148:437–452 Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/xge0000511.

Boldt A, de Gardelle V, Yeung N (2017) The impact of evidence reliability on sensitivity and bias in decision confidence. J Exp Psychol Hum Percept Perform 43:1520–1531.

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436 Available at: http://color.psych.ucsb.edu/psychtoolbox.

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 Gray M, ed. PLoS One 9:e106943 Available at: http://dx.plos.org/10.1371/journal.pone.0106943.

Clarke FR, Birdsall TG, Tanner WP (1959) Two types of ROC curves and definitions of parameters. J Acoust Soc Am 31:629–630 Available at: http://asa.scitation.org/toc/jas/31/5.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215 Available at: http://www.nature.com/articles/nrn755.

Cortese A, Amano K, Koizumi A, Kawato M, Lau H (2016) Multivoxel neurofeedback selectively modulates confidence without changing perceptual performance. Nat Commun 7:13669 Available at: http://www.nature.com/doifinder/10.1038/ncomms13669.

De Martino B, Fleming SM, Garrett N, Dolan RJ (2013) Confidence in value-based choice. Nat Neurosci 16:105–110 Available at: http://www.nature.com/articles/nn.3279.

Desender K, Boldt A, Yeung N (2018) Subjective confidence predicts information seeking in decision making. Psychol Sci 29:761–778 Available at: http://journals.sagepub.com/doi/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 Available at: https://www.sciencedirect.com/science/article/pii/S0028393216300033.

Ferrigno S, Kornell N, Cantlon JF (2017) A metacognitive illusion in monkeys. Proc Biol Sci 284:20171541 Available at: http://www.ncbi.nlm.nih.gov/pubmed/28878068.

Fetsch CR, Kiani R, Newsome WT, Shadlen MN (2014) Effects of cortical microstimulation on confidence in a perceptual decision. Neuron 83:797–804 Available at: https://www.sciencedirect.com/science/article/pii/S0896627314006229.

Fleming SM, Daw ND (2017) Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. Psychol Rev 124:91–114.

Fleming SM, Dolan RJ (2012) The neural basis of metacognitive ability. Philos Trans R Soc Lond B Biol Sci 367:1338–1349 Available at: http://www.ncbi.nlm.nih.gov/pubmed/22492751.

Fleming SM, Maniscalco B, Ko YD, Amendi N, Ro T, Lau H (2015) Action-specific disruption of

perceptual confidence. Psychol Sci 26:89–98 Available at: http://journals.sagepub.com/doi/10.1177/0956797614557697.

Fleming SM, van der Putten EJ, Daw ND (2018) Neural mediators of changes of mind about perceptual decisions. Nat Neurosci:1 Available at: https://doi.org/10.1038/s41593-018-0104-6.

Friston KJ, Ashburner J, Kiebel S, Nichols T, Penny WD (2007) Statistical parametric mapping : the analysis of funtional brain images. Elsevier.

Galvin SJ, Podd J V., Drga V, Whitmore J (2003) Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. Psychon Bull Rev 10:843–876 Available at: http://www.springerlink.com/index/10.3758/BF03196546.

Georgiev D, Rocchi L, Tocco P, Speekenbrink M, Rothwell JC, 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 Stimul 9:601–608 Available at: https://www.sciencedirect.com/science/article/pii/S1935861X16300547?via%3Dihub.

Gorgolewski KJ, Varoquaux G, Rivera G, Schwarz Y, Ghosh SS, Maumet C, Sochat V V., Nichols TE, Poldrack RA, Poline J-B, Yarkoni T, Margulies DS (2015) NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. Front Neuroinform 9.

Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: John Wiley & Sons Ltd.

Gusnard DA, Akbudak E, Shulman GL, Raichle ME (2001) Medial prefrontal cortex and self-

referential mental activity: Relation to a default mode of brain function. Proc Natl Acad Sci 98:4259–4264 Available at: https://doi.org/10.1073/pnas.071043098.

Hampton RR (2001) Rhesus monkeys know when they remember. Proc Natl Acad Sci U S A 98:5359–5362 Available at: http://www.pnas.org/content/98/9/5359.long.

Hangya B, Sanders JI, Kepecs A (2016) A mathematical framework for statistical decision confidence. Neural Comput 28:1840–1858 Available at: http://www.mitpressjournals.org/doi/10.1162/NECO_a_00864.

Hanks TD, Summerfield C (2017) Perceptual decision making in rodents, monkeys and humans. Neuron 93:15–31 Available at: https://www.sciencedirect.com/science/article/pii/S0896627316309424.

Huettel SA, Song AW, McCarthy G (2005) Decisions under Uncertainty: Probabilistic Context Influences Activation of Prefrontal and Parietal Cortices. J Neurosci 25:3304–3311.

Hunt LT, Behrens TEJ, Hosokawa T, Wallis JD, Kennerley SW (2015) Capturing the temporal evolution of choice across prefrontal cortex. Elife 4:e11945 Available at: https://cdn.elifesciences.org/articles/11945/elife-11945-v2.pdf.

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 Available at: https://www.sciencedirect.com/science/article/pii/S1053811917301714.

Jang Y, Wallsten TS, Huber DE (2012) A stochastic detection and retrieval model for the study of metacognition. Psychol Rev 119:186–200 Available at: http://dx.doi.org/10.1037/a0025960.

Kayser AS, Buchsbaum BR, Erickson DT, D'Esposito M (2009) The functional anatomy of a perceptual decision in the human brain. J Neurophysiol 103:1179–1194.

Kepecs A, Uchida N, Zariwala HA, Mainen ZF (2008) Neural correlates, computation and behavioural impact of decision confidence. Nature 455:227–231 Available at: http://www.nature.com/doifinder/10.1038/nature07200.

Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. Science 324:759–764 Available at: http://www.ncbi.nlm.nih.gov/pubmed/19423820.

Kleiner M, Brainard DH, Pelli DG, Ingling A, Murray R, Broussard C (2007) What's new in Psychtoolbox-3? Perception 36:1.

Koizumi A, Maniscalco B, Lau H (2015) Does perceptual confidence facilitate cognitive control? Attention, Perception, Psychophys 77:1295–1306 Available at: http://link.springer.com/10.3758/s13414-015-0843-3.

Lau H, Passingham RE (2006) Relative blindsight in normal observers and the neural correlate of visual consciousness. Proc Natl Acad Sci 103:18763–18768 Available at: http://www.pnas.org/content/103/49/18763.short.

Maniscalco B, Lau H (2016) The signal processing architecture underlying subjective reports of sensory awareness. Neurosci Conscious 2016 Available at: https://academic.oup.com/nc/article/doi/10.1093/nc/niw002/2757122.

Metcalfe J, Shimamura AP (1996) Metacognition: Knowing about knowing. Cambridge, MA, US: The MIT Press.

Middlebrooks PG, Sommer MA (2012) Neuronal correlates of metacognition in primate frontal

cortex. Neuron 75:517–530 Available at: https://www.sciencedirect.com/science/article/pii/S0896627312005247.

Morales J, Lau H, Fleming SM (2018) Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. J Neurosci.

Mueller ST, Weidemann CT (2008) Decision noise: An explanation for observed violations of signal detection theory. Psychon Bull Rev 15:465–494 Available at: http://www.springerlink.com/index/10.3758/PBR.15.3.465.

Pelli DG (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat Vis 10:437–442.

Peters MAK, Thesen T, Ko YD, Maniscalco B, Carlson C, Davidson M, Doyle W, Kuzniecky R, Devinsky O, Halgren E, Lau H (2017) Perceptual confidence neglects decision-incongruent evidence in the brain. Nat Hum Behav 1:0139.

Philiastides MG, Auksztulewicz R, Heekeren HR, Blankenburg F (2011) Causal role of dorsolateral prefrontal cortex in human perceptual decision making. Curr Biol 21:980–983 Available at: https://www.sciencedirect.com/science/article/pii/S0960982211004763.

Pouget A, Drugowitsch JD, Kepecs A (2016) Confidence and certainty: Distinct probabilistic quantities for different goals. Nat Neurosci 19:366–374 Available at: http://www.nature.com/articles/nn.4240.

Rahnev D (2017) Top-Down Control of Perceptual Decision Making by the Prefrontal Cortex. Curr Dir Psychol Sci 26:464–469 Available at: https://journals.sagepub.com/doi/full/10.1177/0963721417709807.

Rahnev D et al. (2020) The Confidence Database. Nat Hum Behav:1–8 Available at:

http://www.nature.com/articles/s41562-019-0813-1.

Rahnev D, Bahdo L, de Lange FP, Lau H (2012a) Prestimulus hemodynamic activity in dorsal attention network is negatively associated with decision confidence in visual perception. J Neurophysiol 108:1529–1536.

Rahnev D, Koizumi A, McCurdy LY, D'Esposito M, Lau H (2015) Confidence leak in perceptual decision making. Psychol Sci 26:1664–1680 Available at: http://journals.sagepub.com/doi/10.1177/0956797615595037.

Rahnev D, Kok P, Munneke M, Bahdo L, de Lange FP, Lau H (2013) Continuous theta burst transcranial magnetic stimulation reduces resting state connectivity between visual areas. J Neurophysiol 110:1811–1821 Available at: http://www.physiology.org/doi/10.1152/jn.00209.2013.

Rahnev D, Lau H, de Lange FP (2011a) Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain. J Neurosci 31:10741–10748 Available at: http://www.ncbi.nlm.nih.gov/pubmed/21775617.

Rahnev D, Maniscalco B, Graves T, Huang E, de Lange FP, Lau H (2011b) Attention induces conservative subjective biases in visual perception. Nat Neurosci 14:1513–1515 Available at: http://www.nature.com/articles/nn.2948.

Rahnev D, Maniscalco B, Luber B, Lau H, Lisanby SH (2012b) Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence. J Neurophysiol 107:1556–1563 Available at: http://www.physiology.org/doi/10.1152/jn.00985.2011.

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

Available at: http://www.ncbi.nlm.nih.gov/pubmed/27162349.

Resulaj A, Kiani R, Wolpert DM, Shadlen MN (2009) Changes of mind in decision-making. Nature 461:263–266 Available at: http://dx.doi.org/10.1038/nature08275.

Rounis E, Maniscalco B, Rothwell JC, Passingham RE, Lau H (2010) Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. Cogn Neurosci 1:165–175 Available at: http://www.tandfonline.com/doi/abs/10.1080/17588921003632529.

Ruff CC, Blankenburg F, Bjoertomt O, Bestmann S, Freeman E, Haynes J-D, Rees G, Josephs O, Deichmann R, Driver J (2006) Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. Curr Biol 16:1479–1488 Available at: https://www.sciencedirect.com/science/article/pii/S0960982206018185.

Ryals AJ, Rogers LM, Gross EZ, Polnaszek KL, Voss JL (2016) Associative recognition memory awareness improved by theta-burst stimulation of frontopolar cortex. Cereb Cortex 26:1200–1210 Available at: https://academic.oup.com/cercor/article-lookup/doi/10.1093/cercor/bhu311.

Samaha J, Barrett JJ, Sheldon AD, LaRocque JJ, Postle BR (2016) Dissociating perceptual confidence from discrimination accuracy reveals no influence of metacognitive awareness on working memory. Front Psychol 7:851 Available at: http://journal.frontiersin.org/Article/10.3389/fpsyg.2016.00851/abstract.

Samaha J, Iemi L, Postle BR (2017) Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy. Conscious Cogn 54:47–55.

Sanders JI, Hangya B, Kepecs A (2016) Signatures of a statistical computation in the human

sense of confidence. Neuron 90:499–506 Available at: https://www.sciencedirect.com/science/article/pii/S0896627316300162.

Saxe R, Kanwisher N (2003) People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind.” Neuroimage 19:1835–1842 Available at: https://www.sciencedirect.com/science/article/pii/S1053811903002301.

Shekhar M, Rahnev D (2018) Distinguishing the roles of dorsolateral and anterior PFC in visual metacognition. J Neurosci 38:5078–5087 Available at: http://www.ncbi.nlm.nih.gov/pubmed/29720553.

Shimamura AP (2000) The role of the prefrontal cortex in dynamic filtering. Psychobiology 28:207–218 Available at: https://link.springer.com/article/10.3758/BF03331979.

So N, Stuphorn V (2015) Supplementary eye field encodes confidence in decisions under risk. Cereb Cortex 26:764–782 Available at: https://academic.oup.com/cercor/article-lookup/doi/10.1093/cercor/bhv025.

Song A, Koizumi A, Lau H (2015) Awareness, a behavioral method to manipulate metacognitive awareness independent of stimulus. In: Behavioral Consciousness Research (Overgaard M, ed). Oxford University Press.

Spence ML, Dux PE, Arnold DH (2015) Computations underlying confidence in visual perception. J Exp Psychol Hum Percept Perform 42:671–682 Available at: http://dx.doi.org/10.1037/xhp0000179.

van den Berg R, Yoo AH, Ma WJ (2017) Fechner’s law in metacognition: A quantitative model of visual working memory confidence. Psychol Rev 124:197–214 Available at: http://psycnet.apa.org/buy/2017-07341-003.

Vlassova A, Donkin C, Pearson J (2014) Unconscious information changes decision accuracy but not confidence. Proc Natl Acad Sci U S A 111:16214–16218 Available at: http://www.ncbi.nlm.nih.gov/pubmed/25349435.

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

Welford AT (1980) Choice of reaction time: basic concepts. In: Reaction times (Welford AT, ed), pp 73–128. New York: Academic Press.

Wokke ME, Cleeremans A, Ridderinkhof KR (2017) Sure I'm sure: Prefrontal oscillations support metacognitive monitoring of decision making. J Neurosci 37:781–789 Available at: http://www.ncbi.nlm.nih.gov/pubmed/28123015.

Yazar Y, Bergström ZM, Simons JS (2014) Continuous theta burst stimulation of angular gyrus reduces subjective recollection. PLoS One 9:e110414 Available at: http://dx.plos.org/10.1371/journal.pone.0110414.

Ye Q, Zou F, Lau H, Hu Y, Kwok SC (2018) Causal evidence for mnemonic metacognition in human precuneus. J Neurosci 38:6379–6387.

Yeung N, Summerfield C (2012) Metacognition in human decision-making: Confidence and error monitoring. Philos Trans R Soc Lond B Biol Sci 367:1310–1321 Available at: http://www.ncbi.nlm.nih.gov/pubmed/22492749.

Zylberberg A, Barttfeld P, Sigman M (2012) The construction of confidence in a perceptual decision. Front Integr Neurosci 6:79 Available at: http://journal.frontiersin.org/article/10.3389/fnint.2012.00079/abstract.

Zylberberg A, Fetsch CR, Shadlen MN (2016) The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. Elife:e17688 Available at:

https://cdn.elifesciences.org/articles/17688/elife-17688-v1.pdf.