**Supporting Information: Appendix**

Distinct encoding of decision confidence in human medial prefrontal cortex

Dan Bang* and Stephen M. Fleming

*Correspondence: danbang.db@gmail.com

## Methods

### Task

*Trial events and timings*

Each trial began with the presentation of a fixation cross at the centre of a circular aperture. After a uniformly sampled delay (prescan: .5-1 s; scan: 1-4 s), subjects viewed a field of moving dots (1 s). Subjects were instructed to fixate during stimulus presentation. After the motion stimulus, subjects were presented with a direction reference which transected the aperture. Subjects were required to press one of two buttons to indicate whether the net direction of dot motion (angle as measured from aperture centre) was clockwise (CW) or counter-clockwise (CCW) to the reference. As a visual aid, the response buttons and the associated arcs of the aperture were coloured orange and blue, with colour assignment counterbalanced across subjects. Once a choice had been made, the colour of the central cross (orange or blue) confirmed the decision (.25 s). In the prescan session, subjects were asked to estimate the confidence in their choice on every trial. In the scan session, subjects were asked to estimate the confidence in their choice every 5-10 trials (15% of trials). Subjects responded by moving a marker along a scale from 50% to 100% in steps of 10%. The marker started randomly in one of the six locations along the scale and was controlled by button press. Once a response had been registered, the marker turned grey (.5 s), before the next trial started.

*Motion stimulus*

The motion stimulus was made up of three sets of dots (each dot was 0.12 degrees in diameter) shown in consecutive frames inside the circular aperture (8 degrees in diameter) centred on the fixation cross (0.2 degrees in diameter). Each set of dots was shown for one frame (about 16 ms) and then replotted again three frames later (about 50 ms) – some dots were displaced in the specified motion direction at a speed of 2 degrees $s^{-1}$ while the rest of the dots were displaced at random locations within the aperture. We refer to the percentage of dots displayed in the specified motion direction as coherence, $k$. The dot density was fixed at 16 dots $degrees^{-2}$ $s^{-1}$. To help subjects maintain fixation, a circular region (0.7 degrees in diameter) at the centre of the aperture was kept free of dots. The motion direction was sampled uniformly from the range 1-360 degrees. The direction of the reference (0.8 degrees in length and 0.08 degrees in width) was within 45 degrees of the motion direction. We refer to the difference between the motion direction and the direction reference as signed distance, $\delta$, with positive values for CW and negative values for CCW. A set of coherences, **K**, and a set of signed distances, **Δ**, were calibrated for each subject and crossed in a factorial design in each session.

*Prescan calibration*

The aim of the calibration procedure was to identify a pair of coherences associated with different levels of sensory reliability and a set of signed distances associated with different levels of boundary difficulty. This aim was achieved in two phases. We did not elicit confidence estimates in either phase. In phase 1, subjects performed the task with a fixed absolute distance, **Δ**: {-10,10}. Coherence was .20 for the first two trials and then updated using a 2-down-1-up procedure: after two correct decisions, coherence was decreased by .01; after one incorrect decision, coherence was increased by .01. We used the median coherence of the last 20 trials of phase 1 to specify a medium coherence, $k_{med}$, constrained to a lower limit of .12 and an upper limit of .50. In phase 2, subjects performed the task with coherence fixed at $k_{med}$. We employed a range of signed distances, **Δ**: {-45, -24, -12, -6, -3, 3, 6, 12, 24, 45}, and fitted a psychometric function for each subject using logit regression. This procedure provided us with a measure of a subject's choice bias and sensitivity. We corrected the psychometric function for choice bias and used the corrected psychometric function to infer a set of positive distances associated with the target accuracies 60%, 72.5% and 85%: $\delta_{60}$, $\delta_{72.5}$, $\delta_{85}$. For the main experiment, we defined the set of coherences as, **K**: {$k_{low}$ = .5 x $k_{med}$, $k_{high}$ = 2 x $k_{med}$}, and the set of signed distances as, **Δ**: {-45, -$\delta_{85}$, -$\delta_{72.5}$, -$\delta_{60}$, $\delta_{60}$, $\delta_{72.5}$, $\delta_{85}$, 45}. The extreme signed distances were probed three times less often; they served as anchor points for construction of psychometric functions. Prescan calibration data are shown in Fig. S1.

*Scan calibration*

Subjects performed one further calibration phase at the start of the scan session during the acquisition of structural images. Coherence was fixed at $k_{med}$ and we calibrated the signed distances associated with 60% and 85% accuracy, $\pm\delta_{60}$ and $\pm\delta_{85}$, using a QUEST procedure (1). For the main experiment, we defined the set of coherences as **K**: {$k_{low}$, $k_{high}$} and the set of signed distances as **Δ**: {-$\delta_{85}$, -$\delta_{60}$, $\delta_{60}$, $\delta_{85}$}. Scan calibration data are shown in Fig. S1.

*Prescan procedure*
The prescan session consisted of five runs. Subjects first completed a tutorial, with motion stimuli shown for variable durations, 1-5 s, and variable coherences, **K**: {.30, 60}. Subjects then practiced the task (40 trials) at high coherence, **K**: {.30, .60}, and high absolute distance, **Δ**: {-30, 30}. In this run only, subjects received trial-by-trial feedback (central cross briefly turned green after correct choices and red after incorrect choices), with the aim to familiarise subjects with direction judgements in continuous space. Next, subjects performed calibration phases 1 (120 trials) and 2 (260 trials). Finally, subjects performed the main experiment (540 trials). The prescan session lasted 2 hours.

*Scan procedure*
The scan session was made up of seven runs. Subjects first performed the calibration phase during the acquisition of structural images (180 trials) and then the main experiment over five runs (5 x 112 = 560 trials). In the scan session only, sampling of motion direction was yoked such that there was a 50% probability that a given trial had the same motion direction as on the previous trial. This manipulation is orthogonal to coherence and distance, and direction does not tell subjects whether the upcoming choice should be CW or CCW. Effects of repetition on direction-specific neural responses will be analysed in a separate paper. In the final scan run, subjects viewed alternating displays (12 s) of static and dynamic dots, $k$ = .50 (2 x 12 = 24 displays).

**Hierarchical drift-diffusion model**
Decision formation was modelled using the drift-diffusion model (DDM). The DDM models two-choice decision making as a process of accumulating noisy evidence over time with a certain speed, or drift rate ($v$), until one of two decision thresholds is crossed and the associated response is executed. Larger threshold separation ($a$) leads to slower responses but more accurate responding. The DDM includes a nondecision component ($t$) which captures time needed for stimulus encoding and motor execution. We employed hierarchical Bayesian estimation of subjects' DDM parameters using the HDDM toolbox (2) (http://ski.clps.brown.edu/hddm_docs/). We fitted non-decision time ($t$), decision threshold ($a$) and drift rate ($v$) separately for each condition of our factorial design (prescan: 8 conditions; scan: 4 conditions). We also included intertrial variability in drift rate ($sv$) and nondecision time ($st$) as free parameters that were constant across conditions. The HDDM toolbox applies Markov Chain Monte-Carlo (MCMC) sampling methods to approximate posterior distributions over the DDM parameters; we ran 1 chain with 5000 samples, with the first 1000 samples discarded as burn-in. We extracted mean group-level posterior estimates for visualisation of DDM parameters in each condition (Fig. S4) and generating posterior predictive values (Fig. 2*A*-*B*). We used the *simuldiff* function from the DMAT toolbox (3) (http://ppw.kuleuven.be/okp/software/dmat/) to generate posterior predictive values: for each condition, we simulated 100,000 trials under the hierarchically estimated DDM parameters and calculated mean choice accuracy and mean choice reaction time in that condition.

**FMRI**
*Acquisition*
MRI data were acquired on a 3T Siemens Trio scanner with a 32-channel head coil. T1-weighted structural images were acquired using a 3D MDEFT sequence: 1x1x1 mm resolution voxels; 176 sagittal slices, 256x224 matrix; TR = 10.55ms; TE = 3.14ms; TI = 680ms. BOLD $T2^*$-weighted functional images were acquired using a gradient-echo EPI pulse sequence: 3x3x3 mm resolution voxels; 48 transverse slices, 64x74 matrix; TR = 3.36; TE = 30ms; slice tilt = 0 degrees, slice thickness = 2 mm; interslice gap = 1mm; ascending slice order. Field maps were acquired using a double-echo FLASH (gradient-echo) sequence: TE1 = 10ms; TE2 = 12.46ms; 64 slices were acquired with 2 mm slice thickness and a 1 mm gap; in-plane field of view is 192x192 $mm^2$ with 3x3 $mm^2$ resolution.

*Preprocessing*
The first 4 volumes of each functional run were discarded to allow for T1 equilibration. Functional images were slice-time corrected, realigned and unwarped using the field maps (4). Structural T1-weighted images were coregistered to the mean functional image of each subject using the iterative mutual-information algorithm. Each subject's structural image was segmented into grey matter, white matter and cerebral spinal fluid using a nonlinear deformation field to map it onto a template tissue probability map (5). These deformations were applied to structural and functional images to create new images spatially normalised to the Montreal Neurological Institute (MNI) space and interpolated to 2x2x2 mm voxels. Normalized images were spatially smoothed using a Gaussian kernel with full-width half-maximum of 8mm. The motion correction parameters estimated from the realignment procedure

and their first temporal derivatives – 12 'motion' regressors in total – were included as confounds in the first-level analysis for each subject.

*Physiological monitoring*
Peripheral measurements of a subject's pulse and breathing were made together with scanner slice synchronisation pulses using a Spike2 data acquisition system (Cambridge Electronic Design Limited, Cambridge UK). The cardiac pulse signal was measured using an MRI compatible pulse oximeter (Model 8600 F0, Nonin Medical, Inc. Plymouth, MN) attached to a subject's finger. The respiratory signal, thoracic movement, was monitored using a pneumatic belt positioned around the abdomen close to the diaphragm. A physiological noise model was constructed to account for artifacts related to cardiac and respiratory phase and changes in respiratory volume using an in-house MATLAB toolbox (6). Models for cardiac and respiratory phase and their aliased harmonics were based on RETROICOR (7) and a similar, earlier method (8). Basis sets of sine and cosine Fourier series components extending to the 3rd harmonic were used to model physiological fluctuations. Additional terms were included to model changes in respiratory volume (9, 10) and heart rate (11). This procedure yielded a total of 14 'biophysical' regressors which were sampled at a reference slice in each image volume. The regressors were included as confounds in the first-level analysis for each subject.

*Regions of interest*
ROI masks for bilateral IPS, pre-SMA and pgACC were created by applying a leave-one-out procedure to activations from GLM1. In particular, for each subject, we performed second-level analyses excluding the subject's data and used the resulting second-level *t* maps (cluster-defining threshold: $P < 0.001$, uncorrected) to specify the subject's ROI masks. This ROI procedure ensured unbiased ROI selection as each subject's neural activity did not contribute to the selection of ROIs for that subject. We used clusters associated with a main effect of coherence (bilateral IPS), a main effect of distance (pre-SMA), and the coherence x distance interaction (pgACC). ROI masks for MT+ were based on the localiser scan; we created a bilateral group mask using the second-level contrast between dynamic and static motion, and then, for each subject, created bilateral MT+ masks (8-mm sphere) around the subject-specific peaks inside the bilateral group mask. ROI masks for bilateral ventral striatum were specified using the Oxford-GSK-Imanova Striatal Structural atlas included with FSL. ROI masks for bilateral rlPFC were specified using the maps provided by Neubert and colleagues (12) (union of '46' and 'fpl').

*Activity time courses*
We transformed each ROI mask from MNI to native space and extracted preprocessed BOLD time courses as the average of voxels within the mask. For each scan run, we regressed out variation due to head motion and biophysical responses, applied a high-pass filter (128 s cut-off) to remove low-frequency drifts, and oversampled the BOLD time course by a factor of ~23 (time resolution of .144 s). For each trial, we extracted activity estimates in a 12 s window (84 time points), time-locked to 1 s before the onset of the motion stimulus or 2 s before the onset of the confidence scale (only for reward-magnitude analysis). We used linear regression to predict ROI activity time courses. More specifically, we applied a linear regression to each time point and then, by concatenating beta-weights across time points, created a beta-weight time course for each predictor of a regression model. We performed this step separately for each subject and pooled beta-weight time courses across subjects for visualisation. We tested group-level significance using a permutation procedure. We repeated the above steps 1,001 times, for each repetition shuffling a subject's trial time courses, and then asked, at each time point, whether the *t*-statistics associated with the empirically observed group-level effect (one-sample *t* test of beta-weights pooled across subjects against zero) was smaller or larger than 97.5% of the *t* statistics obtained from the permutation procedure.

*Single-trial activity estimates*
We estimated single-trial ROI activity estimates as a beta time series. This was achieved by means of an event-related model with a separate regressor for each trial. Regressors were boxcars time-locked to the onset of the motion display and spanning 1 s (Fig. 4*B*) or until choice (Fig. 4*D*) (13). Each regressor was convolved with a canonical hemodynamic response function. We included motion and biophysical parameters as 'nuisance' regressors. Regressors were modelled separately for each scan run and constants were included to account for between-run differences in mean activation and scanner drifts. A high-pass filter (128 s cut-off) was applied to remove low-frequency drifts. One important consideration in using single-trial activity estimates is that a beta for a given trial can be strongly affected by acquisition artefacts that occur together with that trial (e.g., scanner pulse artefacts). Therefore, for each subject, we computed the grand-mean beta estimate across both voxels and trials and excluded

any trial whose mean beta estimate across voxels was 3 SDs below or above this grand mean (14). About 2% of trials were excluded per subject. Finally, we used the ROI masks to extract ROI single-trial activity estimates.

*Software*
MRI data were preprocessed and analysed using SPM12 (www.fil.ion.ucl.ac.uk/spm).

**References**


1. Watson AB, Pelli DG (1983) QUEST: A Bayesian adaptive psychometric method. *Percept Psychophys* 33(2):113–120.
2. Wiecki T V., Sofer I, Frank MJ (2013) HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Front Neuroinform* 7(August):14.
3. Vandekerckhove J, Tuerlinckx F (2008) Diffusion model analysis with MATLAB: A DMAT primer. *Behav Res Methods* 40(1):61–72.
4. Andersson JL, Hutton C, Ashburner J, Turner R, Friston K (2001) Modeling geometric deformations in EPI time series. *Neuroimage* 13(5):903–919.
5. Ashburner J, Friston KJ (2005) Unified segmentation. *Neuroimage* 26(3):839–851.
6. Hutton C, et al. (2011) The impact of physiological noise correction on fMRI at 7T. *Neuroimage* 57(1):101–112.
7. Glover GH, Li T-Q, Ress D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med* 44(1):162–167.
8. Josephs O, Howseman AM, Friston KJ, Turner R (1997) Physiological noise modelling for multi-slice EPI fMRI using SPM. *Proceedings of the 5th Annual Meeting of ISMRM* (Vancouver, Canada).
9. Birn RM, Smith MA, Jones TB, Bandettini PA (2008) The Respiration Response Function: the temporal dynamics of fMRI signal fluctuations related to changes in respiration. *Neuroimage* 40(2):644–654.
10. Birn RM, Diamond JB, Smith MA, Bandettini PA (2006) Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. *Neuroimage* 31(4):1536–1548.
11. Chang C, Glover GH (2009) Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. *Neuroimage* 47(4):1448–1459.
12. Neubert FX, Mars RB, Thomas AG, Sallet J, Rushworth MFS (2014) Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron* 81(3):700–713.
13. Grinband J, Wager TD, Lindquist M, Ferrera VP, Hirsch J (2008) Detection of time-varying signals in event-related fMRI designs. *Neuroimage* 43(3):509–520.
14. Atlas LY, Bolger N, Lindquist MA, Wager TD (2010) Brain mediators of predictive cue effects on perceived pain. *J Neurosci* 30(39):12964–12977.

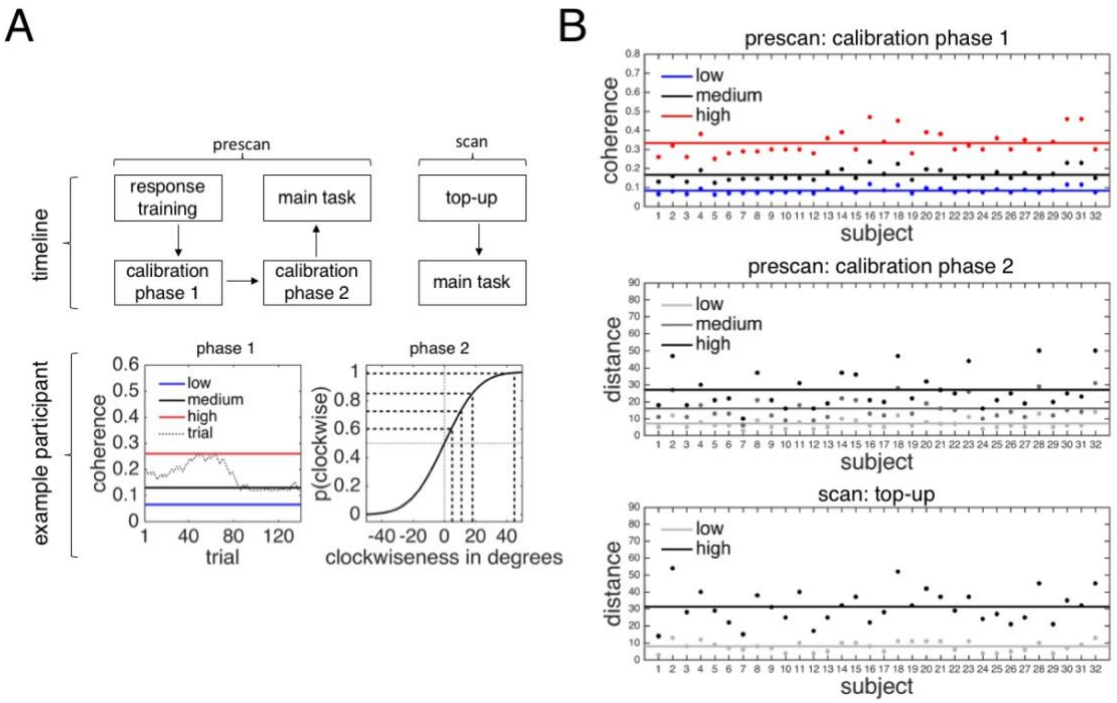


**Fig. S1.** Stimulus calibration. (*A*) Experimental procedure. (*Top*) Subjects took part in separate prescan and scan sessions (2-14 d between sessions). In prescan calibration phase 1, we calibrated coherence (medium) at a fixed distance for each subject using a 2-down-1-up procedure. We used 50% (low) and 200% (high) of the medium coherence value in the main task, in both sessions. In prescan calibration phase 2, we presented each subject with a range of distances at medium coherence so as to construct their psychometric function. We extracted a set of distances associated with target levels of accuracy for the prescan main task. In the scan top-up session, using an adaptive QUEST procedure, we recalibrated a subset of distances from the prescan session at medium coherence for the scan main task. *(Bottom*) Example data from prescan calibration phase 1 and 2. (*B*) Calibration results. Calibrated stimulus values are shown for (*top)* prescan calibration phase 1, (*middle*) prescan calibration phase 2 and (*bottom*) scan top-up.

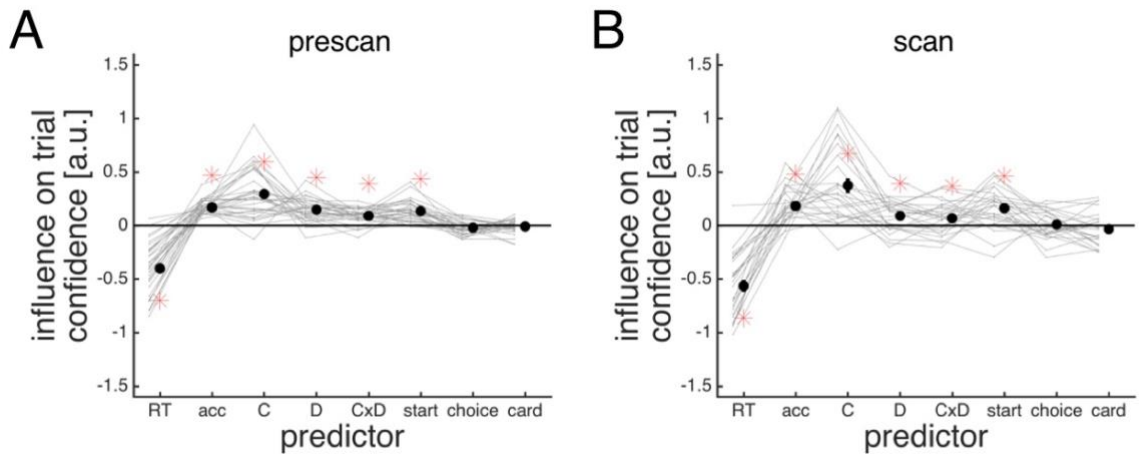


**Fig. S2.** Extended regression analysis of predictors of confidence estimates. (*A*) Prescan session. Coefficients (y-axis) from a trial-by-trial ordinal regression testing the influence of stimulus and choice features (x-axis) on current confidence. Predictors included: choice reaction time (RT), choice accuracy (acc), coherence (C), distance (D), the interaction between coherence and distance (CxD), the initial position of the confidence marker (start), whether the choice was clockwise (choice), and the cardinality of the motion direction (card). (*B*) Scan session. Same regression analysis as in panel A, except that we also included the reward factor and whether the motion direction was the same as on the previous trial; these predictors were not significant and are not shown here. Confidence estimates were elicited every 5-10 trials (84 trials in total) in the scan session. In *A*-*B*, we performed a regression for each subject and tested significance (red asterisk) by comparing the coefficients pooled across subjects to zero (one-sample *t* test). We log-transformed choice reaction time and z-scored predictors. Faint lines connect the coefficients for each subject. Data are represented as group mean ± SEM.

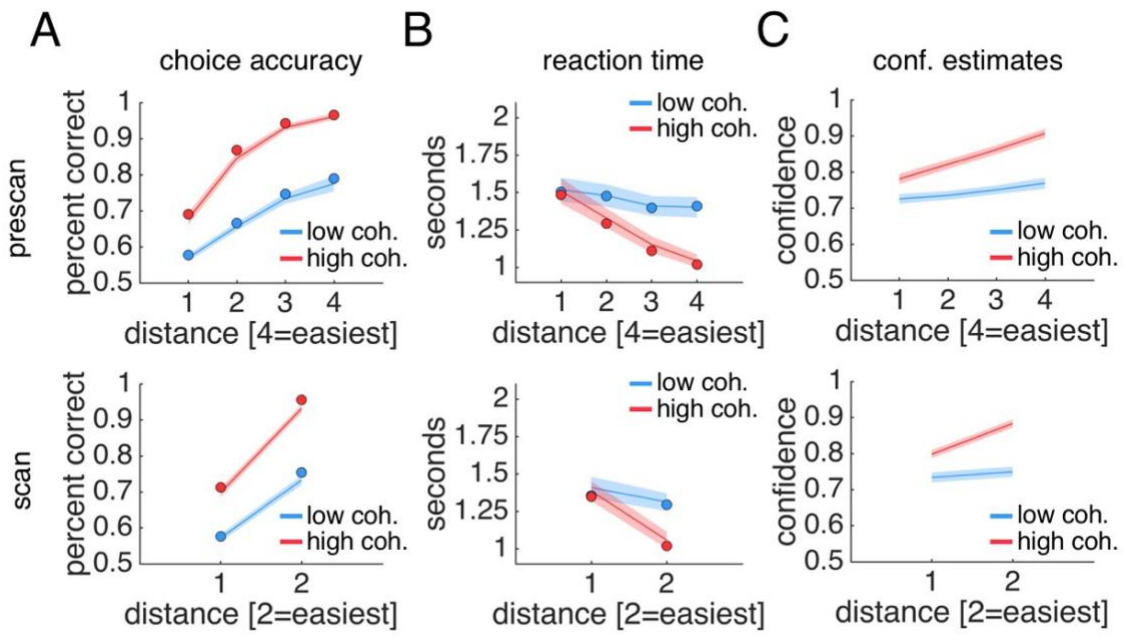


**Fig. S3.** Behavioural results and HDDM fits from prescan and scan sessions. (*A*) Choice accuracy. In the scan session, subjects were more likely to be correct when coherence was high and when distance was high [logistic regression, coherence: $t$ (31) = 13.61, $P < 0.001$, distance: $t$ (31) = 14.76, $P < 0.001$, interaction: $t$ (31) = 11.48, $P < 0.001$]. (*B*) Reaction time measured from reference onset. In the scan session, subjects made faster decisions when coherence was high and when distance was high [linear regression, coherence: $t$ (31) = -6.37, $P < 0.001$, distance: $t$ (31) = -9.66, $P < 0.001$, interaction: $t$ (31) = -8.65, $P < 0.001$]. (*C*) Confidence estimates. In the scan session, subjects were more confident when coherence was high and when distance was high [ordinal regression, coherence: $t$ (31) = 6.22, $P < 0.001$, distance: $t$ (31) = 5.87, $P < 0.001$, interaction: $t$ (31) = 4.63, $P < 0.001$]. In *A*-*B*, solid dots are posterior predictive values from the hierarchical drift-diffusion model fit to subjects' responses separately for each condition. In *A*-*C*, data are represented as mean ± SEM. See main text for test statistics for prescan session.

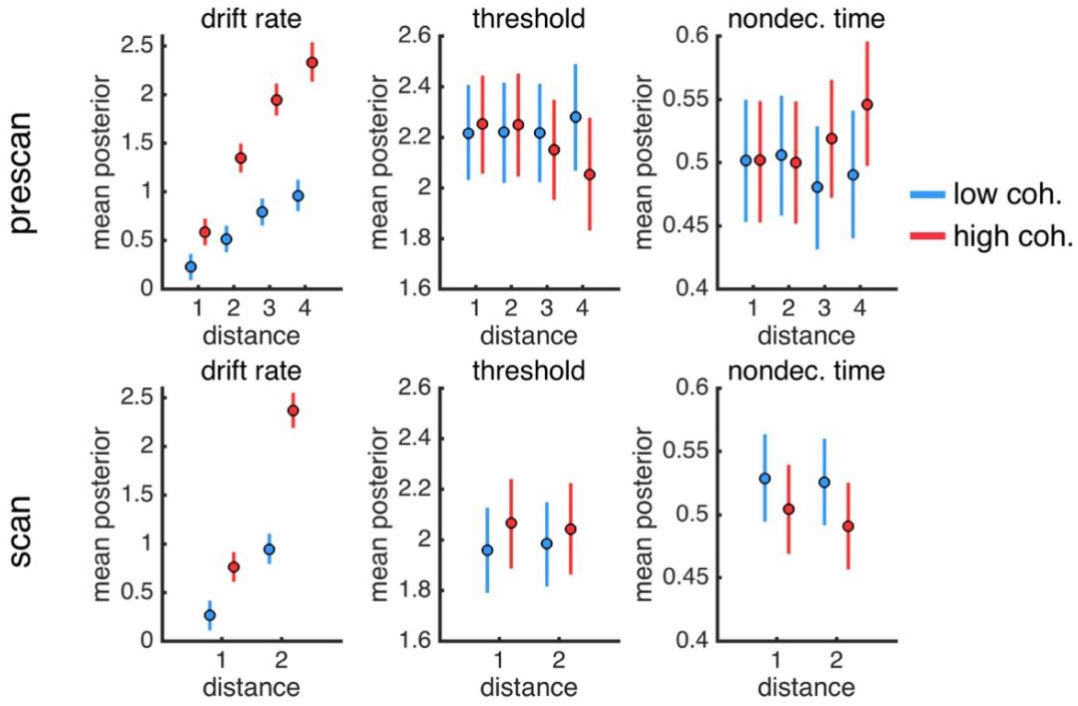


**Fig. S4.** Drift-diffusion model. Mean posterior estimates obtained from hierarchical Bayesian estimation of subjects' drift-diffusion model parameters using the HDDM toolbox for (*top*) the prescan session and (*bottom*) the scan sessions. We fitted drift rate (*v*), decision threshold (*a*) and non-decision time (*t*) separately for each condition of our factorial design (prescan: 8 conditions; scan: 4 conditions). We also included intertrial variability in drift rate (*sv*) and nondecision time (*st*) but these parameters were fitted across conditions (prescan, *sv* = 0.904, *st* = 0.232; scan, *sv* = 1.063, *st* = 0.270). Error bars indicate 95%-confidence intervals as estimated from the posterior distributions over parameters.

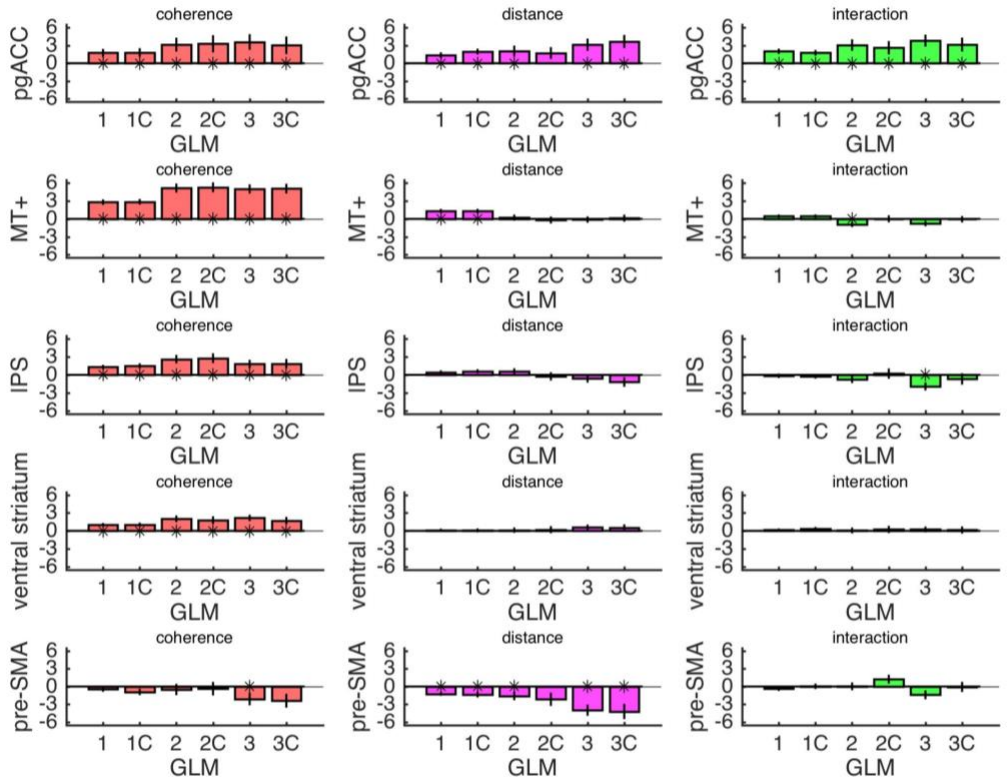


**Fig. S5.** ROI contrast estimates from whole-brain control models. The plots show contrast estimates (y-axis) for coherence, distance and the coherence x distance interaction extracted from a set of control GLMs (x-axis). We tested significance (black asterisk) by comparing the contrast estimates pooled across subjects against zero (one-sample *t* test). The rationale behind the three families of GLMs is as follows: GLM1s assume that choice reaction time is not a confound but reflects relevant neural processing; GLM2s control for within-condition variation in choice reaction time; and GLM3s control for between-condition variation in choice reaction time. C: correct trials only. Data are represented as group mean ± SEM.

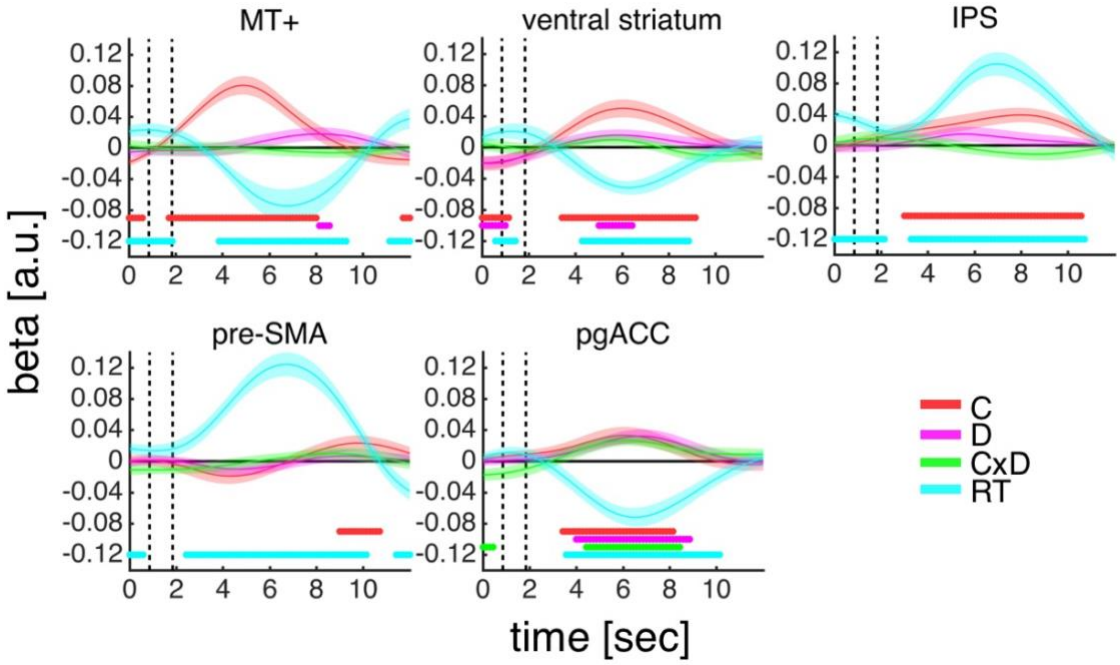


**Fig. S6.** Encoding of factors in ROI activity time courses. General linear model analysis of the effects of coherence (C), distance (D), the coherence x distance interaction (CxD) and choice reaction time (RT) on ROI activity time courses. Dots below time course indicate significant excursion of *t* statistics assessed using two-tailed permutation tests. Vertical dashed lines indicate the onsets of the motion stimulus and the choice phase. Data are represented as group mean ± SEM.

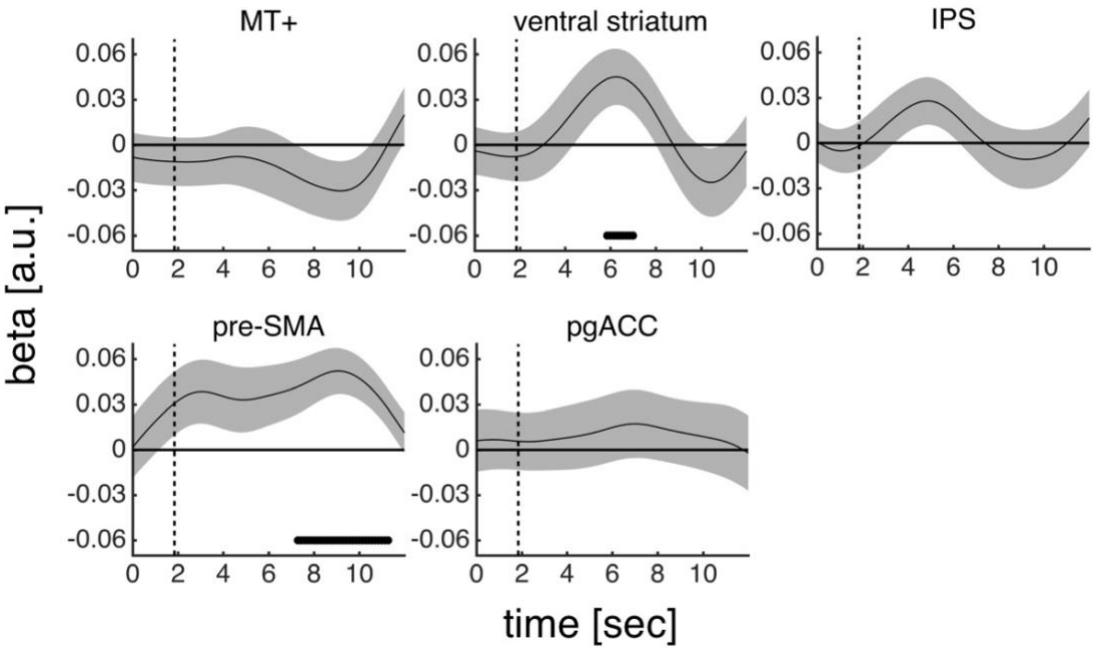


**Fig. S7.** Encoding of reward in ROI activity time courses. General linear model analysis of the effects of reward magnitude (high versus low) on ROI activity time courses. Dots below time course indicate significant excursion of *t* statistics assessed using two-tailed permutation tests. Vertical dashed line indicates the onset of the reward magnitude cue. Data, which only include confidence trials (15% of trials), are represented as group mean ± SEM.

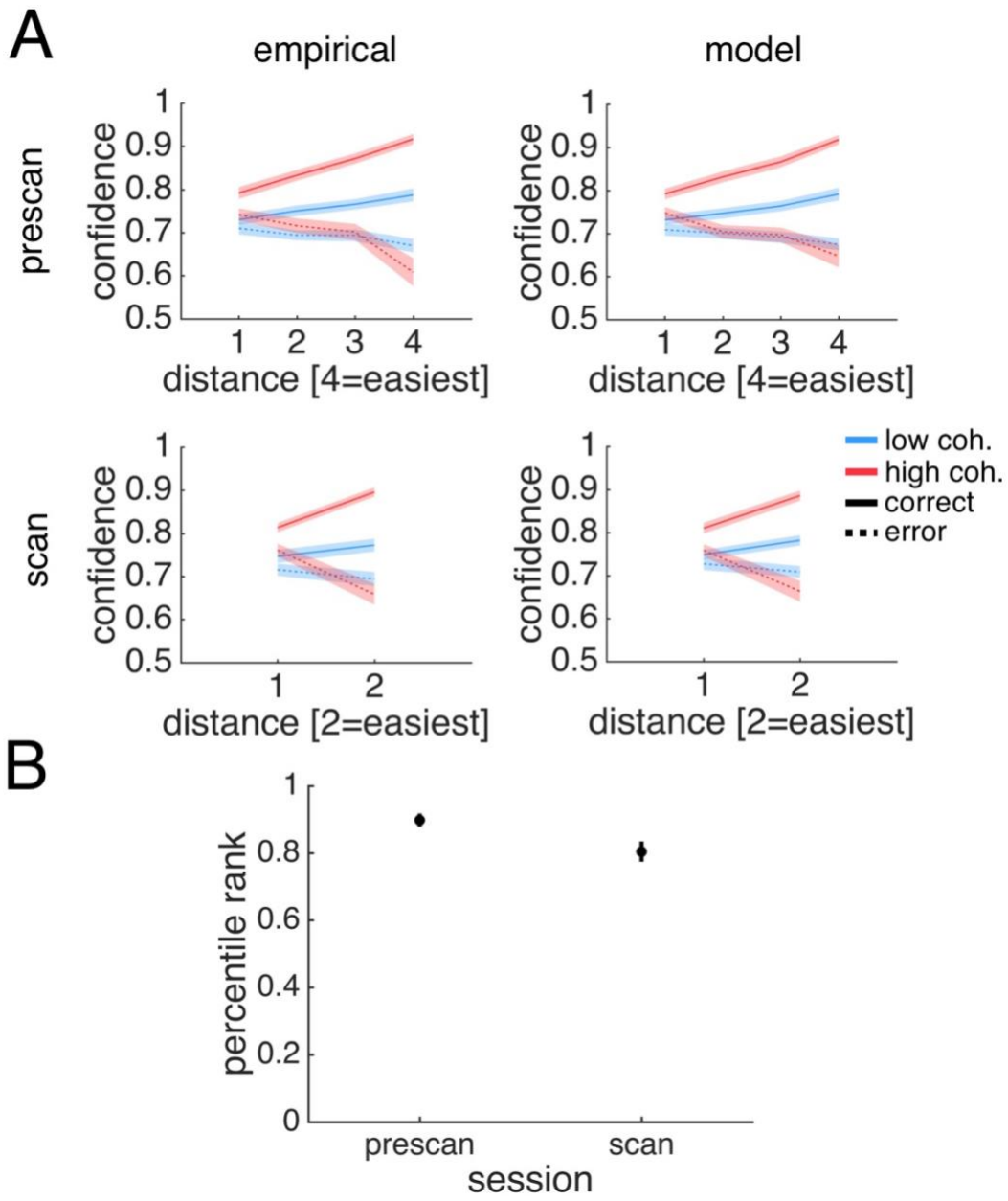


**Fig. S8.** Evaluation of model of subjective confidence. (*A*) Model qualitatively reproduces behaviour. The plots show (*left*) empirically observed and (*right*) model-derived confidence in the conditions of our factorial design split by choice accuracy (distance x coherence x choice accuracy) in (*top*) the prescan and (*bottom*) the scan sessions. The model was fitted to data from the prescan session and then used to create out-of-sample predictions about the scan session. We show empirical and predicted data from confidence trials only (15% of trials) for the scan session. (*B*) Model provided good fits and generalised between sessions. For each subject, $s = i$, we computed the likelihood of the data under a model, $m$, fitted to the subject's own prescan data, $m_{s=i}$, and under the models fitted to the prescan data of every other subject, $m_{s\neq i}$. The plot shows (y-axis) the percentile rank of the likelihood under a subject's own model, $m_{s=i}$, compared to all other models, $m_{s\neq i}$, for (x-axis) the prescan and the scan sessions. Here higher rank indicates better model fits. As described in Fig. 4*A*, the model predicts, for each trial, the probability that a given response is made. Model predictions were evaluated by summing the likelihood of the observed trial-by-trial responses under this probability distribution. In *A*-*B*, data are represented as group mean ± SEM.

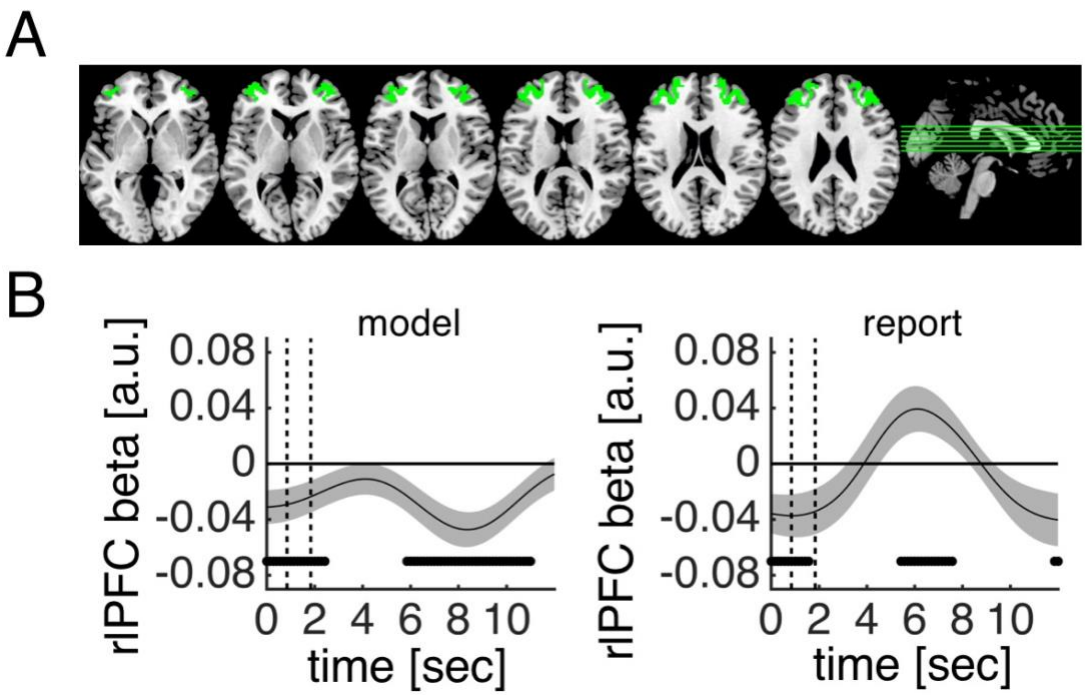


**Fig. S9.** Activity in rlPFC predicts confidence estimates. (*A*) ROI mask for rlPFC. (*B*) General linear model analysis of encoding of model-derived subjective confidence (all trials) and reported confidence (confidence trials) in rlPFC activity time courses. Dots below time course indicate significant excursion of *t* statistics assessed using two-tailed permutation tests. Vertical dashed lines indicate the onsets of the motion stimulus and the choice phase. Data are represented as group mean ± SEM.

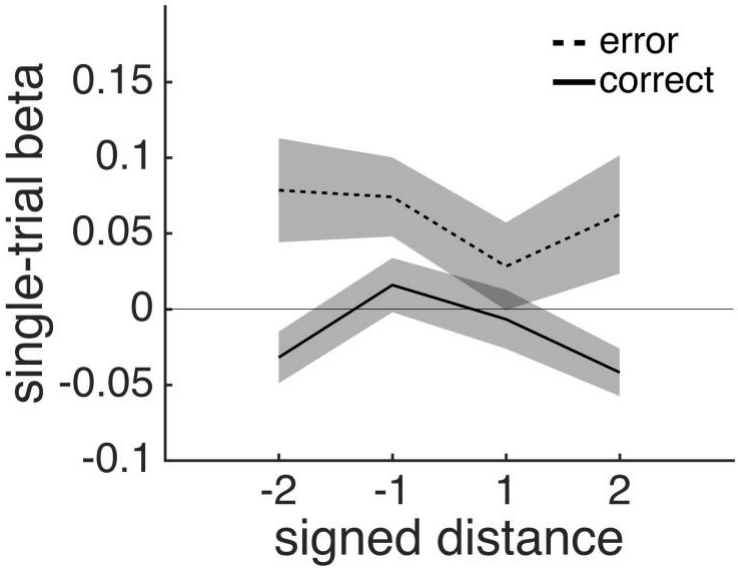


**Fig. S10.** Activity in pre-SMA as a function of boundary distance and choice accuracy. Single-trial pre-SMA activity estimates show an 'X'-pattern as a function of signed distance (negative: CCW; positive: CW) and choice accuracy (dashed: error; solid: correct): activity is, on correct trials, lower for larger distances but, on error trials, higher for larger distances. This 'X'-pattern is expected under a model which tracks the *inverse* of the absolute distance between sensory evidence and a choice boundary. Data are represented as group mean ± SEM.

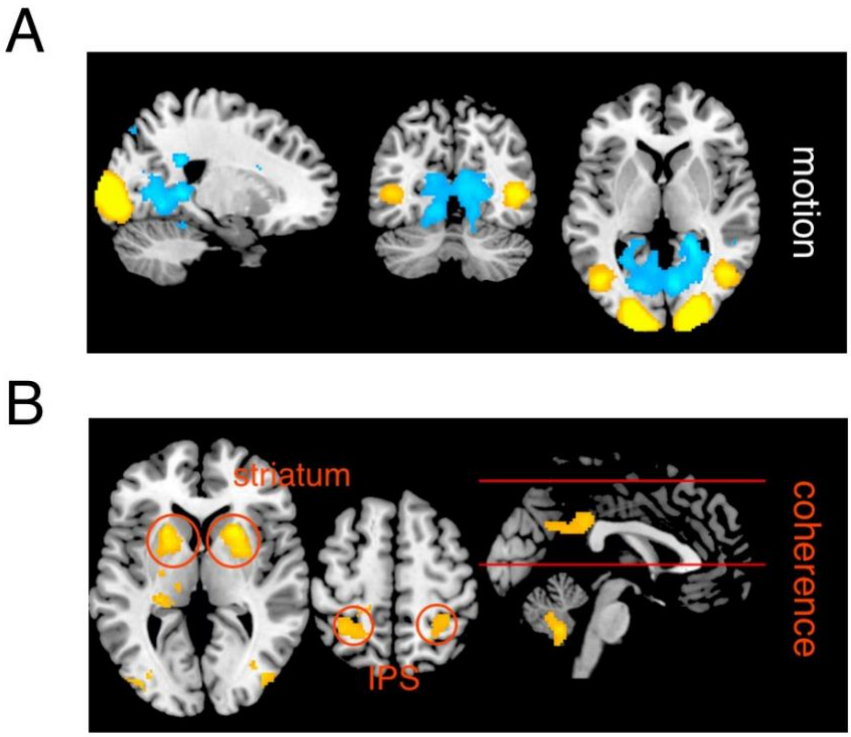


**Fig. S11.** Neural responses to coherently moving dots. (*A*) Whole-brain analysis of contrast between coherently moving and static dots in motion-localiser scan. (*B*) Whole-brain analysis of main effect of coherence masked by coherent-motion contrast. We applied an exclusive mask constructed from the contrast shown in *A* (thresholded at $P < 0.05$, uncorrected) to GLM1. In *A*-*B*, cluster colours denote positive (warm) and negative (cold) effects. Clusters are significant at $P < 0.05$, FWE-corrected for multiple comparisons; the cluster-defining threshold is $P < 0.001$, uncorrected. Images are shown at $P < 0.001$, uncorrected.

| contrast | label | voxels at $P < 0.001$ | peak $z$-score | $P$ (cluster FWE-corrected) | peak voxel MNI coordinates | | | |
|---|---|---|---|---|---|---|---|---|
| high > low coh<br>*exclusive mask: distance and interaction* | striatum | 1423 | 5.41 | <0.001 | 26 | 10 | -6 | R |
| | extrastriate (incl. MT+) | 2675 | 5.36 | <0.001 | 28 | -84 | 8 | R |
| | striatum | 1505 | 5.14 | <0.001 | -24 | 10 | -10 | L |
| | extrastriate (incl. MT+) | 1635 | 4.72 | <0.001 | -40 | -78 | 2 | L |
| | posterior parietal | 373 | 4.54 | 0.001 | 38 | -26 | 40 | R |
| | cerebellum | 428 | 4.42 | <0.001 | -4 | -56 | -42 | LR |
| | posterior cingulate | 336 | 4.39 | 0.001 | 12 | -38 | 30 | LR |
| | posterior parietal | 394 | 4.25 | 0.001 | -32 | -50 | 62 | L |
| high < low coh<br>*exclusive mask: distance and interaction* | ----------none---------- | | | | | | | |
| high > low dis<br>*exclusive mask: coherence and interaction* | posterior cingulate | 214 | 4.91 | 0.014 | -12 | -16 | 48 | LR |
| | early visual | 178 | 3.77 | 0.016 | -16 | -98 | 2 | L |
| high < low dis<br>*exclusive mask: coherence and interaction* | pre-SMA | 169 | 4.34 | 0.020 | 4 | 20 | 48 | LR |
| interaction +<br>*inclusive mask: coherence and distance* | pgACC | 932 | 4.82 | <0.001 | -2 | 44 | 10 | LR |
| interaction –<br>*inclusive mask: coherence and distance* | ----------none---------- | | | | | | | |

**Table S1.** Summary of significant activations for GLM1 post-masking. All reported activations are significant at $P < 0.05$, FWE-corrected for multiple comparisons; the cluster-defining threshold is $P < 0.001$, uncorrected. FWE: familywise error. MNI: Montreal Neurological Institute. L: left. R: right. pre-SMA: pre-supplementary motor area. pgACC: perigenual anterior cingulate cortex.

| contrast | label | voxels at $P < 0.001$ | peak $z$-score | $P$ (cluster FWE-corrected) | peak voxel MNI coordinates | | | |
|---|---|---|---|---|---|---|---|---|
| high > low coh | striatum | 1423 | 5.41 | <0.001 | 26 | 10 | -6 | R |
| | extrastriate (incl. MT+) | 2675 | 5.36 | <0.001 | 28 | -84 | 8 | R |
| | striatum | 1505 | 5.14 | <0.001 | -24 | 10 | -10 | L |
| | extrastriate (incl. MT+) | 1635 | 4.72 | <0.001 | -40 | -78 | 2 | L |
| | posterior parietal | 373 | 4.54 | 0.001 | 38 | -26 | 40 | R |
| | cerebellum | 428 | 4.42 | <0.001 | -4 | -56 | -42 | LR |
| | posterior cingulate | 453 | 4.39 | <0.001 | 12 | -38 | 30 | LR |
| | posterior parietal | 394 | 4.25 | 0.001 | -32 | -50 | 62 | L |
| high < low coh | ----------none---------- | | | | | | | |
| high > low dis | posterior cingulate | 214 | 4.91 | 0.006 | -12 | -16 | 48 | LR |
| | pgACC | 246 | 4.07 | 0.003 | 12 | 34 | 12 | LR |
| | early visual | 330 | 4.02 | 0.006 | 18 | -94 | -2 | R |
| | early visual | 343 | 3.86 | <0.001 | -18 | -90 | 2 | L |
| high < low dis | pre-SMA | 169 | 4.34 | 0.020 | 4 | 20 | 48 | LR |
| interaction + | pgACC | 1093 | 4.82 | <0.001 | -2 | 44 | 10 | LR |
| | temporal | 158 | 4.43 | 0.027 | -60 | -18 | -18 | L |
| interaction – | precuneus | 367 | 4.44 | <0.001 | 0 | -64 | 60 | LR |

**Table S2.** Summary of significant activations for GLM1 pre-masking. All reported activations are significant at $P < 0.05$, FWE-corrected for multiple comparisons; the cluster-defining threshold is $P < 0.001$, uncorrected. FWE: familywise error. MNI: Montreal Neurological Institute. L: left. R: right. pgACC: perigenual anterior cingulate cortex. pre-SMA: pre-supplementary motor area.

| contrast | label | voxels at $P < 0.001$ | peak $z$-score | $P$ (cluster FWE-corrected) | peak voxel MNI coordinates | | | |
|---|---|---|---|---|---|---|---|---|
| high > low coh<br>*exclusive mask: distance and interaction* | extrastriate (incl. MT+) | 791 | 4.89 | <0.001 | 44 | -62 | 4 | R |
| | striatum | 467 | 4.81 | <0.001 | -22 | 12 | -10 | L |
| | striatum | 610 | 4.77 | <0.001 | 20 | 16 | -4 | R |
| high < low coh<br>*exclusive mask: distance and interaction* | ----------none---------- | | | | | | | |
| high > low dis<br>*exclusive mask: coherence and interaction* | pgACC/vmPFC | 570 | 5.00 | 0.001 | 10 | 38 | 12 | LR |
| | insula | 159 | 4.57 | 0.023 | -30 | 20 | -18 | L |
| | insula | 140 | 4.39 | 0.040 | 32 | 14 | -10 | R |
| | posterior cingulate | 150 | 4.38 | 0.030 | 0 | -12 | 40 | LR |
| high < low dis<br>*exclusive mask: coherence and interaction* | ----------none---------- | | | | | | | |
| interaction +<br>*inclusive mask: coherence and distance* | pgACC | 252 | 4.28 | 0.002 | 4 | 44 | 12 | LR |
| interaction –<br>*inclusive mask: coherence and distance* | ----------none---------- | | | | | | | |

**Table S3.** Summary of significant activations for GLM1C (correct trials only) post-masking. All reported activations are significant at $P < 0.05$, FWE-corrected for multiple comparisons; the cluster-defining threshold is $P < 0.001$, uncorrected. FWE: familywise error. MNI: Montreal Neurological Institute. L: left. R: right. pgACC: perigenual anterior cingulate cortex. vmPFC: ventromedial prefrontal cortex.

| contrast | label | voxels at $P < 0.001$ | peak $z$-score | $P$ (cluster FWE corrected) | peak voxel MNI coordinates | | | |
|---|---|---|---|---|---|---|---|---|
| high > low coh | extrastriate (incl. MT+) | 2050 | 4.89 | <0.001 | 44 | -62 | 4 | R |
| | striatum | 618 | 4.81 | <0.001 | -22 | 12 | -10 | L |
| | striatum | 768 | 4.77 | <0.001 | 20 | 16 | -4 | R |
| | extrastriate (incl. MT+) | 1307 | 4.23 | <0.001 | -24 | -88 | 20 | L |
| | posterior parietal | 278 | 4.01 | 0.002 | -24 | -52 | 58 | L |
| | posterior parietal | 172 | 3.97 | 0.027 | 30 | -48 | 62 | R |
| high < low coh | ----------none---------- | | | | | | | |
| high > low dis | pgACC | 610 | 5.00 | <0.001 | 10 | 38 | 12 | LR |
| | insula | 229 | 4.57 | 0.004 | -30 | 20 | -18 | L |
| | insula | 182 | 4.41 | 0.012 | 32 | 14 | -10 | R |
| | posterior cingulate | 293 | 4.38 | <0.001 | 0 | -12 | 40 | LR |
| | premotor | 220 | 4.36 | 0.004 | -46 | -18 | 52 | L |
| high < low dis | ----------none---------- | | | | | | | |
| interaction + | pgACC | 309 | 4.28 | <0.001 | 4 | 44 | 12 | LR |
| interaction – | precuneus | 520 | 4.41 | <0.001 | 10 | -66 | 64 | L |

**Table S4.** Summary of significant activations for GLM1C (correct trials only) pre-masking. All reported activations are significant at $P < 0.05$, FWE-corrected for multiple comparisons; the cluster-defining threshold is $P < 0.001$, uncorrected. FWE: familywise error. MNI: Montreal Neurological Institute. L: left. R: right. pgACC: perigenual anterior cingulate cortex.