NeuroImage 106 (2015) 134–143

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Neural representations of confidence emerge from the process of decision formation during perceptual choices

Sabina Gherman, Marios G. Philiastides *

Institute of Neuroscience and Psychology, University of Glasgow, 58 Hillhead Street, Glasgow, G12 8QB, UK
Centre for Cognitive Neuroimaging, University of Glasgow, 58 Hillhead Street, Glasgow, G12 8QB, UK

ARTICLE INFO

Article history:
Accepted 17 November 2014
Available online 22 November 2014

Keywords:
Decision making
Evidence accumulation
Confidence
Single-trial
EEG
Linear discriminant analysis

ABSTRACT

Choice confidence represents the degree of belief that one's actions are likely to be correct or rewarding and plays a critical role in optimizing our decisions. Despite progress in understanding the neurobiology of human perceptual decision-making, little is known about the representation of confidence. Importantly, it remains unclear whether confidence forms an integral part of the decision process itself or represents a purely post-decisional signal. To address this issue we employed a paradigm whereby on some trials, prior to indicating their decision, participants could opt-out of the task for a small but certain reward. This manipulation captured participants' confidence on individual trials and allowed us to discriminate between electroencephalographic signals associated with certain-vs.-uncertain trials. Discrimination increased gradually and peaked well before participants indicated their choice. These signals exhibited a temporal profile consistent with a process of evidence accumulation, culminating at time of peak discrimination. Moreover, trial-by-trial fluctuations in the accumulation rate of nominally identical stimuli were predictive of participants' likelihood to opt-out of the task, suggesting that confidence emerges from the decision process itself and is computed continuously as the process unfolds. Correspondingly, source reconstruction placed these signals in regions previously implicated in decision making, within the prefrontal and parietal cortices. Crucially, control analyses ensured that these results could not be explained by stimulus difficulty, lapses in attention or decision accuracy.

Introduction

Imagine running in the park on a rainy day, trying to discern whether the person across the lawn is an old friend. The decision to keep concentrating on your stride or change directions to go greet them depends on your level of confidence that it is really them. Choice confidence is crucial not only for such mundane tasks, but also for more biologically and socially complex situations. It provides a probabilistic assessment of expected outcome and can play a key role in how we adjust in ever-changing environments, learn from trial and error, make better predictions, and plan future actions.

In recent years, systems and cognitive neuroscience started to examine the neural correlates underlying perceptual decision making. As a result, many monkey neurophysiology (Gold and Shadlen, 2007; Kim and Shadlen, 1999; Mazurek et al., 2003; Newsome et al., 1989; Shadlen et al., 1996; Shadlen and Newsome, 2001), human neuroimaging (Cheadle et al., 2014; Heekeren et al., 2004, 2006, 2008; Ho et al., 2009; Ploran et al., 2007; Tosoni et al., 2008), and human electrophysiology (de Lange et al., 2010; Donner et al., 2007, 2009; O'Connell et al., 2012; Philiastides et al., 2006; Philiastides and Sajda, 2006; Ratcliff et al., 2009; Wyart et al., 2012) experiments provided converging support that perceptual decisions are characterized by a noisy temporal accumulation of sensory evidence which culminates when an observer commits to a choice. Despite this progress, however, it remains unclear how confidence is represented in the human brain and what its relationship is with the decision process itself.

Current theoretical and experimental accounts have regarded confidence as a meta-cognitive event (i.e. an epiphenomenon of the decision process) that relies on new information arriving beyond the decision point (Fleming et al., 2012; Pleskac and Busemeyer, 2010; Yeung and Summerfield, 2012). Conversely, little has been done in the way of exploring whether confidence might emerge earlier in the decision process and before one commits to a choice. Evidence for the latter has recently emerged from a limited number of animal studies (Kepecs et al., 2008; Kiani and Shadlen, 2009; Shadlen and Kiani, 2013), proposing that choice confidence in perceptual judgments might be an inherent property of the decision process itself and that the same neural generators involved in evidence accumulation also encode choice confidence. To date, it remains unclear whether confidence forms an integral part of the decision process itself and whether it emerges from the same neural generators involved in accumulating evidence for the decision. Similarly, it is unknown whether confidence is reflected in the rate of evidence accumulation itself.

* Corresponding author at: 58 Hillhead Street, Glasgow, G12 8QB, UK.
E-mail address: Marios.Philiastides@glasgow.ac.uk (M.G. Philiastides).

http://dx.doi.org/10.1016/j.neuroimage.2014.11.036
1053-8119/

To address these open questions, we collected electroencephalography (EEG) data during a binary, delayed-response, task in which correct responses were rewarded with monetary incentives. Importantly, on a random half of trials and after forming a decision, participants were given the option to opt out of the task for a smaller but sure reward (a form of post-decision wager (Kiani and Shadlen, 2009)). We expected participants to waive the sure reward when they were certain of their choice, and select it otherwise. This in turn allowed us to use a multivariate single-trial classifier to discriminate between certain-vs.-uncertain trials to identify the temporal characteristics of the neural correlates of choice confidence. Importantly, additional control analyses were carried out to ensure that confidence-related effects could not be explained by stimulus difficulty or trial-by-trial changes in attention.

Materials and methods

Participants

Nineteen subjects (7 males) aged between 18 and 36 years (mean = 23.4 years) participated in the experiment. All had normal or corrected-to-normal vision and reported no history of neurological problems. Written informed consent was obtained in accordance with the School of Psychology Ethics Committee at the University of Nottingham.

Stimuli and task

Stimuli consisted of 20 face (face database, Max Planck Institute for Biological Cybernetics, Tuebingen, Germany) (Troje and Bulthoff, 1996) and 20 car greyscale images obtained from the Web (size 500 × 500 pixels, 8-bits/pixel). Spatial frequency, contrast, and luminance were equalized across all images, and the magnitude spectrum of each image was adjusted to the average magnitude spectrum of all images. We manipulated the phase spectrum of the images to obtain noisy stimuli of varying levels of sensory evidence (i.e. we manipulated the percentage phase coherence of our images) (Dakin et al., 2002). Stimuli were presented centrally on a plain grey background on a computer screen using PsychoPy software (Peirce, 2007). The display was situated 1 m away from the subject, with each stimulus subtending approximately 8 × 8 degrees of visual angle.

We used a training session prior to the main task to identify subject-specific phase coherence values for the stimuli used in the main task. Specifically, during training subjects were required to perform a simple speeded face vs. car categorizations over a total of 600 trials, using images with 7 different phase coherence values (27.5–42.5%, in increments of 2.5%). Each image was presented for 0.1 s, and subjects were allowed a maximum of 1.25 s to make a response. The response was followed by an inter-trial interval, randomized between .75 and 1.5 s. There were an equal number of face and car stimuli, and these were presented in random order. Based on performance during this session, we selected three subject-specific phase coherence levels for the main task (henceforth referred to as low, medium, and high), which spanned psychophysical threshold (in the range 60–80% accuracy).

For the main experiment, subjects performed face vs. car categorizations during a delayed-response, post-decision wagering paradigm designed to discriminate between certain and uncertain trials (Fig. 1A). Importantly, on a random half of the trials, subjects were offered the option to opt-out of the task for a smaller (relative to a correct response) but sure reward (SR). This manipulation encouraged subjects to select the SR option on low confidence trials (Kiani and Shadlen, 2009). Responses were rewarded with points (correct = 10 points, incorrect = 0 points, SR choice = 8 points). The total number of points collected was translated into a monetary payment at the end of the experiment. Each trial began with a face or car stimulus presented for 0.1 s at one of the three possible sensory evidence levels. Stimulus presentation was followed by a forced delay (i.e., the decision time) randomized between 0.9 and 1.4 s. This delay was introduced prior to revealing whether participants could opt-out of the task, to ensure they formed a decision on every trial. Next, a visual response cue (1 s) informed participants whether or not the SR option would be available—this was indicated by a green or red fixation cross, respectively. In addition, the letters "F" (for face) and "C" (for car) where positioned randomly to the left and right of the central fixation cross to indicate the mapping between stimulus and motor effectors (right index and ring fingers, respectively). The latter manipulation aimed at separating the decision process from motor planning and execution. Subjects indicated their choice by pressing one of three buttons on a response box (LEFT/RIGHT for a stimulus choice, MIDDLE for the SR). They were instructed to respond after the response cue was removed from the screen. A response was followed by an inter-trial interval randomized in the range 1–1.5 s. Overall subjects performed 480 trials, divided into two blocks of 240 trials each.

EEG data acquisition

We recorded EEG data during performance of the main task, in an electrostatically shielded room, using a DBPA-1 digital amplifier (Sensorium Inc., VT, USA), at a sampling rate of 1000Hz. We used 117 Ag/AgCl scalp electrodes and three periocular electrodes placed below the left eye and at the left and right outer canthi. Additionally, a chin electrode was used as ground. All channels were referenced to the left mastoid. Input impedance was adjusted to <50 kOhm. To obtain accurate event onset times we placed a photodiode on the monitor to detect the onset of the stimuli. An external response device was used to collect response times. Both signals were collected on two external channels on the EEG amplifiers to ensure synchronization with the EEG data.

EEG data pre-processing

We applied a 0.5–100 Hz band-pass filter to the data to remove slow DC drifts and high frequency noise. These filters were applied non-causally (using MATLAB "filtfilt") to avoid phase related distortions. Additionally, we re-referenced our data to the average of all electrodes. To remove eye movement artifacts, participants performed an eye movement calibration task prior to the main experiment, during which they were instructed to blink repeatedly several times while a central fixation cross was displayed in the center of the computer screen, and to make lateral and vertical saccades according to the position of the fixation cross. We recorded the timing of these visual cues and used principal component analysis to identify linear components associated with blinks and saccades, which were then removed from the EEG data (Parra et al., 2005). Finally, we baseline corrected our EEG data, with the baseline interval defined as the 100 ms prior to stimulus onset.

Single trial EEG analysis

To identify confidence-related activity in the neural data, we used a single-trial multivariate discriminant analysis (Parra et al., 2002, 2005) to estimate linear spatial weightings of the EEG sensors, which discriminated between certain (SR waived) and uncertain (SR selected) trials. We applied our technique to discriminate between the two groups of trials at various time points, in the time range between 100 ms prior to, and 1000 ms following the presentation of the visual stimulus (i.e. during the decision phase of the trial). For each participant we estimated, within short pre-defined time windows of interest, a projection in the multidimensional EEG space (i.e. a spatial filter) that maximally discriminated between the two conditions on stimulus-locked data (Eq. (1)). Unlike conventional, univariate, trial-average event-related potential analysis, our multivariate approach is designed to spatially integrate information across the multidimensional sensor space, rather

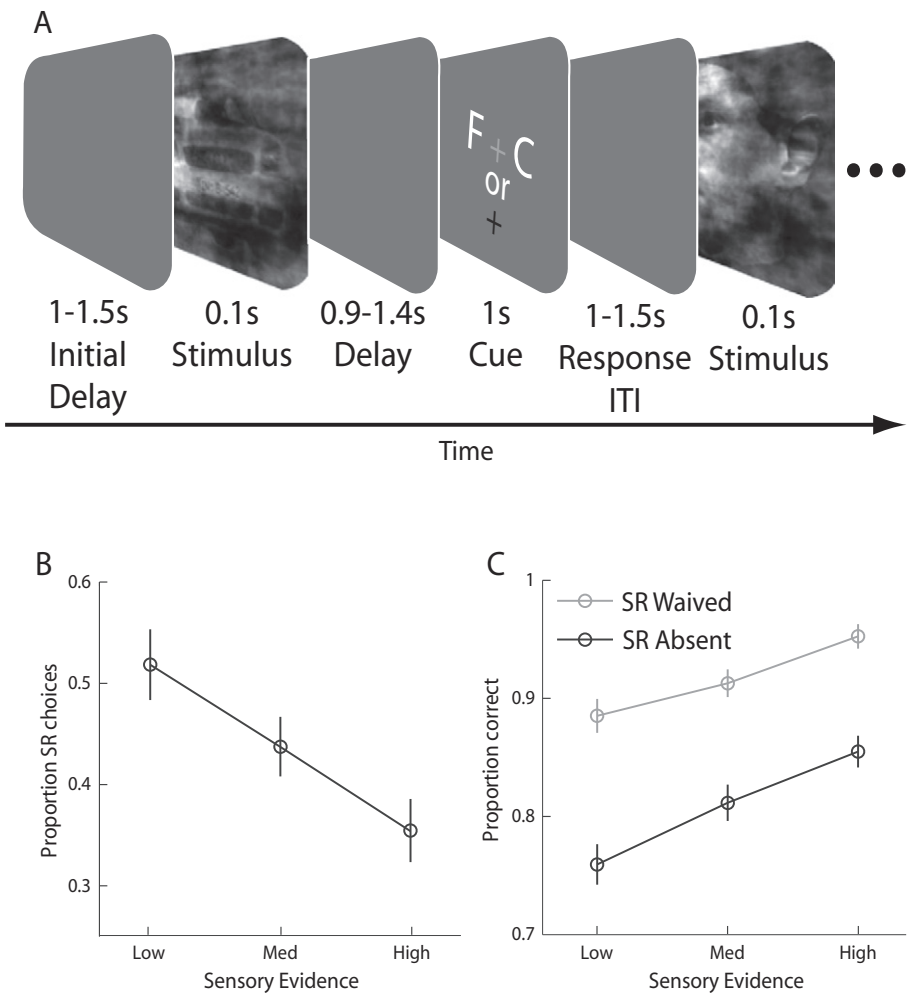


Fig. 1. Experimental design and behavioral performance. A. Schematic representation of the behavioral paradigm. Participants had to categorize a briefly presented (0.1 s) image, at one of three possible levels of sensory evidence, as being a face or car. Stimulus presentation was followed by a random delay (0.9–1.4 s) during which participants had to form a decision. Next, a visual response cue (1 s) informed participants whether a small (relative to a correct choice) but sure reward (SR) was available or not, with either a green (shown in grey) or red (shown in black) cross, respectively. The letters "F" (for face) and "C" (for car) were positioned randomly to the left and right of the fixation cross, indicating the mapping between stimulus and motor effectors (right index and ring fingers respectively). Participants indicated their choice as soon as the response cue was removed from the screen. B. Mean proportion of SR choices (on trials where the SR was offered), across subjects, as a function of sensory evidence. C. Mean proportion of correct responses, across subjects, for SR waived (grey) vs. SR absent (black) trials, as a function of the three levels of sensory evidence. Error bars in B and C represent standard errors across subjects.

than across trials, to increase signal-to-noise ratio while preserving single-trial information.

Specifically, our method aimed to identify a one-dimensional 'discriminating component', $\boldsymbol{y}(t)$, by integrating information across all D electrodes, which maximally discriminated between the two trial groups. We use the term 'component' instead of 'source' to make it clear that this is a projection of all the activity correlated with the underlying source. We did this by applying a weighting vector $\boldsymbol{w}$ (i.e. a spatial filter) to our multidimensional EEG data ($\boldsymbol{x}(t)$), as summarized in the equation below:

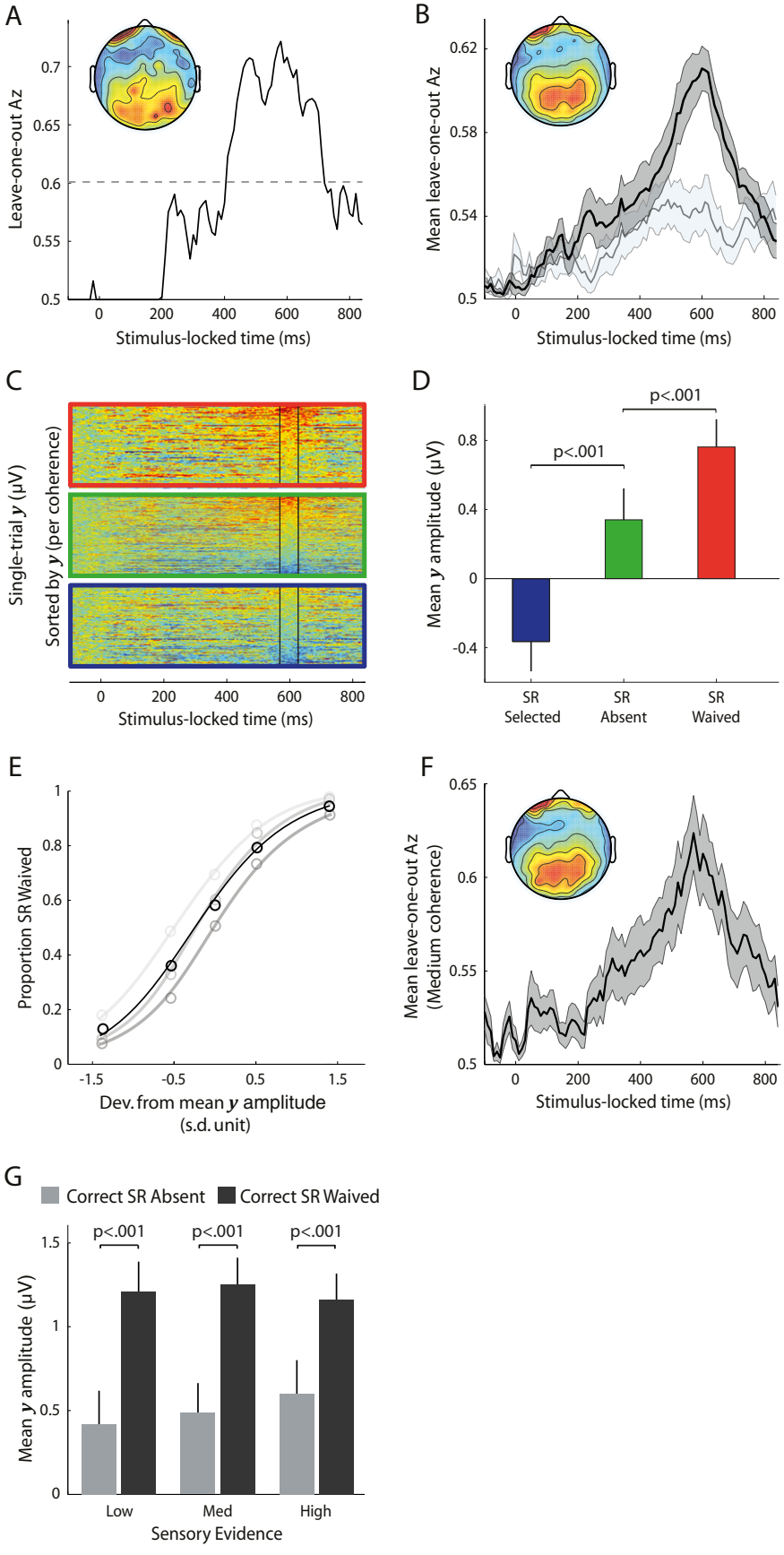
$$\boldsymbol{y}(t) = \boldsymbol{w}^T\boldsymbol{x}(t) = \sum_{i=1}^{D} w_i x_i(t) \quad (1)$$

We used logistic regression and a reweighted least squares algorithm to learn the optimal discriminating spatial weighting vector $\boldsymbol{w}$ (Jordan and Jacobs, 1994). We used this approach to identify a $\boldsymbol{w}$ for several short pre-defined training windows centered at various latencies across our epoch of interest. Specifically, we used a 60 ms training window and stimulus-locked onset times varying from 100 ms before until 1000 ms after the stimulus, in increments of 10 ms. The spatial filters ($\boldsymbol{w}$) obtained this way applied to an individual trial produce a measurement of the component amplitude for that trial. In separating the two groups of trials the discriminator was designed to map the component amplitudes for one condition to positive values and those of the other condition to negative values; note that this mapping was arbitrary. Here, we mapped the high confidence (SR waived) trials to positive values and the low confidence (SR selected) trials to negative values.

We quantified the performance of the discriminator for each time window using the area under a receiver operating characteristic (ROC) curve, referred to as an *Az*-value, using a leave-one-out procedure (Duda et al., 2001). To assess the significance of the discriminator we used a bootstrapping technique where we used a leave-one-out trial approach after randomizing the trial labels. We repeated this randomization procedure 1000 times to produce a probability distribution for *Az*, and estimated the *Az* leading to a significance level of $p < 0.01$.

To visualize the profile of the discriminating component, $\boldsymbol{y}$, across individual trials, we also constructed discriminant component maps (see Fig. 2C for an example). To do so we applied the spatial weighting vector $\boldsymbol{w}$ of the window that resulted in the highest discrimination performance between SR waived vs. SR selected trials, across an extended time range (100 ms before until 1000 ms after the stimulus). Each row of one such discriminant component map represents a single trial

across time. We also sorted trials (i.e., the rows of these maps) based on the amplitude of the discriminating component in the time window of maximum discrimination. We also used this approach to compute the temporal profile of the discriminating component, ***y***, along the sensory evidence dimension to look for evidence of a gradual build-up of activity leading up to the point of maximum discrimination and to extract



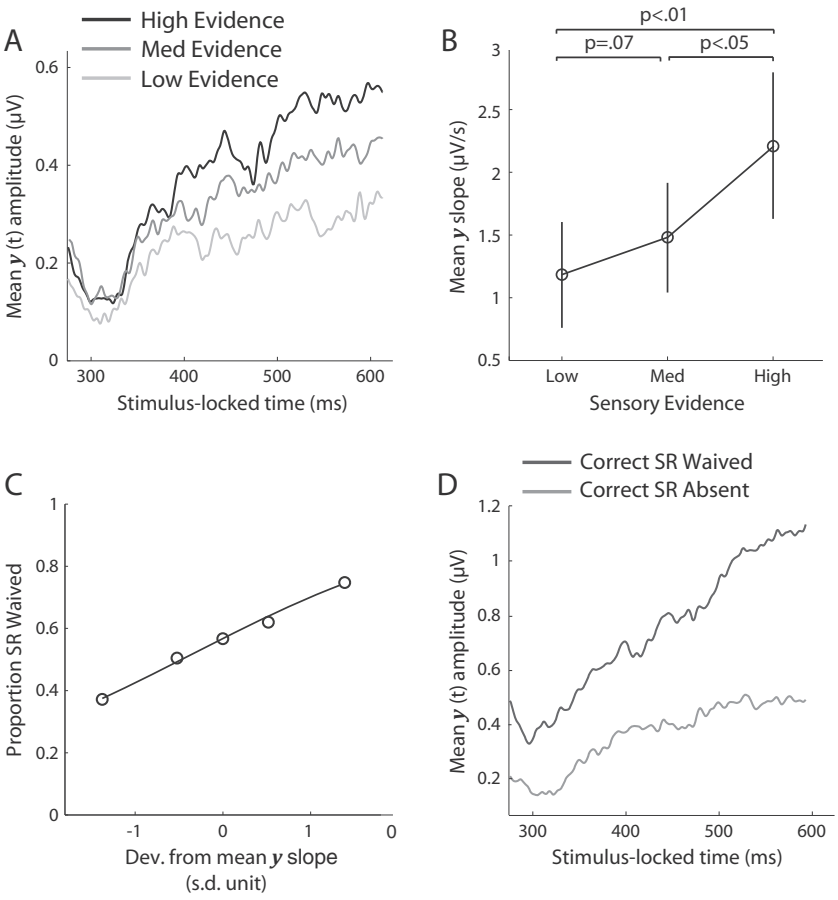


Fig. 3. Choice confidence and evidence accumulation. A. Subjecting our data through the same spatial distribution of component activity estimated during confidence discrimination (i.e., Figs. 2A, B) revealed a gradual build-up of activity (i.e. accumulating activity) earlier in the trial that was modulated by the amount of sensory evidence (i.e. % stimulus phase coherence). Trials were locked to the onset of the stimulus and averaged across subjects. B. Mean slope of the accumulating activity across subjects was positively correlated with the amount of sensory evidence. Slopes were estimated by computing linear fits through the data based on subject-specific onset and peak accumulation times. Error bars represent standard errors across subjects. C. Trial-by-trial deviations from the mean accumulation slope were positively correlated with the probability of waiving the SR. To visualize this association the data points were computed by grouping trials into five bins based on the deviations in the slope of the accumulating activity. Importantly, the curve is a fit of Eq. (4) to individual trials. D. Confidence-related accumulating activity was steeper for correct SR waived (confident) trials (dark grey) than for correct SR absent (on average, less confident) trials (light grey).

single-trial slopes of this accumulating activity. Slopes were computed using linear regression between the onset- and peak times of the accumulating activity extracted from individual participants. Specifically, we extracted subject-specific accumulation onset-times by selecting (through visual inspection) the time point at which the discriminating component activity began to rise in a systematic fashion after an initial dip in the data following any early (non-discriminative) evoked responses present in the data (as seen in Fig. 3A). Peak accumulation times were selected as the time points of maximum discrimination across individual participants. To justify our choice for a linear model in estimating accumulation slopes, we fit three additional models (exponential, logarithmic and power-law) to the individual subject accumulation patterns, using the same onset and peak accumulation times. We compared the goodness of fit to the data (mean square error) and found that the linear model provided the best fit to the accumulating activity, across all levels of sensory evidence.

Fig. 2. Neural representation of choice confidence. A. Classifier performance (Az) during high-vs.-low confidence discrimination (i.e. SR waived vs. SR chosen) of stimulus-locked single-trial data, for a representative subject. The dotted line represents the subject-specific Az value leading to a significance level of $p = 0.01$, estimated using a bootstrap test. The scalp topography is associated with the discriminating component estimated at time of maximum discrimination. B. Mean classifier performance and scalp topography across subjects during confidence discrimination (i.e. SR waived vs. SR chosen) (dark grey). For comparison, mean classifier performance during accuracy discrimination (i.e. correct vs. incorrect) for SR absent trials is also shown (light grey). Shaded areas represent standard errors across subjects. C. Single-trial discriminant component maps, for a representative subject, obtained by applying the subject-specific spatial projections estimated at the time of maximum discrimination (black window) to an extended time range relative to the onset of the stimulus and across all trials (including SR absent trials that were independent of those used to train the classifier). Each row in these maps represents discriminant component amplitudes, $y(t)$, for a single trial across time. Within each trial group (top to bottom panel: SR waived, SR absent, SR selected), trials are sorted by mean component amplitude (y) at time of maximum discrimination. Red represents positive and blue negative component amplitudes, respectively. D. Mean component amplitude for the SR absent group was situated between those of the high and low confidence groups (SR waived and SR selected). This is consistent with a mixture of "certain" and "uncertain" trials in the SR absent group as can be seen in C for one participant (i.e. a mixture of red and blue component amplitudes). Error bars are standard errors across subjects. E. Trial-by-trial deviations from the mean component amplitude at time of maximum confidence discrimination were positively correlated with the probability of waiving the SR. To visualize this association the data points were computed by grouping trials into five bins based on the deviations in component amplitude. Importantly, the curve is a fit of Eq. (3) to individual trials. Grey curves are fits of Eq. (3) to each of the three levels of sensory evidence separately (light to dark grey represents high to low sensory evidence. F. Mean classifier performance and scalp topography across subjects within an individual level of sensory evidence (medium phase coherence; results looked very similar for the other two levels). Note that the patterns are qualitatively very similar to those shown in B for which classification was performed over all trials. Shaded area represents standard errors across subjects. G. Mean component amplitude for correct SR waived (confident) trials (dark grey) and correct SR absent (on average, less confident) trials (light grey), split by level of sensory evidence. Error bars are standard errors across subjects.

Given the linearity of our model we also computed scalp projections of the discriminating components resulting from Eq. (1) by estimating a forward model for each component:

$$a = \frac{\boldsymbol{X}\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}} \tag{2}$$

where the EEG data ($\boldsymbol{X}$) and discriminating components ($\boldsymbol{y}$) are now in a matrix and vector notation, respectively, for convenience (i.e., both $\boldsymbol{X}$ and $\boldsymbol{y}$ now contain a time dimension). Eq. (2) describes the electrical coupling of the discriminating component $\boldsymbol{y}$ that explains most of the activity in $\boldsymbol{X}$. Strong coupling indicates low attenuation of the component $\boldsymbol{y}$ and can be visualized as the intensity of vector $\mathbf{a}$. We used these scalp projections as a means of localizing the underlying neuronal sources (see next section).

Distributed source reconstruction

To spatially localize the resultant discriminating component activity related to choice confidence we used a distributed source reconstruction approach based on empirical Bayes (Friston et al., 2008) as implemented in SPM8 (http://www.fil.ion.ucl.ac.uk/spm/). The method allows for an automatic selection of multiple cortical sources with compact spatial support that are specified in terms of empirical priors, while the inversion scheme allows for a sparse solution for distributed sources (refer to (Friston et al., 2008) for details). We used a three-sphere head model, which comprised of three concentric meshes corresponding to the scalp, the skull and the cortex. The electrode locations were co-registered to the meshes using fiducials in both spaces and the head shape of the average MNI brain.

To compute the electrode activity to be projected onto these locations, we applied Eq. (2) to extract, at each time point, the scalp activity that was correlated with the confidence discriminating component $\boldsymbol{y}$ estimated during peak discriminator performance (i.e. we computed a forward model indexed by time,$\mathbf{a}(t)$). We estimated $\mathbf{a}(t)$ in 1 ms data increments in the time range between 300 and 880 ms after stimulus onset (i.e. around the peak discrimination time).

Analysis of neural data

We used different logistic regressions to examine how neural activity correlated with participants' behavioral performance. To factor out the effect of task difficulty in our analyses, we first z-scored, at each level of sensory evidence separately, both the single-trial confidence component amplitudes (i.e., $\boldsymbol{y}$ at the end of the accumulation process) and the single-trial slopes of the accumulating activity itself (Acc. slopes). Subsequently, we proceeded to perform different regression analyses on these trial-to-trial residual fluctuations (i.e. deviations from mean $\boldsymbol{y}$ and Acc. slopes). Regression analyses were performed separately for each subject.

To assess how the fluctuations in discriminant component amplitude $\boldsymbol{y}$ (estimated from discriminating certain vs. uncertain trials) influenced participants' likelihood of waiving the sure reward (SR), on trials where this option was available, we performed the following regression analysis:

$$P_{SR\ Waived} = \left[1 + e^{-(\beta_0 + \beta_1 y)}\right]^{-1} \tag{3}$$

We expected a positive correlation between the two quantities (as larger fluctuations in $\boldsymbol{y}$ amplitudes are expected to reflect more confident trials), and thus we tested whether the regression coefficients resulting across subjects ($\boldsymbol{\beta}_1$s in Eq. (3)) came from a distribution with mean larger than zero (using a one-tailed t-test). We also repeated this analysis for each level of sensory evidence separately and tested whether $\boldsymbol{y}$ remained a significant predictor of participants' likelihood to waive the SR in each of the three levels. Moreover, we tested for differences in explanatory power across the three levels by comparing the resulting regression coefficients (using one-tailed paired t-test).

To assess how the slope of the accumulating activity influenced behavioral performance, we used the same rationale as with the previous analysis. Specifically, we used the accumulation slopes as a predictor for the probability of waiving the SR, on trials where this option was available:

$$P_{SR\ Waived} = \left[1 + e^{-(\beta_0 + \beta_1\ \text{Acc. Slopes})}\right]^{-1} \tag{4}$$

We hypothesized that, if confidence is an inherent property of the accumulation process itself, then accumulation slopes would be positively correlated with the probability of waiving the SR (i.e., $\boldsymbol{\beta}_1 > 0$), and we performed a one-tailed t-test to formally test for this hypothesis.

Next, we investigated whether accumulation slopes provided additional explanatory power for the probability of waiving the SR than what was already conferred by the discriminant component amplitude $\boldsymbol{y}$ (i.e. whether a significant positive correlation with accumulation slopes would still be present if the discriminant component amplitude $\boldsymbol{y}$ was included as an additional predictor in the regression):

$$P_{SR\ Waived} = \left[1 + e^{-(\beta_0 + \beta_1\ y + \beta_2\ \text{Acc. Slopes})}\right]^{-1} \tag{5}$$

As before, we performed a one-tailed t-test to assess whether regression coefficients for accumulation slopes ($\boldsymbol{\beta}_2$s in Eq. (5)) came from a distribution with mean larger than zero.

To rule out the possibility that confidence effects are driven by changes in attention across trials we included two additional predictors in the previous regression model, corresponding to two well-known neural signatures of attention; 1) pre-stimulus EEG power in the α band ($\boldsymbol{\alpha_{prestim}}$), which was linked to top-down control of attention (Wyart and Tallon-Baudry, 2009) and was shown to correlate with visual discrimination performance (Thut et al., 2006; van Dijk et al., 2008), resulting from the analysis described in the next section and 2) an evoked component appearing 220 ms post-stimulus ($\boldsymbol{y_{220}}$), which was shown (in the same task used here) to index allocation of attentional resources required for the decision (Philiastides et al., 2006), and was localized in areas of the frontoparietal attention network (Philiastides and Sajda, 2007).

$$P_{SR\ Waived} = \left[1 + e^{-(\beta_0 + \beta_1 y + \beta_2\ \text{Acc. Slopes} + \beta_3 \alpha_{prestim} + \beta_4\ y_{220})}\right]^{-1} \tag{6}$$

We expected the fluctuations associated with confidence in both discriminant component amplitude $\boldsymbol{y}$ and accumulation slopes to remain significant positive predictors of the likelihood of waiving the SR, and thus we tested whether the resulting regression coefficients across subjects ($\boldsymbol{\beta}_1$s and $\boldsymbol{\beta}_2$s in Eq. (6)) came from a distribution with mean larger than zero (using a one-tailed t-test).

Single-trial power analysis

Pre-stimulus alpha power was obtained using a wavelet transform as in (Mazaheri and Jensen, 2006; Tallon-Baudry et al., 1996). In short, single trials were convolved by a complex Morlet wavelet $w(t,f_0) = Aexp(-t^2/2\sigma_t)exp(2i\pi f_0 t)$, where $\sigma_t = m/2\pi f_0$, and i is the imaginary unit. $A = (\sigma_t\sqrt{\pi})^{-1/2}$ is a normalization term, whereas the constant m defines the time-frequency resolution tradeoff and was set to 7. The wavelet transformation produces a complex time series for the frequencies f_0 of interest (here 8-12 Hz). Single-trial power was calculated by averaging the squared absolute values of the convolutions in the 500 ms preceding the onset of the stimulus at the subject-specific peak alpha frequency and occipitoparietal sensor with the highest overall alpha power.

Results

Our participants' behavioral performance indicated that our paradigm was successful in capturing choice confidence. Specifically,

our participants selected the SR more frequently in more difficult trials ($F(2, 36) = 55.87$, $p < .001$, *post hoc* paired t-tests, all $p < .001$, Fig. 1B), consistent with previous reports showing that confidence scales with the amount of sensory evidence (Vickers and Packer, 1982). Importantly, there was no difference in the frequency of choosing the SR across face and car trials ($t(18) = 1.7$, $p = 0.11$) ensuring this effect was not driven by one of the two stimulus categories.

More interestingly, accuracy on trials in which participants were offered the SR and rejected it was significantly higher compared to the trials in which the SR was not available ($F(1, 18) = 100.26$, $p < .001$, Fig. 1C). This effect was present for all levels of sensory evidence suggesting that participants waived the SR based on a sense of confidence on each trial rather than on the level of stimulus difficulty. Overall there was no significant difference in accuracy between face and car trials indicating that there was no category-specific choice bias ($t(18) = 0.76$, $p = 0.46$). As expected (Blank et al., 2013; Philiastides et al., 2006; Philiastides and Sajda, 2006), there was also a main effect of stimulus difficulty ($F(2, 36) = 28.99$, $p < .001$, *post hoc* paired t-tests, all $p < .001$, Fig. 1C), with accuracy increasing with the amount of sensory evidence. Finally, we note, that due to the delayed-response paradigm employed here, there were no significant differences in response time between certain (SR waived) and uncertain (SR selected) trials (420 ms and 406 ms respectively, $t(18) = 0.99$, $p = .33$).

To identify confidence-related activity in the neural data, we used a single-trial multivariate approach to discriminate between certain (SR waived) and uncertain (SR selected) trials. We observed that the discriminator's performance increased gradually after 300 ms (i.e. after early encoding of the stimulus) and peaked around 600 ms post-stimulus, on average. This pattern of discriminator performance was visible in individual data (Fig. 2A) as well as in the group average (Fig. 2B), consistent with the idea that confidence develops gradually as the decision process unfolds and culminates before one commits to a choice (Ding and Gold, 2013; Kiani and Shadlen, 2009). To visualize the temporal profile of this discriminating component activity across trials, we also constructed single-trial component maps by applying our subject-specific spatial projections estimated in the time window yielding maximum confidence discrimination (using Eq. (1)) to an extended time window. These maps clearly highlight the overall difference in component amplitude $\boldsymbol{y}$ between SR waived and SR selected trials and the temporally broad response profile of the discriminating component, both of which contributed to the discriminator's performance. The maps also highlight the trial-by-trial variability in the amplitude and temporal spread of this component, providing qualitative support that decision confidence might represent a graded quantity (Fig. 2C).

To provide further support linking this discriminating component to choice confidence, we considered trials in which the SR was not available (i.e. SR absent), and participants were forced to make a face/car response. Importantly, these trials can be considered as "unseen" data (they are independent of those used to train the classifier), and can be subjected through the same neural generators (i.e. spatial projections) estimated during discrimination of SR waived vs. SR selected trials. We expected that these trials would contain a mixture of confidence levels and therefore the resulting mean component amplitude at the time of peak discrimination would be situated between those of the certain and uncertain trial groups (i.e. SR waived > SR absent > SR selected). Indeed, this was the case and the mean SR absent activity was significantly different from both the SR selected ($t(18) = 7.53$, $p < .001$) and SR waived ($t(18) = -7.71$, $p < .001$) (Fig. 2D). The mixture of both high and low confidence trials within the SR absent group can be further appreciated by inspecting the resulting single-trial component amplitudes (Fig. 2C; middle panel).

A potential concern is that subjects' choice to waive or select the SR (and consequently our discriminator's performance) is driven primarily by the physical properties of the stimulus (i.e. stimulus difficulty). This is unlikely, as changes in early stimulus encoding would have produced significant discrimination performance earlier in the trial (i.e. around 170–200 ms post-stimulus, driven by EEG components known to be affected by stimulus evidence—N170/P200 (Jeffreys, 1996; Liu et al., 2000; Philiastides et al., 2006)), which was absent in our data (see discriminator performance at the relevant time windows in Fig. 2A, B). Nonetheless, we performed additional analyses to ensure that stimulus difficulty could not explain the observed effects.

We first removed the overall influence of stimulus difficulty by computing the trial-to-trial deviations around the mean discriminating component activity, separately for each level of sensory evidence, and used these residual fluctuations as predictors of participants' choices to waive the SR in a single-trial logistic regression analysis (Eq. (3)). We found a significant positive correlation ($t(18) = 15.19$, $p < .001$) between component amplitudes and the probability of waiving the SR (i.e. bigger amplitudes, higher probability of SR waived; Fig. 2E). Crucially, we also repeated this regression analysis separately for each level of sensory evidence and found that our component amplitudes remained a significant predictor of subjects' opt-out behavior within each level of stimulus difficulty (all $p < .001$), without significant differences in explanatory power across the three levels (all $p \geq .2$; Fig. 2E). Similarly, we repeated the discrimination between certain-vs.-uncertain trials using observations from individual levels of sensory evidence and demonstrated that our discriminator performance remained virtually unchanged compared to our main analysis (compare Fig. 2B with 2 F for a single level of difficulty).

To identify the spatial extent of our confidence component, we first computed a forward model of the discriminating activity (Eq. (2)), which can be visualized in the form of a scalp map (Fig. 2A, B). Importantly, we used these forward models as a means of localizing the underlying neural generators using a Bayesian distributed source reconstruction technique (Friston et al., 2008). The source analysis revealed sources in areas in the dorsolateral prefrontal cortex with a pronounced left bias and in regions of the posterior parietal cortex, bilaterally (Fig. 4; explained variance > 97%), areas which have previously been implicated

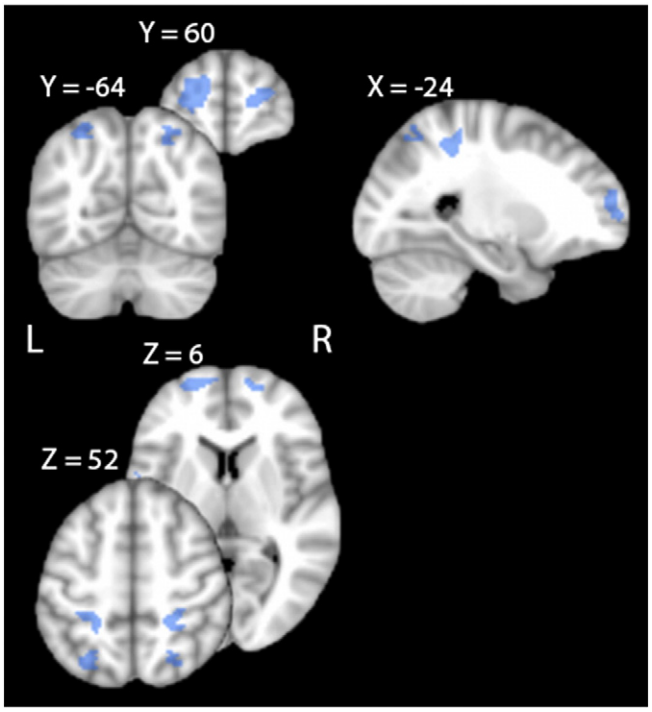


Fig. 4. Spatial representation of choice confidence. A distributed source reconstruction technique (Friston et al., 2008) revealed neural generators associated with choice confidence in dorsolateral prefrontal cortex (with a left bias) and in distinct clusters in parietal cortex, bilaterally (along the intraparietal sulcus). Slice coordinates are given in mm's in MNI space.

in perceptual decision making and evidence accumulation, both in the human (Heekeren et al., 2006; Ploran et al., 2007; Tosoni et al., 2008) and primate (Kiani and Shadlen, 2009; Kim and Shadlen, 1999; Shadlen and Newsome, 2001) brains. These results, coupled with the gradual build-up of confidence-related discriminating activity (Figs. 2A, B), suggest that choice confidence might be encoded in the same brain areas supporting evidence accumulation and decision formation. Moreover, they raise the intriguing possibility that confidence is computed continuously as the decision process unfolds, thus being reflected in the slope of the process of evidence accumulation itself (Ding and Gold, 2013).

To formally test these predictions, we subjected the data through the same neural generators (i.e. spatial projections) estimated for the confidence discrimination but stratified our trials along the sensory evidence dimension instead. In doing so, we observed ramp-like activity starting, on average, at 300 ms post-stimulus, which built up gradually to the time of peak confidence discrimination (Fig. 3A), and whose slope was parametrically modulated by the amount of sensory evidence ($F(2,36) = 10.6$, $p < 0.001$, Fig. 3B), consistent with a process of evidence accumulation (Kelly and O'Connell, 2013; Philiastides et al., in press). Importantly, this finding suggests that choice confidence and evidence accumulation share common neural generators. To investigate whether confidence emerges from the decision process itself, we tested whether the trial-by-trial build-up rates of the accumulating activity were predictive of participants' opt-out behavior. Specifically, we used single-trial slope estimates of the accumulating activity to predict participants' decisions to waive the SR in a new logistic regression model (Eq. (4)). As in the previous analysis, overall stimulus difficulty effects were removed from individual trials. We found a significant positive correlation ($t(18) = 11.94$, $p < .001$) between the slope of accumulation and the probability of waiving the SR (i.e. steeper slopes, higher probability of SR waived, Fig. 3C).

A potential confound of the previous analysis is that the slope of the accumulating activity simply echoes the confidence effects we identified earlier on the amplitude of our discriminating component, as the latter were extracted, on average, near the end of the accumulating activity. Crucially, we found that the two quantities were only partially correlated ($r = .39$, $p < .001$), due to the high degree of inter-trial variability in internal components of decision processing as has been described previously by accumulation-to-bound models (Bogacz et al., 2006; Mulder et al., 2014; Ratcliff et al., 2009; van Maanen et al., 2011). As such we found that each exerted a separate influence on our participants' opt-out behavior (Eq. (5), $t(18) = 2.96$, $p = .008$), which suggests that traces of confidence begin to develop as early as the decision process itself and continue to be reflected in the process of evidence accumulation, becoming progressively more robust as the decision unfolds.

Importantly, to rule out that our confidence effects are driven by changes in attention across individual trials we exploited two well-known neural signatures of attention (pre-stimulus alpha (Wyart and Tallon-Baudry, 2009) and a post-stimulus evoked response indexing allocation of attentional resources (Philiastides et al., 2006)), which we used as additional predictors of our participants' opt-out behavior in a different logistic regression model (Eq. (6)). Crucially, we found that our original confidence component amplitudes and accumulation slopes remained significant predictors of the likelihood of waiving the SR (component amplitudes: $t(18) = 14.51$, $p < .001$, one-tailed t-test; slopes: $t(18) = 2.15$, $p < .05$). Furthermore, to test whether local fluctuations in attention could further explain our findings, we used a serial autocorrelation regression analysis to predict our discriminator component amplitudes ($\boldsymbol{y}$) on the current trial using those on the immediately preceding five trials and found no significant effects (all $p > 0.1$). Taken together, these results provide compelling evidence that our observed effects could not be explained purely by changes in attention.

To ensure that accuracy, which was previously shown to correlate partially with decision confidence (Vickers and Packer, 1982; Vickers et al., 1985), is not responsible for the reported effects, we performed two additional control analyses. First, we used SR absent trials, which contained trial-to-trial accuracy information and trained a separate classifier to discriminate between correct and incorrect trials. If our confidence effects were a mere manifestation of differences between correct and incorrect trials then classification performance would have been comparable to that obtained along the confidence dimension. Instead, classifier performance was significantly reduced relative to our SR waived-vs.-SR selected discrimination (Fig. 2B, $t(18) = 5.1$, $p < .001$, paired t-test). In addition, we correlated the resulting trial-by-trial component amplitudes with those estimated for the same set of trials (SR absent) using the spatial projections from the original confidence discrimination and found only a partial correlation ($r = 0.21$, $p < .001$).

Finally, we performed an analysis in which we subjected the data through the same neural generators (i.e. spatial projections) estimated for the confidence discrimination and partitioned our trials in two groups in a way that ensured accuracy remained constant while confidence was altered across the groups. Specifically, we compared component activity between correct SR waived trials (confident trials) and correct SR absent trials (which, were on average, less confident as they contained a mixture of confident and non-confident choices). We found that the component amplitudes for the more confident group of trials were significantly higher with persistent effects across all levels of sensory evidence (Fig. 2G, paired t-tests, all $p < .001$). Correspondingly, the temporal profile of this component activity revealed that accumulation slopes for more confident trials were significantly higher across all levels of sensory evidence (Fig. 3D, paired t-tests, all $p < .01$), providing further support that confidence is reflected in the rate of evidence accumulation itself. Taken together, these results endorse the notion that our reported confidence effects could not be explained purely by differences in decision accuracy.

Discussion

Here, we used a multivariate single-trial EEG approach, coupled with a distributed source reconstruction technique, to provide a mechanistic account on how decision confidence is represented in the human brain. We showed that a neural representation of confidence arises as early as the decision process itself and becomes progressively more robust as the decision unfolds, culminating shortly before one commits to a choice. Importantly, we demonstrated that this representation is reflected in the rate of evidence accumulation, thereby linking the development of choice confidence to the same neural mechanism used to form the decision itself. Consistent with this interpretation, source reconstruction placed confidence-related activity in regions previously implicated in evidence accumulation and decision making in human prefrontal and parietal cortices (Filimon et al., 2013; Heekeren et al., 2006; Ploran et al., 2007; Tosoni et al., 2008).

Together, these findings lend support to the idea that there exists a general-purpose decision making network involved in accumulating evidence for a decision while simultaneously encoding the confidence in that decision. Overall, our findings are in line with a recent report showing that neurons in lateral intraparietal cortex of the primate brain represent the formation of the decision as well as the degree of confidence underlying that decision (Kiani and Shadlen, 2009). Similarly, a growing body of evidence from animal neurophysiology suggests that when the brain forms a decision it does so in a way that resembles a Bayesian inference, in the sense that even for binary choices, a decision is formed by sampling and gradually accruing information from probability distributions rather than single estimates representing each of the alternatives (Ma et al., 2006; Zemel et al., 1998). In this framework, a measure of confidence arising directly from the decision process itself can therefore be thought of as a graded quantity, representing degree of belief that an impending choice will be correct.

Key to establishing a quantitative association between decision confidence and neural activity was our ability to exploit single-trial

information within each class of nominally identical stimuli, thereby controlling for confounding effects of stimulus difficulty and attention. Specifically, we demonstrated that trial-by-trial fluctuations in confidence-related neural activity remained predictive of opt-out behavior even after accounting for the overall amount of task difficulty as well as when extracted and tested separately for each level of sensory evidence. Similarly, we addressed the possibility that our confidence effects merely reflected changes in participants' attentional state on each trial, either prior to, or during the decision process.

In doing so, we considered two neural measures, which have previously been hypothesized to reflect top-down influences of attention on the decision process during visual discriminations, and investigated the extent to which they predicted participants' choice confidence (i.e., opt-out behavior). Importantly, we showed that neither of these measures hindered the explanatory power of the confidence discriminating neural activity. Likewise, we also showed that local fluctuations in attention across trials, as assessed via a serial autocorrelation regression analysis, could not provide an adequate account of our findings. While we do not dismiss the possibility that trial to trial variability in attention may exert a top-down influence on the efficiency of stimulus encoding and/or decision process, and ultimately on the level of confidence in one's choice, our findings render a purely attentional account of the observed confidence effects unlikely.

In another control analysis, we also tested the potential influence of decision accuracy on the observed confidence-related effects. Specifically, we demonstrated that discrimination performance along our confidence dimension was significantly higher than that of a separate classifier trained to discriminate accuracy (correct-vs.-error SR absent trials). Correspondingly, spatial discrimination projections estimated from the two classifiers and applied independently to SR absent trials produced single-trial component amplitudes that were only partially correlated (Vickers and Packer, 1982; Vickers et al., 1985). Finally, confidence effects persisted even after accuracy was accounted for (i.e. comparing confident vs. less-confidence trials), confirming that accuracy alone could not explain our confidence effects during the decision phase.

Although we designed our experiment to discourage explicit updating of reward expectations (i.e. we did not provide feedback as to whether a choice was correct or not) it remains possible that our representation of choice confidence can be explained by the expected value of the chosen option in so far as the latter is correlated with one's belief that their choice is correct. This alternative interpretation, however, is unlikely, as expected value signals are thought to reflect the consequence of the decision process, in the sense that value can only be encoded once a decision has been formed (Glascher et al., 2009). Here, we clearly demonstrated that confidence arises with the decision process itself, and it develops gradually as the process unfolds. In addition, the regions we found associated with choice confidence are outside the human reward network, which is known to reflect expected reward and value signals (Dreher et al., 2006; Kable and Glimcher, 2007; Knutson et al., 2005; Philiastides et al., 2010; Rangel et al., 2008; Rolls et al., 2008; Rushworth and Behrens, 2008).

Our findings that confidence signals appear as early as the process of evidence accumulation itself constitute strong evidence against a purely metacognitive (post-decisional) account of decision confidence, consistent with a recent report of pre-decisional signals of self-reported confidence during perceptual choices (Graziano et al., 2014; Zizlsperger et al., 2014). Importantly, however, our results do not exclude the possibility that confidence representations persist beyond the decision point and after a behavioral choice was made (Fleming et al., 2012; Pleskac and Busemeyer, 2010). Nonetheless, these metacognitive representations are captured using post-decisional subjective confidence reports, which are likely to be subjected to additional influences arriving after the decision point (e.g. internal noise, expected reward etc.). In addition, the extent to which these post-decisional signals influence metacognitive assessment and subsequent choices remains unclear. Future studies designed to investigate how pre- and post-decisional confidence signals interact to shape behavior would be necessary. In particular, understanding how confidence traces arising from the process of decision formation are communicated to regions implicated in metacognitive appraisal would be required (De Martino et al., 2013; Fleming et al., 2012; Hebart et al., 2014).

In summary, choice confidence represents the degree of belief that one's actions are likely to be correct and as such can play a critical role in how we interact with the world around us. Here, we provided a mechanistic account on how confidence is represented in the human brain and provided strong evidence that linked the development of choice confidence to the same mechanism and neural generators used to form the decision itself. These results could provide the foundation upon which future computational studies could continue to interrogate the mechanistic details of the influence of confidence on decision making (Zylberberg et al., 2012). Crucially, our findings coupled with our ability to exploit the relevant neural signatures non-invasively and on a trial-by-trial basis, may have direct implications for decision-making problems that rely on inconclusive or partially ambiguous evidence. Specifically, they can provide the platform for developing cognitive interfaces that can help facilitate, and ultimately optimize decision making (Sajda et al., 2007, 2009).

Acknowledgements

This work was supported by a British Academy (SG121587) and Royal Society (RG110054) research grants to MGP.

Conflict of interest

The authors declare no competing financial interests.

References

Blank, H., Biele, G., Heekeren, H.R., Philiastides, M.G., 2013. Temporal characteristics of the influence of punishment on perceptual decision making in the human brain. J. Neurosci. 33, 3939–3952.
Bogacz, R., Brown, E., Moehlis, J., Holmes, P., Cohen, J.D., 2006. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychol. Rev. 113, 700–765.
Cheadle, S., Wyart, V., Tsetsos, K., Myers, N., de Gardelle, V., Herce Castanon, S., Summerfield, C., 2014. Adaptive Gain Control during Human Perceptual Choice. Neuron 81, 1429–1441.
Dakin, S.C., Hess, R.F., Ledgeway, T., Achtman, R.L., 2002. What causes non-monotonic tuning of fMRI response to noisy images? Curr. Biol. 12, R476–R477 (author reply R478).
de Lange, F.P., Jensen, O., Dehaene, S., 2010. Accumulation of evidence during sequential decision making: the importance of top-down factors. J. Neurosci. 30, 731–738.
De Martino, B., Fleming, S.M., Garrett, N., Dolan, R.J., 2013. Confidence in value-based choice. Nat. Neurosci. 16, 105–110.
Ding, L., Gold, J.I., 2013. The basal ganglia's contributions to perceptual decision making. Neuron 79, 640–649.
Donner, T.H., Siegel, M., Oostenveld, R., Fries, P., Bauer, M., Engel, A.K., 2007. Population activity in the human dorsal pathway predicts the accuracy of visual motion detection. J. Neurophysiol. 98, 345–359.
Donner, T.H., Siegel, M., Fries, P., Engel, A.K., 2009. Buildup of choice-predictive activity in human motor cortex during perceptual decision making. Curr. Biol. 19, 1581–1585.
Dreher, J.C., Kohn, P., Berman, K.F., 2006. Neural coding of distinct statistical properties of reward information in humans. Cereb. Cortex 16, 561–573.
Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern classification. John Wiley, New York (Section 10, 1).
Filimon, F., Philiastides, M.G., Nelson, J.D., Kloosterman, N.A., Heekeren, H.R., 2013. How embodied is perceptual decision making? Evidence for separate processing of perceptual and motor decisions. J. Neurosci. 33, 2121–2136.
Fleming, S.M., Huijgen, J., Dolan, R.J., 2012. Prefrontal contributions to metacognition in perceptual decision making. J. Neurosci. 32, 6117–6125.
Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J., 2008. Multiple sparse priors for the M/EEG inverse problem. Neuroimage 39, 1104–1120.
Glascher, J., Hampton, A.N., O'Doherty, J.P., 2009. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. Cereb. Cortex 19, 483–495.
Graziano, M., Parra, L.C., Sigman, M., 2014. Neural Correlates of Perceived Confidence in a Partial Report Paradigm. J Cogn Neurosci http://dx.doi.org/10.1162/jocn_a_00759.
Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. Annu. Rev. Neurosci. 30, 535–574.

Hebart, M.N., Schriever, Y., Donner, T.H., Haynes, J.D., 2014. The Relationship between Perceptual Decision Variables and Confidence in the Human Brain. Cereb. Cortex http://dx.doi.org/10.1093/cercor/bhu181.
Heekeren, H.R., Marrett, S., Bandettini, P.A., Ungerleider, L.G., 2004. A general mechanism for perceptual decision-making in the human brain. Nature 431, 859–862.
Heekeren, H.R., Marrett, S., Ruff, D.A., Bandettini, P.A., Ungerleider, L.G., 2006. Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. Proc. Natl. Acad. Sci. U. S. A. 103, 10023–10028.
Heekeren, H.R., Marrett, S., Ungerleider, L.G., 2008. The neural systems that mediate human perceptual decision making. Nat. Rev. Neurosci. 9, 467–479.
Ho, T.C., Brown, S., Serences, J.T., 2009. Domain general mechanisms of perceptual decision making in human cortex. J. Neurosci. 29, 8675–8687.
Jeffreys, D.A., 1996. Evoked potential studies of face and object processing. Vis. Cogn. 3, 1–38.
Jordan, M.I., Jacobs, R.A., 1994. Hierarchical Mixtures of Experts and the Em Algorithm. Neural Comput. 6, 181–214.
Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. Nat. Neurosci. 10, 1625–1633.
Kelly, S.P., O'Connell, R.G., 2013. Internal and external influences on the rate of sensory evidence accumulation in the human brain. J. Neurosci. 33, 19434–19441.
Kepecs, A., Uchida, N., Zariwala, H.A., Mainen, Z.F., 2008. Neural correlates, computation and behavioural impact of decision confidence. Nature 455, 227–231.
Kiani, R., Shadlen, M.N., 2009. Representation of confidence associated with a decision by neurons in the parietal cortex. Science 324, 759–764.
Kim, J.N., Shadlen, M.N., 1999. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. Nat. Neurosci. 2, 176–185.
Knutson, B., Taylor, J., Kaufman, M., Peterson, R., Glover, G., 2005. Distributed neural representation of expected value. J. Neurosci. 25, 4806–4812.
Liu, J., Higuchi, M., Marantz, A., Kanwisher, N., 2000. The selectivity of the occipitotemporal M170 for faces. Neuroreport 11, 337–341.
Ma, W.J., Beck, J.M., Latham, P.E., Pouget, A., 2006. Bayesian inference with probabilistic population codes. Nat. Neurosci. 9, 1432–1438.
Mazaheri, A., Jensen, O., 2006. Posterior alpha activity is not phase-reset by visual stimuli. Proc. Natl. Acad. Sci. U. S. A. 103, 2948–2952.
Mazurek, M.E., Roitman, J.D., Ditterich, J., Shadlen, M.N., 2003. A role for neural integrators in perceptual decision making. Cereb. Cortex 13, 1257–1269.
Mulder, M.J., van Maanen, L., Forstmann, B.U., 2014. Perceptual decision neurosciences - A model-based review. Neuroscience 277C, 872–884.
Newsome, W.T., Britten, K.H., Movshon, J.A., 1989. Neuronal correlates of a perceptual decision. Nature 341, 52–54.
O'Connell, R.G., Dockree, P.M., Kelly, S.P., 2012. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. Nat. Neurosci. 15, 1729–1735.
Parra, L., Alvino, C., Tang, A., Pearlmutter, B., Yeung, N., Osman, A., Sajda, P., 2002. Linear spatial integration for single-trial detection in encephalography. Neuroimage 17, 223–230.
Parra, L.C., Spence, C.D., Gerson, A.D., Sajda, P., 2005. Recipes for the linear analysis of EEG. Neuroimage 28, 326–341.
Peirce, J.W., 2007. PsychoPy–Psychophysics software in Python. J. Neurosci. Methods 162, 8–13.
Philiastides, M.G., Sajda, P., 2006. Temporal characterization of the neural correlates of perceptual decision making in the human brain. Cereb. Cortex 16, 509–518.
Philiastides, M.G., Sajda, P., 2007. EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. J. Neurosci. 27, 13082–13091.
Philiastides, M.G., Ratcliff, R., Sajda, P., 2006. Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. J. Neurosci. 26, 8965–8975.
Philiastides, M.G., Biele, G., Heekeren, H.R., 2010. A mechanistic account of value computation in the human brain. Proc. Natl. Acad. Sci. U. S. A. 107, 9430–9435.
Philiastides, M.G., Heekeren, H.R., Sajda, P., 2014. Human scalp potentials reflect a mixture of decision-related signals during perceptual choices. J. Neurosci. http://dx.doi.org/10.1523/JNEUROSCI.3012-14.2014 (in press).
Pleskac, T.J., Busemeyer, J.R., 2010. Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. Psychol. Rev. 117, 864–901.
Ploran, E.J., Nelson, S.M., Velanova, K., Donaldson, D.I., Petersen, S.E., Wheeler, M.E., 2007. Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. J. Neurosci. 27, 11912–11924.
Rangel, A., Camerer, C., Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision making. Nat. Rev. Neurosci. 9, 545–556.
Ratcliff, R., Philiastides, M.G., Sajda, P., 2009. Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. Proc. Natl. Acad. Sci. U. S. A. 106, 6539–6544.
Rolls, E.T., McCabe, C., Redoute, J., 2008. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. Cereb. Cortex 18, 652–663.
Rushworth, M.F., Behrens, T.E., 2008. Choice, uncertainty and value in prefrontal and cingulate cortex. Nat. Neurosci. 11, 389–397.
Sajda, P., Gerson, A.D., Philiastides, M.G., Parra, L.C., 2007. Single-trial analysis of EEG during rapid visual discrimination: Enabling cortically-coupled computer vision. In: Dornhege, G., Mullan, J.R., Hinterberger, T., McFarland, D.J., Muller, K.R. (Eds.), Toward Brain-Computer Interfacing. MIT Press, Cambridge, MA, pp. 423–439.
Sajda, P., Philiastides, M.G., Parra, L.C., 2009. Single-trial analysis of neuroimaging data: inferring neural networks underlying perceptual decision-making in the human brain. IEEE Rev. Biomed. Eng. 2, 97–109.
Shadlen, M.N., Kiani, R., 2013. Decision making as a window on cognition. Neuron 80, 791–806.
Shadlen, M.N., Newsome, W.T., 2001. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. 86, 1916–1936.
Shadlen, M.N., Britten, K.H., Newsome, W.T., Movshon, J.A., 1996. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J. Neurosci. 16, 1486–1510.
Tallon-Baudry, C., Bertrand, O., Delpuech, C., Pernier, J., 1996. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J. Neurosci. 16, 4240–4249.
Thut, G., Nietzel, A., Brandt, S.A., Pascual-Leone, A., 2006. Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. J. Neurosci. 26, 9494–9502.
Tosoni, A., Galati, G., Romani, G.L., Corbetta, M., 2008. Sensory-motor mechanisms in human parietal cortex underlie arbitrary visual decisions. Nat. Neurosci. 11, 1446–1453.
Troje, N.F., Bulthoff, H.H., 1996. Face recognition under varying poses: the role of texture and shape. Vision Res. 36, 1761–1771.
van Dijk, H., Schoffelen, J.M., Oostenveld, R., Jensen, O., 2008. Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. J. Neurosci. 28, 1816–1823.
van Maanen, L., Brown, S.D., Eichele, T., Wagenmakers, E.J., Ho, T., Serences, J., Forstmann, B.U., 2011. Neural correlates of trial-to-trial fluctuations in response caution. J. Neurosci. 31, 17488–17495.
Vickers, D., Packer, J., 1982. Effects of alternating set for speed or accuracy on response time, accuracy and confidence in a unidimensional discrimination task. Acta Psychol. (Amst) 50, 179–197.
Vickers, D., Smith, P., Burt, J., Brown, M., 1985. Experimental Paradigms Emphasizing State or Process Limitations.2. Effects on Confidence. Acta Psychol. (Amst) 59, 163–193.
Wyart, V., Tallon-Baudry, C., 2009. How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. J. Neurosci. 29, 8715–8725.
Wyart, V., de Gardelle, V., Scholl, J., Summerfield, C., 2012. Rhythmic fluctuations in evidence accumulation during decision making in the human brain. Neuron 76, 847–858.
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.
Zemel, R.S., Dayan, P., Pouget, A., 1998. Probabilistic interpretation of population codes. Neural Comput. 10, 403–430.
Zizlsperger, L., Sauvigny, T., Handel, B., Haarmeier, T., 2014. Cortical representations of confidence in a visual perceptual decision. Nat. Commun. 5, 3940.
Zylberberg, A., Barttfeld, P., Sigman, M., 2012. The construction of confidence in a perceptual decision. Front. Integr. Neurosci. 6, 79.