



Neural basis of non-conscious visual working memory



Agneish Dutta^a, Kushal Shah^a, Juha Silvanto^b, David Soto^{a,*}

^a Imperial College London, Department of Medicine, Division of Brain Sciences, London, UK

^b Department of Psychology, Faculty of Science and Technology, University of Westminster, 309 Regent Street, London W1B 2HW, UK

ARTICLE INFO

Article history:

Accepted 11 January 2014

Available online 17 January 2014

Keywords:

Consciousness
Working memory
Vision
Prefrontal

ABSTRACT

Recent research indicates that human observers can perform high-level cognitive tasks typically associated with working memory processes (e.g. learning of complex item sequences, reading, arithmetic or delayed visual discrimination) independently of conscious awareness of the relevant information. However, the neural basis of this phenomenon is not known. Here we show neuroimaging and neurostimulation evidence that the dorsolateral and anterior prefrontal cortex can operate on non-conscious information in a manner that goes beyond automatic forms of sensorimotor priming and which may support implicit working memory processes and higher-level cognitive function.

© 2014 Elsevier Inc. All rights reserved.

Introduction

Working memory (WM) processes, for example, the maintenance of information for later use in decision-making, are classically thought to operate on the contents of conscious awareness (Baars and Franklin, 2003; Baddeley, 2003). Recent research has however provided intriguing evidence that reading, doing arithmetic, complex visuo-spatial learning and working memory operations may occur independently of conscious awareness of the critical information (Rosenthal et al., 2010; Sklar et al., 2012; Soto et al., 2011) suggesting that, at least under certain conditions, non-conscious information can be committed to working memory systems. These recent investigations beg the question of how the brain can undergo computations using non-conscious information in the service of high-level cognition. Here we investigate these mechanisms by combining functional MRI and neurostimulation techniques.

Addressing the neural substrates that mediate working memory processes outside awareness is relevant from theoretical and also clinical neuroscience perspectives; for instance, ‘blindsight’ patients following damage to the primary visual cortices can display astonishingly high levels of non-conscious perception in forced-choice paradigms (i.e. high discrimination accuracy of stimuli that go undetected), but they may be unable to use non-conscious cues following a delay period to influence behaviour — as if the influence of non-conscious cues were short-lived and only affected behaviour in “real-time” (Striemer et al., 2009).

The behavioural paradigm used here followed a recent study (Soto et al., 2011) showing that human observers can encode the orientation of a Gabor cue, which could be rendered non-conscious by masking, and keep it online even in the presence of intervening distracters to perform significantly above chance in a delayed cue-target discrimination test. Further evidence for this non-conscious form of WM was found when 2 masked items had to be retained for the delayed test. In the present neuroimaging study we elected to use the simpler version of this paradigm with a delayed cue-target discrimination to delineate for the first time the neural basis of this phenomenon. The behavioural task is depicted in Fig. 1 (see also Methods for further details). Observers were presented with a brief oriented grating (the cue) which was masked and followed by a delay period before the onset of the memory orientation test, namely, to respond whether the test grating was tilted clock- or anticlockwise relative to the cue. Observers were encouraged to attend and hold the cue grating in their memory and try their best in the memory test even if they could not consciously perceive it. Following the discrimination judgment, observers rated their awareness of the initial cue on a 1–3 scale: 1 — “Did not see the cue at all”; 2 — “Saw a brief glimpse but not the orientation”; and 3 — “Saw the cue orientation”. Awareness ratings are critical to assess conscious experience (Overgaard et al., 2010).

This experimental paradigm can therefore be employed to assess whether information that goes undetected can nevertheless permeate into working memory systems that may enable maintenance and later use of the non-conscious information in subsequent tasks (i.e. cue-target delayed discrimination responses).

One hypothesis is that non-conscious working memory-like computations may be driven by primitive perceptual memory systems tuned to specific features of the non-conscious cue, for example, orientation channels in early visual cortex (Magnussen, 2000; Pasternak and

* Corresponding author at: Imperial College London, Division of Brain Sciences, Charing Cross Campus, St. Dunstan’s Road, London W6 8RP, UK.
E-mail address: d.soto@imperial.ac.uk (D. Soto).

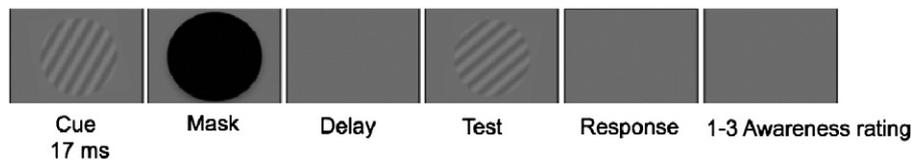


Fig. 1. Illustration of the behavioural paradigm. The cue could be tilted 10, 40, 70, 100, 130 or 160° from the vertical. The test grating could be tilted 30° either clock- or anticlockwise relative to the initial grating. Participants were asked to indicate the direction of the tilt and provide a 1–3 rating of their awareness of the memory cue. There were also catch trials on which no memory cue was presented (see also *Methods* for the fMRI and tDCS experiments).

Greenlee, 2005) which may provide input for delayed perceptual decision-making. Alternatively, there may be hardwired operations of prefrontal (PFC) regions typically associated with the maintenance/manipulation of relevant information in the focus of the mind's eye (Curtis and D'Esposito, 2003) and perceptual decision-making (Heekeren et al., 2004). According to this hypothesis the operation of PFC mechanisms in WM tasks may occur to some extent outside of conscious awareness.

Methods

Experiment 1: Functional MRI

Participants

24 healthy adult observers with normal or corrected to normal vision (14 females) participated in return for £30 for their time. They were aged between 20 and 24 (mean age: 22). Participants had no reported neurological or psychiatric disorders and were naive regarding the experimental aims and hypotheses. The study was approved by the West London Ethics committee and all experiments were undertaken with the understanding and prior written consent of each participant.

Apparatus

MRI scanning was performed in a Siemens Avanto 1.5 T MRI scanner and a 32-channel head coil. An out-bore screen was mounted at the end of the scanner bore and visual stimuli were displayed by a projector in the adjacent room to the scanner. A mirror was placed above the headrest at 45° enabling participants to see the screen looking straight upwards.

Imaging acquisition and scanning parameters

A T2 weighted echo planar imaging sequence was used to obtain 29 contiguous slices covering the whole brain. 287 volumes were taken with a field of view of 205 × 205 mm, TR of 2500 ms, TE of 44 ms and slice thickness of 3.2 mm. We used the prospective acquisition correction algorithm (Thesen et al., 2000) to attenuate motion-induced effects on magnetization history by adjusting slice position and orientation prospectively. We also collected a six minute T1-weighted structural scan.

Experimental task and procedure

The task was programmed and presented with E-Prime v2.0 (Psychology Software Tools Inc., Pittsburgh, USA; www.pstnet.com/epime.cfm). Each trial began with a fixation point appearing in the centre of the screen for 500 ms followed by a blank screen showing for 500 ms. The memory cue appeared next for 16.67 ms. This memory cue was a grating with 0.1 Michelson contrast (97 cd/m²) presented on a grey background (93 cd/m²). There were 6 gratings varying in orientation in steps of 30°. Cue orientation could be 10, 40, 70, 100, 130, or 160° from the vertical (note for example that when the bottom-right quadrant of a grating is tilted 160° from vertical, its top-left counterpart is tilted 340°, which is only 30° apart from the top-left quadrant of the grating tilted 10° from the vertical). Grating orientation was randomly selected on each trial. The cue was followed by a blank screen of 16.67 ms and a black mask (4 cd/m²) appearing for 100 ms. Another

blank screen followed for 1400 ms before the test grating appeared. The test could be rotated 30° either clock- or counter-clockwise relative to the cue with the same probability. Participants were instructed about this contingency and encouraged to try their best to discriminate between the two alternatives even if they could not consciously perceive the cue. The response period followed, where participants were asked to indicate whether the test grating was rotated clock- or anticlockwise relative to the initial cue by pressing a button as accurately as they could within an interval of 1500 ms. Following this they were asked to rate how aware they were of the presence of the memory cue during a time window of 1500 ms, by pressing a button using the following scale: 1 – “Did not see the cue at all”; 2 – “Saw a brief glimpse but not the orientation”; 3 – “Saw the cue orientation”. We did not encourage participants to distribute their responses evenly across the three possible rating categories but rather to use the ratings that best reflected their conscious experience. A jittered inter-trial interval between 2.5 and 4.5 s followed. This inter-trial interval had an exponential distribution to facilitate the estimation of the BOLD response i.e. 50% 2.5 s, 25% 3 s, 12% 3.5 s, 6% 4 s, and 6% 4.5 s. There were 90 trials per participant. Sixty trials presented a test grating tilted either clock- or counter-clockwise relative to the initial cue. There were also 30 catch trials without a cue, enabling the calculation of cue perceptual sensitivity on 1-rating trials (see below).

The total imaging session per participant was around 30 min (including some training trials inside the scanner, 6 min for the structural scan and 13 min for the experimental block of 90 trials). The relatively low number of trials followed from our prior study (Soto et al., 2011) which also comprised a similar amount in some of the experiments, the rationale being that it may often be difficult for participants to sustain the motivation and attention for long periods in this type of tasks with sub- and near-threshold stimuli on which there is little insight on the correctness of responses and no performance feedback.

Prior to the scanning session, observers practiced the task. They received instructions and were shown examples of the trials on the computer screen. Then they carried out two blocks of 24 trials. In one half of these trials the duration of the memory cue was 16.67 ms and in the other half it was 216.67 ms, randomly selected. The practice trials used the same cue-target delay of 1400 ms (plus a 100 ms mask) as in the actual fMRI experiment. Observers were encouraged to try their best in the task even if they could not consciously perceive the cues.

fMRI analysis

fMRI data analysis was performed using FEAT (fMRI Expert Analysis Tool) Version 6, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The first 4 volumes of the EPI scan were removed from the dataset, leaving 283 volumes. The following image pre-processing was applied: non-brain removal using BET (Smith, 2002); motion correction using MCFLIRT (Jenkinson et al., 2002). Translational movement parameters never exceeded 1 voxel in any direction for any participant (absolute mean 0.185 mm). Spatial smoothing used a Gaussian kernel of FWHM 6.0 mm and highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 50.0 s).

Statistical analyses were performed by modelling explanatory variables (EVs) within the context of the general linear model on a

voxel-by-voxel basis. The trials were modelled from the onset of the cue and the following regressors were included: correct responses on '1', '2' and '3' rating trials separately; errors (i.e. incorrect response with cue present on '1', '2' and '3' rating trials); onset of the confidence rating period; and catch trials. Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Jenkinson et al., 2002). Subsequently, we performed group-level analyses using FLAME 1 + 2 (FMRIB's Local Analysis of Mixed Effects) as implemented in FEAT. Each individual's EPI scans were registered to high-resolution structural images using FLIRT Boundary-Based Registration (Greve and Fischl, 2009), and were then co-registered and transformed to standard (Montreal Neurological Institute) space. We report maps of BOLD responses thresholded using clusters determined by a voxelwise Z threshold of 2.3 and a cluster significance threshold of $P = 0.05$, corrected for multiple comparisons across the whole brain (Worsley, 2001).

Covariate analyses

These analyses tested for regions that correlated with inter-individual differences in memory performance. The covariate effect was computed using a one-sample *t* test in a Mixed Effects analyses design which modelled the group mean and the additional effect of the covariate, which was entered in the model as a regressor orthogonalised with regard to the group mean regressor. Covariate analyses also included a further explanatory variable with individual cue sensitivity (d').

Two participants could not be entered in the fMRI analyses because they did not report unaware trials. The number of participants included on each of the fMRI contrasts varied depending on whether a given participant had trials in all the relevant cells for the contrast in question. For example, covariate analyses of memory performance on unaware trials included 22 participants. For the covariate analyses of memory performance in the context of the linear contrasts analyses across the awareness ratings, 7 participants had to be excluded because they lacked either 2-rating or 3-rating trials. To avoid floor effects in the analyses of correct > error delayed memory responses on unaware trials (without the covariate effect), we only included the participants that scored above 50% correct on unaware trials. Remarkably, the different analyses performed led to a consistent pattern of results despite leaving different subjects in and out across the different analysis contrasts.

Signal detection analyses

These were performed on trials where observers reported being unaware of the cue (i.e. 1-rating trials in the awareness scale). For these analyses, the 'signal' was defined as the absence of the cue and the 'noise' as the presence of the cue. Thus, "1" rating responses when the cue was absent were labelled as hits and the same responses when the cue was present were labelled as false alarms. We used the individual probabilities of hits and false alarms to compute d' , a sensitivity measure of performance based on signal detection theory (Wickens, 2002).

We used this approach to get a measure of the observer's sensitivity to the presence/absence of the cue on those trials rated as unaware. Since 2 and 3 visibility ratings are associated with awareness of the cue, these trials are non-informative for the purposes of our study, namely, to understand non-conscious processes in the service of working memory. Hence we did not use the 2 and 3 rating trials to compute a measure of cue detection sensitivity.

We assessed delayed memory performance by using the proportion of correct discriminations. We note that the area under the ROC curve equals the proportion correct obtained by an unbiased observer in a two-alternative forced choice task (Green, 1964). Hence, proportion correct works well as a measure of discrimination sensitivity. It should be noted that even if the participants had a response bias for one of the orientation response options (e.g., left-tilted over right-tilted) this would be averaged out in the calculation of proportion correct.

Experiment 2: Transcranial direct current stimulation (tDCS) study

Participants

These were 16 healthy volunteers (8 females), ranging from ages 21 to 22. None of them had any history of neurological or psychological disorders and all had normal or corrected to normal vision. All participants were right-handed. None of them had participated in the fMRI study. They were naive to the specific hypothesis and had not previously participated in a transcranial direct current stimulation study. All participants gave informed consent for a protocol approved by the West London ethics committee, which conformed to the declaration of Helsinki. One participant was excluded from the analyses because he/she did not attempt to discriminate the cue-target orientation difference (i.e. consistently responding "left" on unaware trials). The pattern of results however remained the same with all subjects entered in the analyses.

Apparatus, experimental task and procedure

These were similar to in Experiment 1, except that stimuli were presented on a Dell 17" cathode ray tube monitor. The luminance of the gratings was 5.67 cd/m² and their background: 5.52 cd/m². The luminance of the mask was 1.34 cd/m². The experiment took place in an illuminated room. Pilot testing indicated that masking of the cue under the above conditions could be achieved without the 16.67 ms blank screen between cue offset and mask onset employed in Experiment 1. Hence this blank screen was not included, similar to our prior study (Soto et al., 2011). There were 144 trials (72 were catch trials).

tDCS protocol

tDCS was applied to the scalp by a pair of 4 cm × 4 cm rubber electrodes, which were housed in 5 × 5 cm sponges. The current was delivered by a battery-driven stimulator (Magstim). The left PFC was stimulated by either an anode or a cathode electrode depending on experimental condition which was placed over F3 according to the 10–20 international system for electroencephalogram electrode placement. This method has been confirmed as an accurate method of dorso-lateral PFC localisation by neuronavigation techniques (Herwig et al., 2003) and this montage is known to influence the excitability of DLPFC and aPFC regions (Keeser et al., 2011). The reference electrode was placed on the left shoulder. The electrodes were held in place by rubber bands that were placed in coronal (under the chin) and transverse (around the head) orientations. Current was applied at 1.5 mA, with current density of 60 μA/cm². The stimulation was applied for 15 min while participants were at rest, prior to the task. Task performance was assessed following the 15 min of tDCS. Each of the participants tested received either anodal or cathodal tDCS in separate sessions, in a counterbalanced order. The sessions were separated by at least one day.

Table 1

Mean proportion of awareness responses on cue present trials. The values in brackets indicate SEM. (A) Experiment 1 (B) tDCS.

(A)			
fMRI study	Response 1	Response 2	Response 3
	0.341 (0.063)	0.377 (0.045)	0.282 (0.058)
(B)			
tDCS study	Response 1	Response 2	Response 3
Anodal	0.619 (0.076)	0.279 (0.048)	0.1 (0.041)
Cathodal	0.59 (0.086)	0.215 (0.041)	0.193 (0.078)

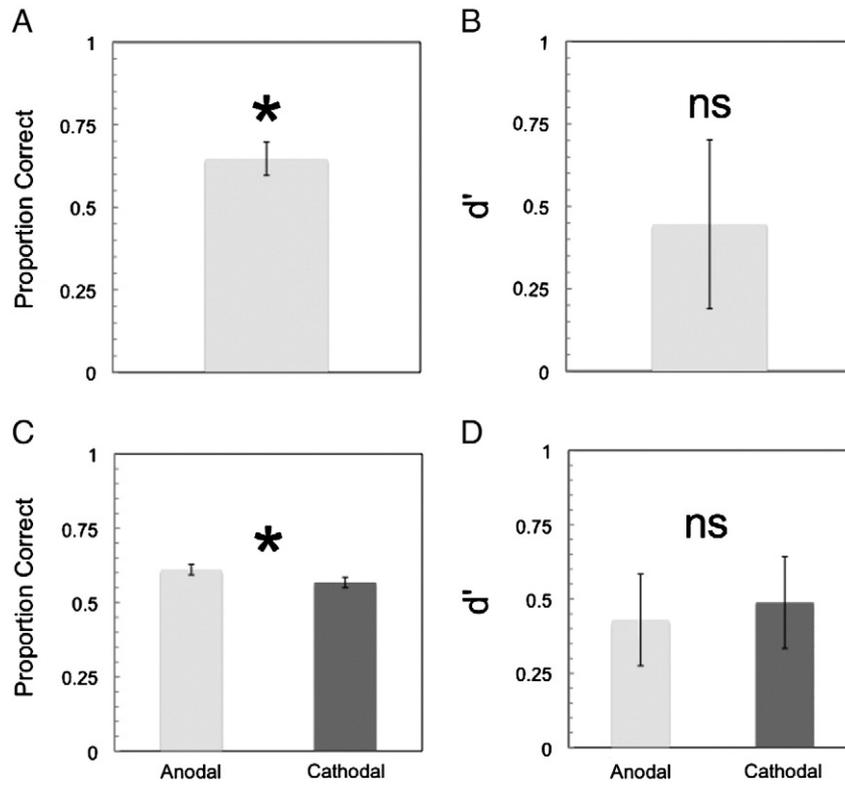


Fig. 2. Behavioural results. (A) Proportion of correct discrimination responses on unaware trials in the delayed memory test of the fMRI study. (B) Perceptual sensitivity of the cue on unaware trials in the fMRI study. (C) Proportion of correct discrimination responses on unaware trials in the delayed memory test across the anodal and cathodal tDCS conditions. (D) Cue detection sensitivity (d') on unaware trials across tDCS conditions.

Results

Experiment 1: Functional MRI

The proportion of awareness responses on cue present trials is given in Table 1A. The proportion of correct responses in the delayed memory orientation test on unaware (1-rating) trials was 0.65 correct (catch trials were not included in this computation), which was significantly above 0.5 chance level ($N = 22$; $t(21) = 2.66$, $P = 0.014$, two-tailed, Fig. 2A) indicating successful non-conscious delayed memory performance (Soto et al., 2011). Table 2A illustrates memory performance across the different awareness ratings.

Additional psychophysical analyses showed that detection sensitivity of the cue on unaware (1-rating) trials was poor (Fig. 2B; mean $d' = 0.44$, s.e.m = 0.256; see Table 3 for an indication of the relevant performance scores in the SDT analyses) and did not significantly differ from chance ($t(21) = 1.74$, $P = 0.096$). Cue detection sensitivity and memory performance on unaware trials did not correlate (Pearson correlation = 0.08, $P = 0.72$) so participants with higher cue detection sensitivity did not exhibit higher memory performance on unaware trials. Further,

Table 2

Proportion correct of memory orientation performance across the different awareness ratings. (A) fMRI study (B) tDCS study. Note that some participants reported little or no 3-rating trials and hence the means are taken from the data available.

(A)			
fMRI study	Response 1	Response 2	Response 3
	0.63	0.80	0.94
(B)			
tDCS study	Response 1	Response 2	Response 3
Anodal	0.62	0.71	0.91
Cathodal	0.57	0.74	0.95

regression analyses with d' as predictor and memory orientation accuracy as the predicted variable showed a significant intercept in the regression equation (0.642; $P < 0.0001$); hence memory performance remained above chance even when detection d' for the cue presence/absence is extrapolated to zero (for a justification of this method see Greenwald et al., 1996; Hannula et al., 2005). These results indicate that memory accuracy was dissociated from cue detection sensitivity.

Detection tasks may be associated with unequal variance of signal and noise distributions, hence limiting the applicability of signal detection measures that assume equal variance (i.e. d' computed based on a single pair of Hits and FA). To address this issue we calculated A' , which has been shown to be more robust than d' when the variance of signal and noise distributions is not equal (Donaldson, 1993). Replicating the results found with d' we showed that A' did not significantly differ from chance level (mean $A' = 0.59$, $t(21) = 1.59$; $P = 0.12$, one-sample t -test against chance $A' = 0.5$).

Neuroimaging results are reported at a voxelwise $Z = 2.3$, $P < 0.05$ whole-brain corrected, unless otherwise noted. First, a network involving cingulate regions, occipital and bilateral superior parietal lobe showed response increases with higher-level of awareness (i.e. 1-rating < 3-rating awareness trials; Fig. 3A, in green). Further, the contrast between correct vs. incorrect responses on trials displaying awareness of the cue ('2' and '3' ratings) showed activation in the right superior parietal lobe (Fig. 3A, in orange).

Table 3

Experiment 1 and 2. Mean P(Hits), P(False Alarms), and d' and A' scores for the presentation of the cue on unaware '1' rating trials.

	Experiment 1 fMRI	Experiment 2 Anodal	Experiment 2 Cathodal
Mean P(Hits)	0.57	0.575	0.584
Mean P(FA)	0.42	0.424	0.415
d'	0.44	0.406	0.482
A'	0.59	0.608	0.558

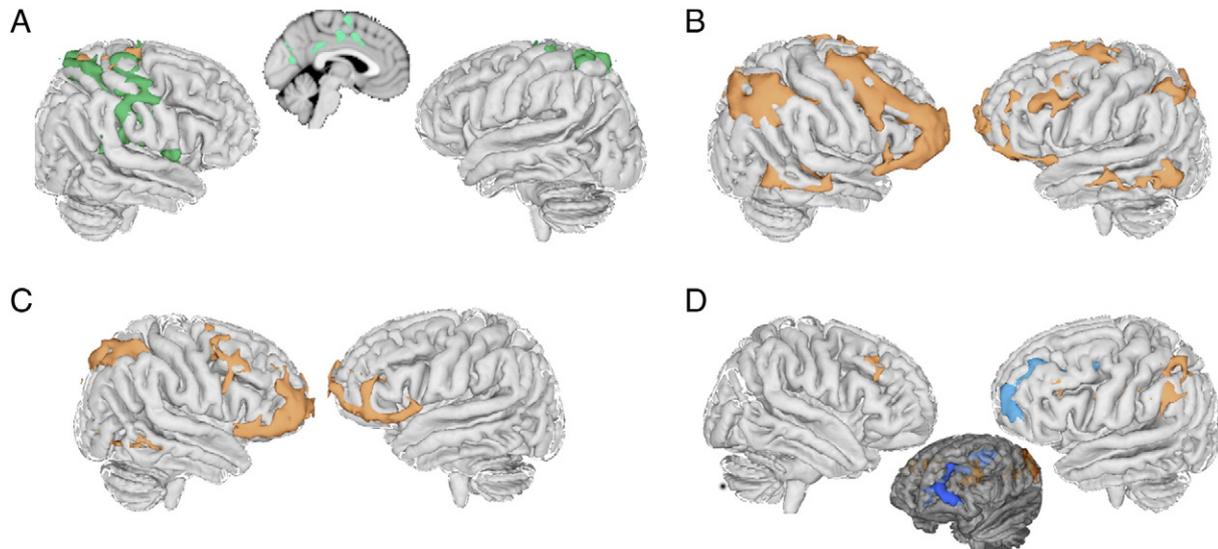


Fig. 3. fMRI results. (A) Green clusters showing increased activity with increased awareness including the bilateral precentral gyrus (MNI Coordinates: $\pm 28-24 62$), bilateral superior parietal lobe ($\pm 24-56 62$), occipital regions ($-6-70 14$) and anterior cingulate ($2 12 36$) ($Z > 2.3$, whole-brain corrected); Orange regions showed increased responses on correct relative to incorrect discriminations on aware trials ('2' and '3' ratings), including the right superior parietal lobe ($32-46 58$) ($Z > 1.96$, corrected). (B) Orange clusters ($Z > 2.3$, corrected) co-varied with delayed discrimination performance on unaware trials: bilateral aPFC regions included right BA47 ($42, 46, -14$), BA10 ($-30, 58, -2$) and BA11 ($\pm 32, 52, -10$), extending into left BA45 ($\pm 44, 36, 26$) and the PPC around the angular gyrus ($\pm 42-54 48$). (C) Regions that co-varied with non-conscious discrimination performance in the awareness contrast for '1' rating > '3' rating trials ($Z > 2.3$, whole-brain corrected). (D) The blue regions showed greater response on correct relative to incorrect discrimination on unaware trials and were located in BA10 ($-24 56 10$) BA11 ($-28 62 2$) extending more dorsally into BA9 ($-20, 46, 32$), BA46 ($-28 40 32$) and BA44 ($-30 14 42$) ($Z > 2.3$, whole-brain corrected). The orange clusters depict regions that correlated with individual variability in memory performance in the contrast for non-conscious hits > error responses ($Z > 2.3$, whole-brain corrected). The activation maps for the blue and orange contrasts overlapped in the left DLPFC (MNI $-30 16 42$).

Remarkably, we found clusters of activity in the dorsolateral prefrontal cortex (DLPFC) and anterior PFC (aPFC) along with posterior parietal regions (PPC) that were positively correlated with inter-individual variability in non-conscious memory performance on unaware trials (Fig. 3B), such that increased activity was associated with better memory performance on unaware trials. We found no regions that were more active with decreasing levels of awareness of the cue (i.e. 1-rating > 3-rating trials). Hence the mechanism supporting non-conscious delayed memory performance was not merely modulated by increased sensory uncertainty but was modulated by the accuracy of the response on the non-conscious trials (see below). To further test that the correlation between individual memory performance on non-conscious trials and neural response was indeed specific to responses on the non-conscious trials, we computed contrast of parameters estimates for BOLD responses on '1' unaware rating > '3' aware rating trials and correlated them with individual memory performance on the non-conscious trials. We found that responses in the aPFC, DLPFC, and PPC were again positively correlated with memory performance on unaware trials, such that increased BOLD responses were associated with better performance (Fig. 3C).

Our main aim here was to assess whether memory signals associated with the delayed discrimination performance can be reliably observed in the non-conscious trials. Nevertheless we also assessed correlations between individual memory performance on partially aware (2-rating) trials and BOLD responses and likewise in the aware (3-rating trials). While no significant correlations between brain activation and memory performance in the aware (3-rating) trials were apparent even at a very lenient threshold ($P = 0.01$, uncorrected), we found a significant correlation between responses in the right rostral part of PFC (MNI: $44 48-16$ BA46; $P < 0.05$ corrected) and memory performance in 2-rating trials (i.e. 'saw a brief glimpse but not the orientation of the cue'). We further assessed the presence of significant differences between neural responses associated to memory performance on 1-rating unaware trials relative to the 2-rating trials (in which participants reported a brief

glimpse of the cue but not its orientation), but we only found voxels in the PFC highlighted above at a more relaxed statistical threshold (data not shown).

Critically, we also conducted analyses independent of the performance covariate to contrast correct vs. incorrect responses on unaware trials (to avoid floor effects these analyses included only observers whose delayed memory discrimination performance was just above 0.5 chance; see Methods). We found that PFC was more active on hits relative to error responses on non-conscious trials (Fig. 3D, in blue; note also the orange clusters in Fig. 3D in the context of the same contrast, now for the covariate effect of individual memory performance including even subjects that scored <50% correct on unaware trials). Hence, we found reliable PFC memory signals in the non-conscious trials both in the correlations between individual neural responses and behavioural performance and also in the contrast between hits and error neural responses on the non-conscious trials.

Notably, we found no regions above threshold that co-varied with individual cue detection sensitivity (d') in any of the above contrasts. We also performed a direct comparison between the BOLD responses covarying with non-conscious memory performance vs. cue sensitivity and replicated the pattern found in Fig. 3B. Hence, DLPFC, aPFC and PPC activations were more linked to non-conscious memory discrimination performance.

Experiment 2: Transcranial direct current stimulation (tDCS)

Experiment 2 delivered tDCS over the left PFC (1.5 mA for 15 min; see also Methods) in a new sample of sixteen participants to obtain causal evidence that PFC supports delayed memory performance on unaware trials.

The proportion of awareness responses on cue present trials is given in Table 1B. Delayed memory performance on unaware trials was above chance, both following anodal ($t(14) = 3.13$, $P < 0.007$, two tailed) and cathodal stimulation ($t(14) = 2.68$, $P < 0.018$, two tailed) but more

Table 4
Experiment 2. Mean reaction times (ms) of correct responses.

	Anodal	Cathodal
Unaware trials	1570.83	1575.86
All trials	1353.24	1368.87

critically, performance was higher in the anodal relative to the cathodal condition (62% vs 57%; $t(14) = 2.51$, $P < 0.025$, two tailed). These results are depicted in Fig. 2C. Table 2B illustrates memory performance across the different awareness ratings and tDCS conditions. Of note, no tDCS effect was found when participants reported partial awareness of the cue (e.g. on 2 rating trials) even though memory performance was not at ceiling (anodal accuracy = 71%; cathodal = 74%, $P = 0.49$).

We found no evidence of tDCS effects on response latencies (all P s > 0.7 for the comparison of anodal and cathodal RTs of the correct responses on 1-rating trials and also when all 1–2-, and –3 rating trials were considered together). (We also note that the same pattern of results was observed when all RTs both from correct and incorrect responses were considered). Table 4 illustrates the mean RTs across the different conditions.

Together, this pattern of results suggests that the tDCS modulation was specific to memory accuracy on non-conscious trials. Furthermore, while polarization of the left PFC modulated delayed cue–target discrimination accuracy (i.e. the memory performance) on the non-conscious trials, there was no evidence of tDCS effects on the ability of participants to detect the cue itself ($t(14) = -0.466$, $P = 0.648$; anodal $d' = 0.406$; cathodal $d' = 0.482$). Fig. 2D illustrates cue detection sensitivity across tDCS conditions. Similar results were obtained with A' ($t(14) = -0.884$, $P = 0.392$; anodal $A' = 0.608$; cathodal $A' = 0.558$). The relevant SDT performance scores are reported in Table 3.

Objective cue detection sensitivity (d') on unaware trials was found in this experiment significantly above chance ($d' = 0$; $P < 0.017$ following anodal and $P < 0.034$ following cathodal tDCS) but we argue that this result may have been triggered by non-conscious processes. Critically, the output of a regression analyses with delayed discrimination as the dependent variable and d' as the predictor showed that memory performance following anodal tDCS was significantly 3.1% above chance ($P < 0.001$) when cue d' was extrapolated to null performance (for a justification of this method see Greenwald et al., 1996; Hannula et al., 2005).

Discussion

Behavioural results showed that visual stimuli below the threshold for conscious report can nevertheless be maintained and accessed for later use in a delayed memory discrimination test. The current neuroimaging results, from both functional MRI signals and the effects of neurostimulation, indicate that the PFC can support visual working memory operations independently of conscious awareness of the critical information. This conclusion was supported (i) by significant correlations between individual memory performance and PFC activations in the non-conscious trials, even when individual cue detection sensitivity was accounted for in the model (ii) by PFC sensitivity to the correctness of the discrimination responses on the non-conscious trials (iii) by the modulation of memory performance on unaware trials by PFC–tDCS, which was not apparent on the more aware trials.

The present results suggest that the role of human PFC in working memory may not be restricted to visible items but may also extend to non-conscious processes in the service of visual working memory. It may be argued that our task involved a visual WM load of a single item for a primitive visual feature (i.e. orientation). In our previous behavioural studies (Soto et al., 2011), we showed that non-conscious processing for WM successfully occurred even when distracters (either visible or masked) appeared during a 5 s delay between the cue and the memory test, and even when 2 items were held in memory for the

subsequent discrimination test. Further research ought to fully characterise the neural substrates of non-conscious visual WM processes in other settings that involve higher memory loads or intervening distracters that may engage additional control processes in WM.

We note that cue detection sensitivity on unaware trials did not significantly differ from chance levels in Experiment 1 and further note that cue detection sensitivity and memory performance on unaware trials were uncorrelated in Experiment 1. Hence participants with higher cue sensitivity did not display higher memory performance on unaware trials here, which would be expected if cue awareness mediated the level of memory performance. Furthermore, while cue detection sensitivity was found to be above chance in Experiment 2, regression analyses showed a significant intercept for memory performance on unaware trials when cue detection sensitivity was extrapolated to zero chance. On these grounds, delayed memory performance in the present study was dis-associated from conscious awareness. Critically, it should be noted that objective sensitivity measures derived from signal detection theory do not provide an accurate indication of the participants' awareness of stimuli because objective sensitivity measures (i.e. d') clearly tend to overestimate conscious perception (Dehaene and Changeux, 2011; Persaud et al., 2007; Sergent and Dehaene, 2004). According to objective signal detection theory (SDT) approaches, non-consciousness is merely equated with null performance (i.e. $d' = 0$) and consciousness is hence associated with any process that leads to just about or higher than chance performance. Based on theoretical and empirical grounds, more recent approaches advocate for the use of subjective ratings to evaluate the actual level of awareness (Overgaard et al., 2010; Sandberg et al., 2010; see also Dienes and Perner, 1994; Lau, 2008). We focused on the analyses of memory performance on the 1-rating trials to prevent the influence of criterion biases in reporting awareness, on the grounds that 1-rating trials may represent the unaware state more truly (e.g.) than the 2-rating trials (see Overgaard et al., 2010; Sandberg et al., 2010 for further discussion on the efficacy of perceptual awareness scales to assess the actual level of awareness).

We therefore used a combination of both objective and subjective measures of visual performance to track the relationship between awareness and delayed memory performance, which taken together indicate that processes mediating memory performance may be dissociable from the processes linked to perceptual awareness.

Below we discuss further the implications of the present findings for current accounts of PFC function, working memory and consciousness. Intriguingly, we found little evidence here for PFC activation associated with memory accuracy in the more aware (3-rating) trials. We note however that frontoparietal engagement in working memory tasks tends to increase with higher levels of memory load (Curtis and D'Esposito, 2003; Todd and Marois, 2004; Xu and Chun, 2006). Hence the apparent reduction of PFC activity in these aware trials may have been due to the relatively low working memory load imposed by the task, which only required a single oriented grating to be maintained for the later delayed cue–target discrimination. Evidence suggests that the visual cortex, found activated here on correct delayed discrimination with increased awareness, may be sufficient to solve delayed discrimination tests under low memory loads with visible and single cues (Baumann et al., 2008; Serences et al., 2009).

Considering the well-established role of fronto-parietal cortex in visuospatial attention (Corbetta and Shulman, 2002), it may be argued that the association between fronto-parietal activity and non-conscious delayed discrimination performance might reflect increased attentional deployment on those (unaware) trials with strongest masking of the cue. While it is difficult to rule out the mediation of attention, because it has been shown that selective attention can modulate processing of non-conscious stimuli (Bahrami et al., 2007; Kentridge et al., 2008; Sumner et al., 2006) it is worth noting that a well-known effect of attention is to boost perceptual signal strength (Carrasco et al., 2004; Carrasco et al., 2000; Downing, 1988; Hawkins et al., 1990; Lee et al., 1999). Hence attention would be expected to modulate cue sensitivity in our study but

we found no brain regions where responses significantly covaried with cue sensitivity (d'). Remarkably, the more anterior PFC in the frontal pole was sensitive to correct (vs. error) discrimination responses in the non-conscious trials, yet this region falls outside the classical fronto-parietal cortex associated with preparatory control and orienting of visual attention to relevant targets (Corbetta and Shulman, 2002).

Classical priming paradigms indicate that the presentation of non-conscious primes preceding a conscious target, which may be associated with the same response as the task-relevant target, can indeed influence participants' responses (Dehaene and Changeux, 2011). Previous studies have reported that superior frontal and dorsolateral PFC regions may be engaged in a bottom-up manner by non-conscious primes in the context of high-level cognitive control tasks involving response conflict (Lau and Passingham, 2007; van Gaal et al., 2010) but, unlike here, did not address the substrates that mediate working memory processes such as delayed discrimination. Tsushima et al. (2006) showed that non-conscious distracters can produce enhanced interference relative to conscious counterparts. They also found that PFC activity in the presence of distracters was attenuated with non-conscious relative to conscious distracters and suggested that PFC cognitive control mechanisms were less successful on the non-conscious distracters. Our findings suggest however that non-conscious signals that are associated with current behavioural goals may permeate into working memory systems and engage PFC.

Furthermore, we note that non-conscious priming effects reported in the literature are typically found with brief prime-target delays in the order of hundreds of milliseconds (Kouider and Dehaene, 2007; though see Bar and Biederman, 1998; Kunst-Wilson and Zajonc, 1980). More critically, non-conscious priming has been typically found under conditions which do not require observers to actively engage in processing the non-conscious primes (i.e. the primes are task-irrelevant). In the present paradigm, however, the interval between the cue and the test is in the range of seconds (see also Soto et al., 2011) and participants are explicitly required to treat the masked cues as goal-relevant for a subsequent delayed discrimination.

The human PFC is known to subserve the mental blackboard for high-level cognition such as cognitive control, working memory and decision-making but the extent to which this is reliant on conscious awareness of the critical information has remained elusive. For example, the aPFC has been associated with explicit maintenance and retrieval of concurrent task goals (Koechlin and Hyafil, 2007), top-down conscious visual perception (Bar et al., 2006), perceptual decision making based on a sense of intuition or gist (Volz and von Cramon, 2006), decision making under uncertainty (Hsu et al., 2005) and metacognitive processes that evaluate the subjective confidence of perceptual decisions (Fleming et al., 2010). However, PFC engagement here did not reflect metacognitive processing or levels of decision confidence/uncertainty because PFC signals were strictly associated with memory discrimination accuracy on non-conscious trials. Our evidence for aPFC involvement in non-conscious visual working memory is reminiscent of neuroimaging evidence from single-case study of patients with perceptual awareness deficits (i.e. 'blindsight', spatial neglect or prosopagnosia) which showed aPFC activity associated with implicit detection of non-conscious signals (Sahraie et al., 1997; Valdes-Sosa et al., 2011; Vuilleumier et al., 2002).

The human PFC has been shown to play a role in conscious visual experience (Del Cul et al., 2009; Lau and Passingham, 2006) and the fronto-parietal cortex is broadly conceived as a key substrate of the neural global workspace, 'broadcasting' sensory information in a recurrent manner (Dehaene and Naccache, 2001) and leading to conscious access of sensory contents in visual selection tasks (Beck et al., 2001; Del Cul et al., 2007; Lumer and Rees, 1999). Our findings, however, indicate that PFC regions extending into the more rostral PFC are also functionally relevant for processes operating on non-conscious information, which critically go beyond automatic sensorimotor priming to support working memory processes and higher-level cognitive function.

Acknowledgments

This work was supported by a grant from the Medical Research Council (UK, 89631). We thank the staff at the Birkbeck-UCL Centre for Neuroimaging for their help with the scanning protocol.

Conflict of interest

The authors declare no conflict of interest.

References

- Baars, B.J., Franklin, S., 2003. How conscious experience and working memory interact. *Trends Cogn. Sci.* 7, 166–172.
- Baddeley, A., 2003. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839.
- Bahrami, B., Lavie, N., Rees, G., 2007. Attentional load modulates responses of human primary visual cortex to invisible stimuli. *Curr. Biol.* 17, 509–513.
- Bar, M., Biederman, I., 1998. Subliminal visual priming. *Psychol. Sci.* 9, 464–469.
- Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmid, A.M., Dale, A.M., Hamalainen, M.S., Marinkovic, K., Schacter, D.L., Rosen, B.R., Halgren, E., 2006. Top-down facilitation of visual recognition. *Proc. Natl. Acad. Sci. U. S. A.* 103, 449–454.
- Baumann, O., Endestad, T., Magnussen, S., Greenlee, M.W., 2008. Delayed discrimination of spatial frequency for gratings of different orientation: behavioral and fMRI evidence for low-level perceptual memory stores in early visual cortex. *Exp. Brain Res.* 188, 363–369.
- Beck, D.M., Rees, G., Frith, C.D., Lavie, N., 2001. Neural correlates of change detection and change blindness. *Nat. Neurosci.* 4, 645–650.
- Carrasco, M., Penpeci-Talgar, C., Eckstein, M., 2000. Spatial covert attention increases contrast sensitivity across the CSF: support for signal enhancement. *Vis. Res.* 40, 1203–1215.
- Carrasco, M., Ling, S., Read, S., 2004. Attention alters appearance. *Nat. Neurosci.* 7, 308–313.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.
- Curtis, C.E., D'Esposito, M., 2003. Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423.
- Dehaene, S., Changeux, J.P., 2011. Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227.
- Dehaene, S., Naccache, L., 2001. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79, 1–37.
- Del Cul, A., Baillet, S., Dehaene, S., 2007. Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol.* 5, e260.
- Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., Slachevsky, A., 2009. Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain* 132, 2531–2540.
- Dienes, Z., Perner, J., 1994. Assumptions of subjective measures of unconscious mental states: three mappings. Higher order theories of consciousness. John Benjamins Publishers, Amsterdam.
- Donaldson, W., 1993. Accuracy of d' and A' as estimates of sensitivity. *Bull. Psychon. Soc.* 31, 271–274.
- Downing, C.J., 1988. Expectancy and visual-spatial attention: effects on perceptual quality. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 188–202.
- Fleming, S.M., Weil, R.S., Nagy, Z., Dolan, R.J., Rees, G., 2010. Relating introspective accuracy to individual differences in brain structure. *Science* 329, 1541–1543.
- Green, D.M., 1964. General prediction relating yes-no and forced-choice results. *J. Acoust. Soc. Am.* 36, 1042.
- Greenwald, A.G., Draine, S.C., Abrams, R.L., 1996. Three cognitive markers of unconscious semantic activation. *Science* 273, 1699–1702.
- Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* 48, 63–72.
- Hannula, D.E., Simons, D.J., Cohen, N.J., 2005. Imaging implicit perception: promise and pitfalls. *Nat. Rev. Neurosci.* 6, 247–255.
- Hawkins, H.L., Hillyard, S.A., Luck, S.J., Mouloua, M., Downing, C.J., Woodward, D.P., 1990. Visual attention modulates signal detectability. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 802–811.
- 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.
- Herwig, U., Satrapi, P., Schonfeldt-Lecuona, C., 2003. Using the international 10–20 EEG system for positioning of transcranial magnetic stimulation. *Brain Topogr.* 16, 95–99.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., Camerer, C.F., 2005. Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841.
- Keeser, D., Meindl, T., Bor, J., Palm, U., Pogarell, O., Mulert, C., Brunelin, J., Moller, H.J., Reiser, M., Padberg, F., 2011. Prefrontal transcranial direct current stimulation changes connectivity of resting-state networks during fMRI. *J. Neurosci.* 31, 15284–15293.
- Kentridge, R.W., Nijboer, T.C., Heywood, C.A., 2008. Attended but unseen: visual attention is not sufficient for visual awareness. *Neuropsychologia* 46, 864–869.
- Koechlin, E., Hyafil, A., 2007. Anterior prefrontal function and the limits of human decision-making. *Science* 318, 594–598.
- Kouider, S., Dehaene, S., 2007. Levels of processing during non-conscious perception: a critical review of visual masking. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 857–875.

- Kunst-Wilson, W.R., Zajonc, R.B., 1980. Affective discrimination of stimuli that cannot be recognized. *Science* 207, 557–558.
- Lau, H.C., 2008. *Are We Studying Consciousness Yet?* Oxford University Press.
- Lau, H.C., Passingham, R.E., 2006. Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proc. Natl. Acad. Sci. U. S. A.* 103, 18763–18768.
- Lau, H.C., Passingham, R.E., 2007. Unconscious activation of the cognitive control system in the human prefrontal cortex. *J. Neurosci.* 27, 5805–5811.
- Lee, D.K., Itti, L., Koch, C., Braun, J., 1999. Attention activates winner-take-all competition among visual filters. *Nat. Neurosci.* 2, 375–381.
- Lumer, E.D., Rees, G., 1999. Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1669–1673.
- Magnussen, S., 2000. Low-level memory processes in vision. *Trends Neurosci.* 23, 247–251.
- Overgaard, M., Timmermans, B., Sandberg, K., Cleeremans, A., 2010. Optimizing subjective measures of consciousness. *Conscious. Cogn.* 19, 682–684 (discussion 685–686).
- Pasternak, T., Greenlee, M.W., 2005. Working memory in primate sensory systems. *Nat. Rev. Neurosci.* 6, 97–107.
- Persaud, N., McLeod, P., Cowey, A., 2007. Post-decision wagering objectively measures awareness. *Nat. Neurosci.* 10, 257–261.
- Rosenthal, C.R., Kennard, C., Soto, D., 2010. Visuospatial sequence learning without seeing. *PLoS One* 5, e11906.
- Sahraie, A., Weiskrantz, L., Barbur, J.L., Simmons, A., Williams, S.C., Brammer, M.J., 1997. Pattern of neuronal activity associated with conscious and unconscious processing of visual signals. *Proc. Natl. Acad. Sci. U. S. A.* 94, 9406–9411.
- Sandberg, K., Timmermans, B., Overgaard, M., Cleeremans, A., 2010. Measuring consciousness: is one measure better than the other? *Conscious. Cogn.* 19, 1069–1078.
- Serences, J.T., Ester, E.F., Vogel, E.K., Awh, E., 2009. Stimulus-specific delay activity in human primary visual cortex. *Psychol. Sci.* 20, 207–214.
- Sergent, C., Dehaene, S., 2004. Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychol. Sci.* 15, 720–728.
- Sklar, A.Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., Hassin, R.R., 2012. Reading and doing arithmetic nonconsciously. *Proc. Natl. Acad. Sci. U. S. A.* 109, 19614–19619.
- Smith, S.M., 2002. Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155.
- Soto, D., Mantyla, T., Silvanto, J., 2011. Working memory without consciousness. *Curr. Biol.* 21, R912–R913.
- Striener, C.L., Chapman, C.S., Goodale, M.A., 2009. “Real-time” obstacle avoidance in the absence of primary visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 106, 15996–16001.
- Sumner, P., Tsai, P.C., Yu, K., Nachev, P., 2006. Attentional modulation of sensorimotor processes in the absence of perceptual awareness. *Proc. Natl. Acad. Sci. U. S. A.* 103, 10520–10525.
- Thesen, S., Heid, O., Mueller, E., Schad, L.R., 2000. Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. *Magn. Reson. Med.* 44, 457–465.
- Todd, J.J., Marois, R., 2004. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754.
- Tsushima, Y., Sasaki, Y., Watanabe, T., 2006. Greater disruption due to failure of inhibitory control on an ambiguous distractor. *Science* 314, 1786–1788.
- Valdes-Sosa, M., Bobes, M.A., Quinones, I., Garcia, L., Valdes-Hernandez, P.A., Iturria, Y., Melie-Garcia, L., Lopera, F., Asencio, J., 2011. Covert face recognition without the fusiform-temporal pathways. *Neuroimage* 57, 1162–1176.
- van Gaal, S., Ridderinkhof, K.R., Scholte, H.S., Lamme, V.A., 2010. Unconscious activation of the prefrontal no-go network. *J. Neurosci.* 2010, 4143–4150.
- Volz, K.G., von Cramon, D.Y., 2006. What neuroscience can tell about intuitive processes in the context of perceptual discovery. *J. Cogn. Neurosci.* 18, 2077–2087.
- Vuilleumier, P., Armony, J.L., Clarke, K., Husain, M., Driver, J., Dolan, R.J., 2002. Neural response to emotional faces with and without awareness: event-related fMRI in a parietal patient with visual extinction and spatial neglect. *Neuropsychologia* 40, 2156–2166.
- Wickens, T.D., 2002. *Elementary signal detection theory*. Oxford University Press, Oxford, UK.
- Worsley, K.J., 2001. *Functional MRI: An Introduction to Methods*. Oxford University Press.
- Xu, Y., Chun, M.M., 2006. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440, 91–95.