Relating metacognitive sensitivity to human brain structure: a combined psychophysics-MRI study

Stephen M. Fleming[1], Rimona Weil[1,2], Zoltan Nagy[1], Raymond J. Dolan[1] & Geraint Rees[1,2]

[1]Wellcome Trust Centre for Neuroimaging & [2]Institute of Cognitive Neuroscience, University College London

UCL

Introduction

- **Humans have the ability to generate thoughts about their own internal states, known as higher-order or metacognitive beliefs (1).**
- Metacognitive sensitivity may underpin everyday awareness and sense of self (2), and is crucial for the active guidance of future behaviour (3).
- Generation of metacognitive judgments can be formalised using computational descriptions of post-decision processes (4), but the biological foundations of this ability are unknown.
- **Is variability in metacognitive ability related to brain structure?**

Methods

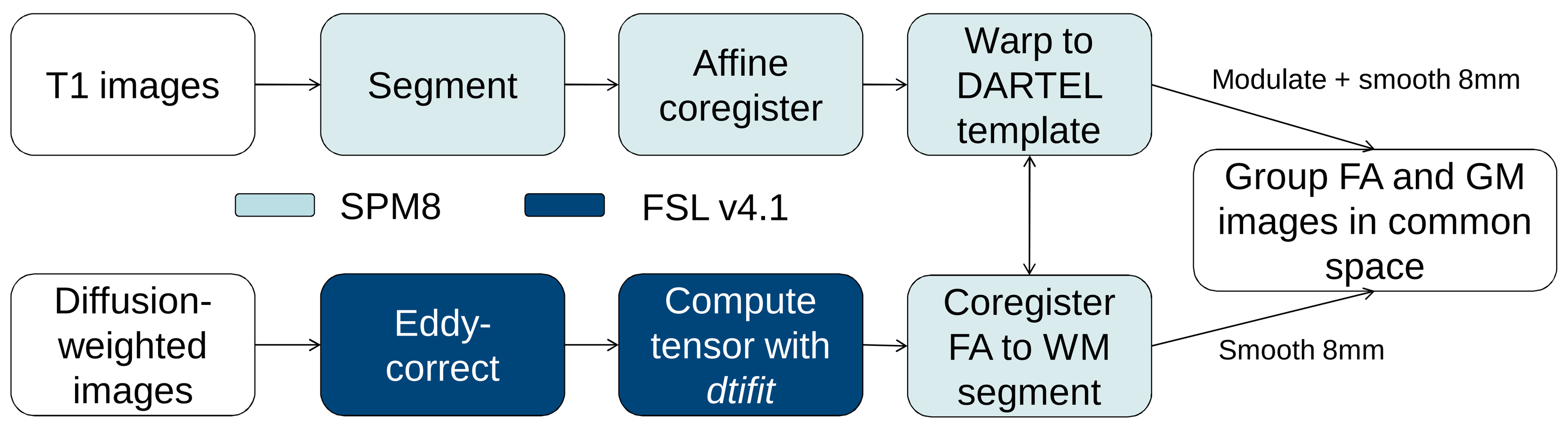
- ***Participants*** - 32 participants (15 males; aged 19 – 37 years; mean age 26.4 years). 1 participant was excluded due to aberrant task performance (d' > 3SD from the group mean).

- ***Psychophysics*** - temporal 2AFC task (see Task figure), 600 trials split into 6 blocks. Participants were required to decide whether the higher contrast "pop-out" Gabor had appeared in the first or the second interval; no feedback was given. Trial-by-trial ratings of decision confidence were made on a scale of 1-6.
    - Pop-out Gabor contrast was dynamically adjusted using a 1-up 2-down staircase procedure, leading to convergence on 71% accuracy.

- ***Behavioural analysis*** – Type II signal detection analysis was used to quantify metacognitive sensitivity. A "hit" is defined as a high confidence response after a correct decision and a "false alarm" as a high confidence response after an incorrect decision. Area under the ROC (A_{roc}) was quantified using non-parametric methods (5).

- ***Structural brain imaging*** - A 1.5T Sonata scanner (Siemens Medical Systems, Erlangen, Germany) was used to acquire all images for each participant. T1-weighted anatomical whole-brain scans were acquired for voxel-based morphometry (VBM) analysis (176 slices, echo time = 3.56ms, TR = 12.24ms, voxel size 1mm isotropic). The diffusion tensor imaging (DTI) dataset comprised of 68 images with 60 slices and 2.3 mm isotropic resolution (b = 1000 s/mm^2, echo time = 90ms, FOV = 220ms).

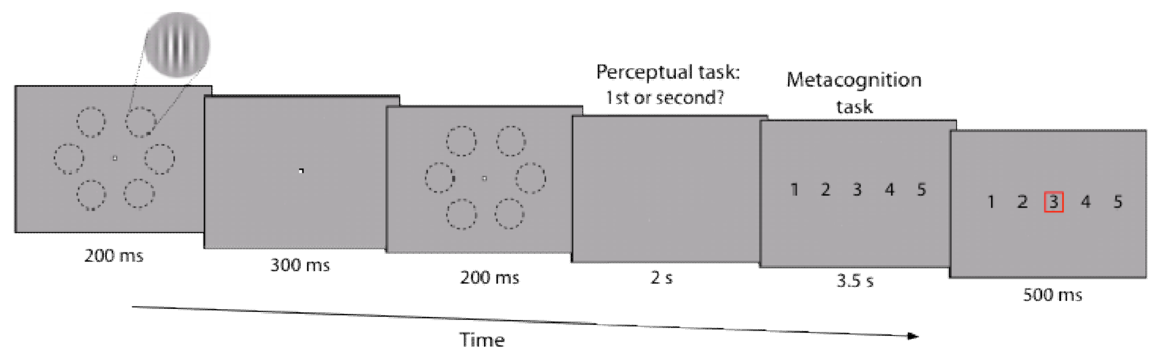
- ***Preprocessing pipeline:***



- ***Regression analysis*** - A_{roc}, d', Type II criterion (B_{roc}), the unsigned value of the Type I criterion ($|c|$) and gender (M = 1; F = 0) were included in the model. Cluster-based statistics were used to locate significant regions after applying an initial cluster-defining threshold of $P < 0.001$. Nonstationary correction of expected cluster size was applied using the *NS* toolbox (http://www.fmri.wfubmc.edu/cms/NS-General), which for smooth data and high d.f. leads to adequate control over family-wise false positives ($P < 0.05$) (6).

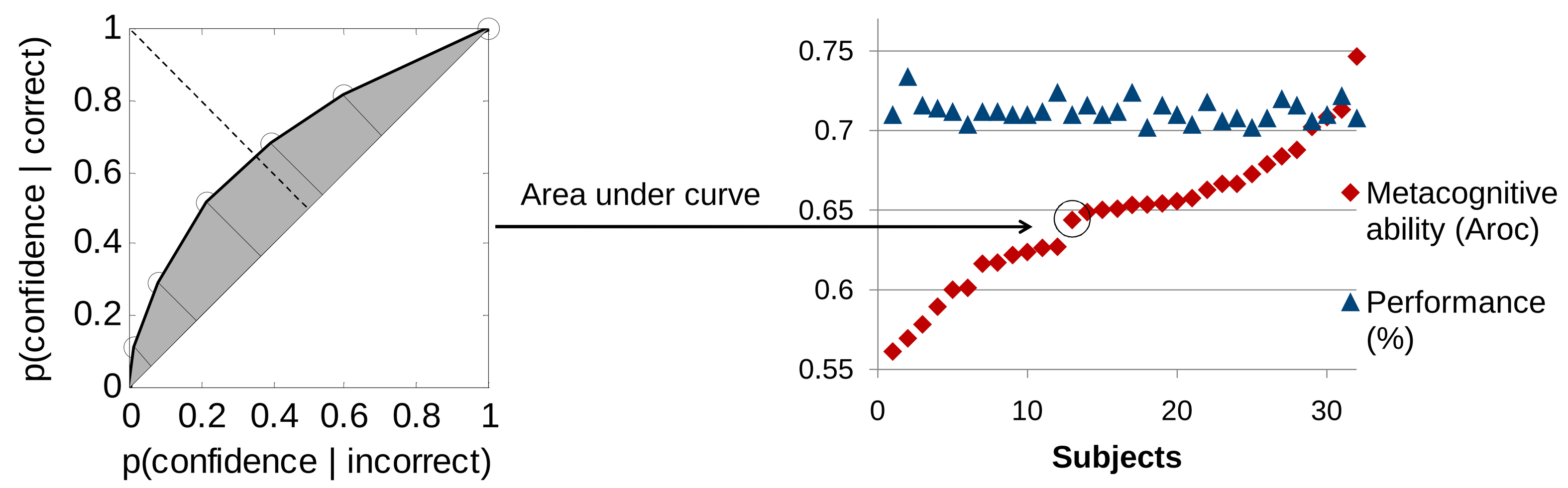
Behaviour

Task



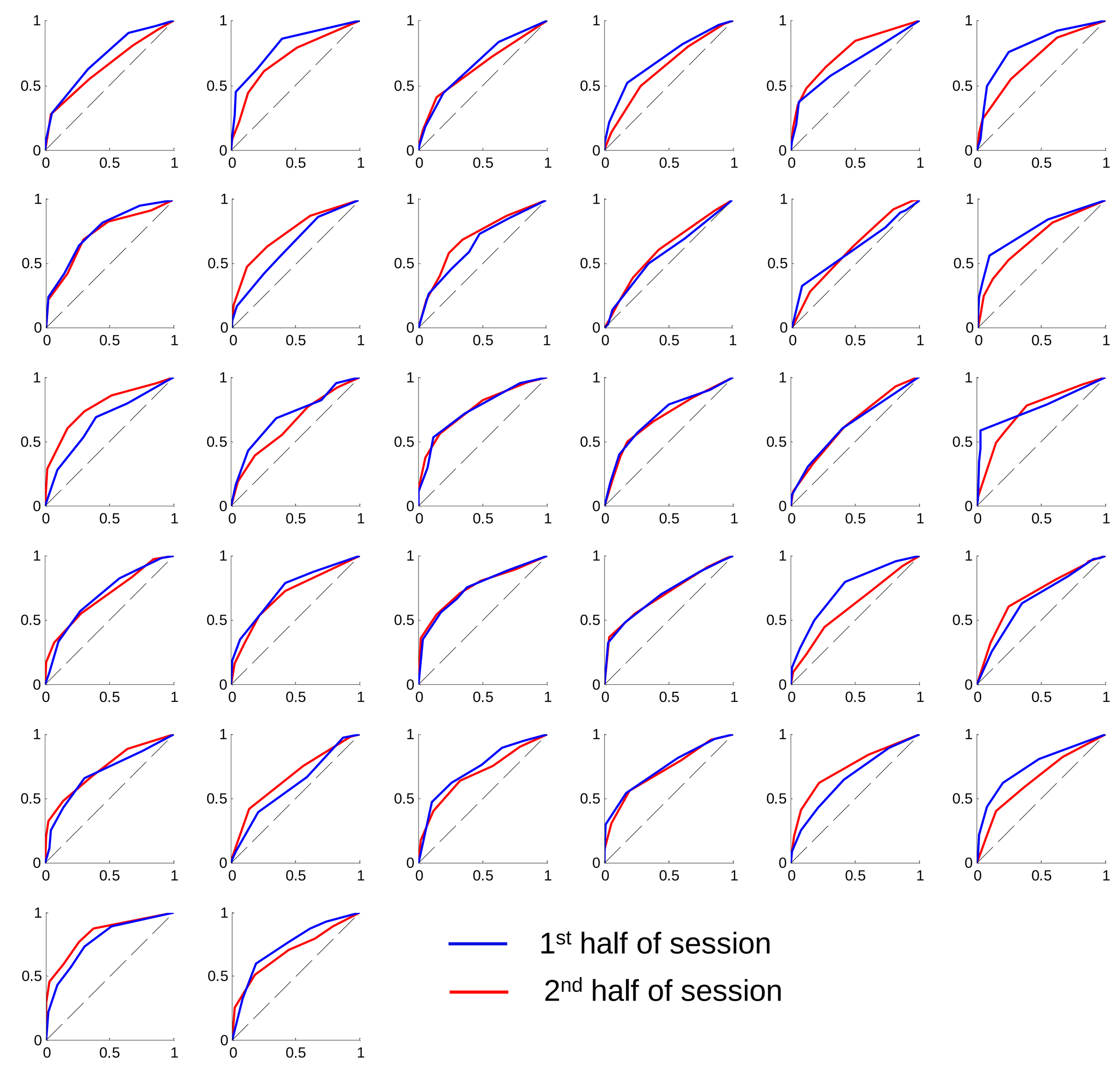
Temporal 2AFC + dynamic staircase; trial-by-trial ratings of decision confidence

"Type II" ROC analysis



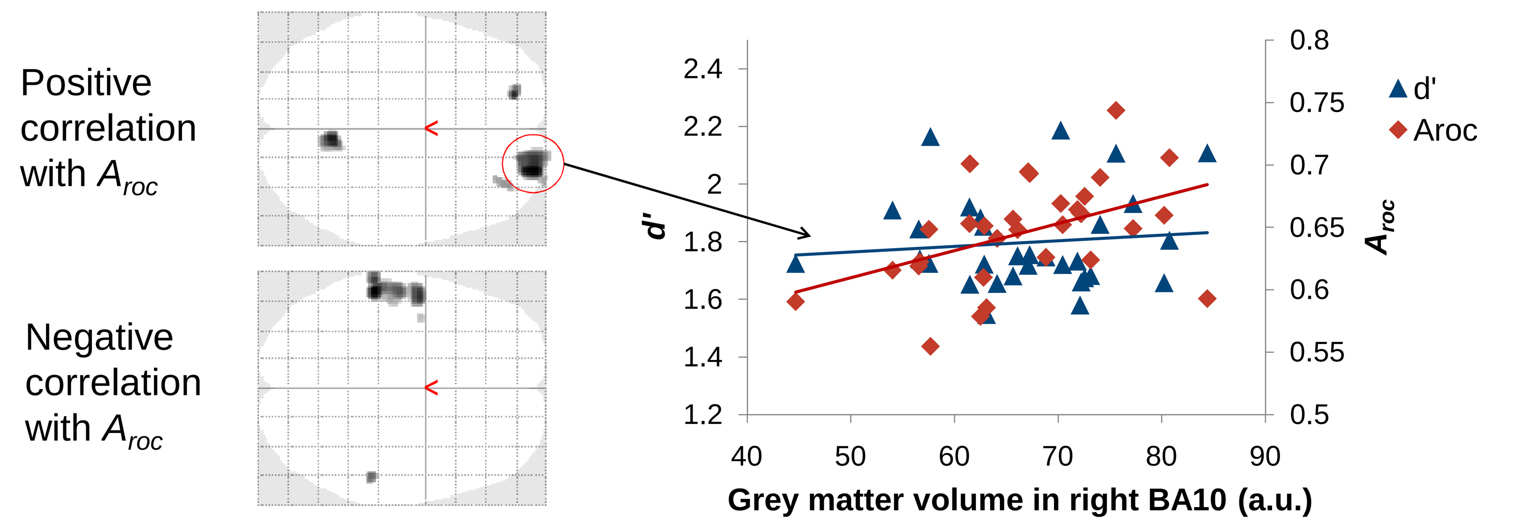
Area under ROC (A_{roc}) reflects an individual's ability to link subjective confidence to changes in objective performance

Individual Type II ROCs

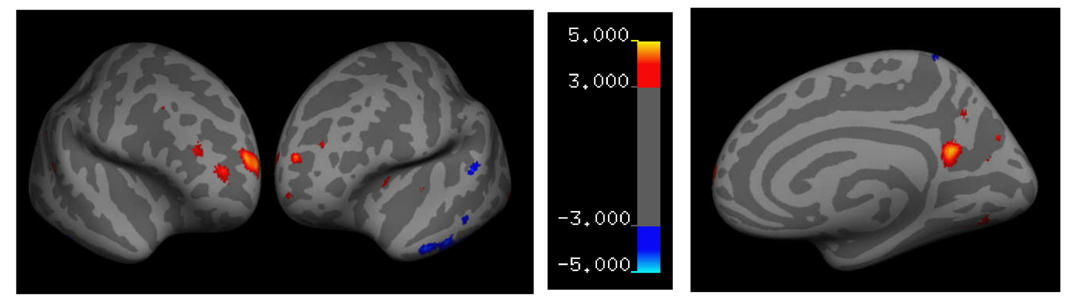


Intra-subject consistency in A_{roc}, $r = 0.51$, $P = 0.003$

Grey matter correlating with A_{roc}

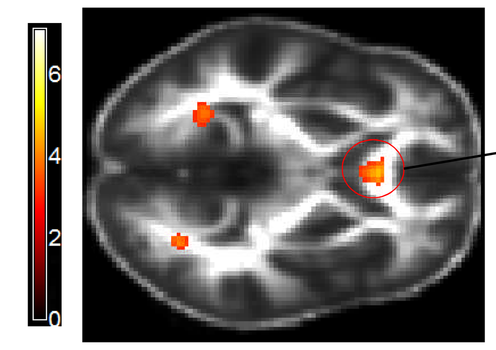


Images displayed at $P < 0.001$, uncorrected. R-BA10 and L-ITG, $P < 0.05$, whole-brain cluster-FWE corrected

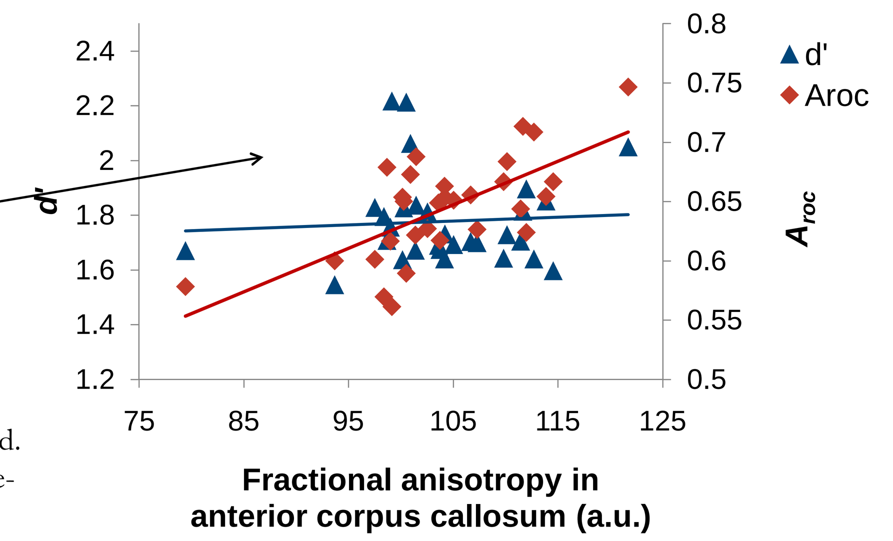


White matter (FA) correlating with A_{roc}

Positive correlation with A_{roc}



Images displayed at $P < 0.001$, uncorrected. Anterior corpus callosum, $P < 0.05$, whole-brain cluster-FWE corrected



Conclusions

- Individual differences exist in metacognitive sensitivity (A_{roc}) despite holding perceptual performance (d') constant.
- Both grey matter volume and white matter integrity (fractional anisotropy) in prefrontal cortex (BA10 and anterior corpus callosum) are predictive of an individual's metacognitive ability.
- The functional contributions of these regions to post-decision judgments remains to be examined.
- **Our findings indicate a circumscribed network including prefrontal and inferior temporal cortex may contribute to metacognitive computations in healthy individuals.**

(1) J. Metcalfe, *Metacognition: Knowing About Knowing.* (MIT Press: 1996).
(2) H.C. Lau, A higher order Bayesian decision theory of consciousness. *Prog. Brain Res.* **168**, 35-48 (2008).
(3) J. Metcalfe & B. Finn, Evidence that judgments of learning are causally related to study choice. *Psychon. Bull. Rev.* **15**, 174-179 (2008).
(4) S.M. Fleming, R.J. Dolan, Effects of loss aversion on post-decision wagering: Implications for measures of awareness. *Conscious. Cogn.* **19**, 352-363 (2010).
(5) D.E. Kornbrot, Signal detection theory, the approach of choice: model-based and distribution-free measures and evaluation. *Percept. Psychophys.* **68**, 393-414 (2006).
(6) S. Hayasaka, K.L. Phan, I. Liberzon, K.J. Worsley, T.E. Nichols, Nonstationary cluster-size inference with random field and permutation methods. *NeuroImage* **22**, 676-687 (2004).

This work was supported by MRC funding within the UCL 4-Year PhD in Neuroscience (SF), an MRC Clinical Research Fellowship (RW) and the Wellcome Trust (GR, RJD, ZN).

For further information please contact: s.fleming@fil.ion.ucl.ac.uk