Title Page

Double-dissociation of frequency-specific contributions of dorso-lateral and dorso-medial prefrontal cortex to familiarity and recollective processes in the primate

Abbreviated Title: Prefrontal oscillations and recognition memory

Zhemeng Wu (吴哲萌)[1], Martina Kavanova[1], Lydia Hickman[1], Fiona Lin[1], Erica Boschin[1], Juan M. Galeazzi[1], Lennart Verhagen[1,2], Mark J. Buckley[1]

1 Department of Experimental Psychology, University of Oxford, Oxford OX1 3SR, United Kingdom

2 Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

Correspondence Author:

Mark J. Buckley, M,A. D. Phil. E-mail: buckley@psy.ox.ac.uk

Zhemeng Wu, M,A. E-mail: zhemeng.wu@queens.ox.ac.uk

Conflict of interest statement The authors declare no competing financial interests.

Author contribution

Z.W. and M.J.B. designed research; Z.M., M.K., L.H., F.L, E.B., J.M.G, C.P., L.V., and M.J.B. performed research; Z.M., M.K., L.H., F.L. analyzed data; Z.M. and M.J.B. wrote the paper.

Acknowledgements

This work was supported by Wellcome Trust Strategic Award Grant (Ref: WT101092MA) and MRC Project Grant (Ref: MR/K005480/1). L.V. is supported by funding from the Wellcome Trust (WT100973AIA). We thank the team of expert animal technicians and veterinary staff and anaesthetists for their very high standards of animal care and husbandry throughout. We thank Carlos Pedreira and Martin O'Neill for NHP support in setting up the electrophysiology lab. We thank Olof van der Werf and Alberto Lazari for T.M.S. support and we thank the Oxford Centre for Human Brain Activity (OHBA) and Matthew Rushworth for access to TMS facilities.

Abstract

According to dual-process theories, recognition memory draws upon both familiarity and recollection. It remains unclear how primate prefrontal cortex (PFC) contributes to familiarity and recollection processes but frequency-specific neuronal oscillations are considered to play a key role. Here, non-human primate (NHP) electrophysiological recordings first showed that a specific subregion of macaque PFC (i.e., dorsolateral PFC, dlPFC) was implicated at a specific frequency (i.e., increased beta oscillatory power) in a specific phase of a recognition memory task (i.e., during sample presentation). Then, to assess generalization to humans we targeted left human dlPFC (BA 9/46) as well as left dorsomedial prefrontal cortex (BA 8/9) for comparison, and also vertex as a control, with low-beta frequency transcranial magnetic stimulation (compared to non-frequency-specific stimulation, and to no-stimulation control) during occasional sample presentations within a similar task; hence we investigated the causal importance for human memory of a location-specific, frequency-specific, and task-epoch-specific intervention derived directly from the NHP electrophysiological observations. Using a dual-process signal detection (DPSD) model based on analysing receiver operating characteristics (ROC) curves, we showed beta-frequency TMS caused decreased recollection when targeted to human dlPFC, but enhanced familiarity when targeted to dorsomedial prefrontal cortex. Non-frequency-specific patterns of stimulation to all sites, and beta-frequency stimulation to vertex, were all without behavioural effect. This study provides causal evidence that PFC-mediated contributions to object recognition memory are modulated by beta-frequency oscillations; more broadly it provides translational evidence bridging NHPs and humans by emphasizing functional roles of beta oscillations in homologous brain regions in recognition memory.

Key words: electrophysiology; TMS; prefrontal cortex; recollection; familiarity; recognition-memory

Significance Statement

Recognition memory is assumed to draw upon both familiarity and recollection. Little is known yet about how neural oscillations in different sub-regions of primate prefrontal cortex (PFC) underlies these two processes. By recording macaque local field potentials we first observed modulation of beta-frequency oscillatory power in dorsolateral PFC during sample stimulus encoding. Then, in an analogous recognition memory task in humans, we used targeted transcranial magnetic stimulation (TMS) to causally influence PFC mnemonic function during encoding. We revealed a double dissociation between the effects of beta-frequency stimulation to dorsolateral PFC in suppressing recollection process and dorsomedial PFC in enhancing familiarity process. These cross-species translational studies demonstrate a causal role of beta-frequency oscillations in PFC to primate recognition memory performance.

Introduction

Early neuronal recording studies in macaque prefrontal cortex (PFC) suggested sustained neuronal spiking might encode stimuli across delays in object recognition memory tasks (Fuster and Alexander, 1971; Rao et al., 1997; Wilson et al., 1993); whilst other mnemonic mechanisms in PFC have since been proposed (Lundqvist et al., 2016; Stokes, 2015), the necessity of PFC for object recognition memory has been confirmed by macaque lesion studies (Bachevalier and Mishkin, 1986; Kowalska et al., 1991; Levy and Goldman-Rakic, 1999; Meunier et al., 1997; Mishkin and Manning, 1978; Parker, 1998; Petrides, 2000, 1995). A significant body of research maintains that human recognition memory draws upon two processes: familiarity and recollection (Wixted, 2007; Yonelinas, 2002, 2001, 1994; Yonelinas et al., 2002, 1998; Yonelinas and Parks, 2007). One influential neuropsychology review concluded that PFC patients were numerically worse than controls at recognition but deficits in recall were more profound (Tulving et al., 1995). However, there is no consensus as to whether different sub-areas of PFC contribute differently to familiarity and recollection. Using paradigms that aim to dissociate recollection from familiarity (e.g. receiving operating characteristic (ROC) curve plotting across confidence levels, or dissociations of subjective reports of 'remembering' versus 'knowing'), human neuropsychological investigations present conflicting results. In some, lateral PFC lesions cause deficits in familiarity rather than recollection (Aly et al., 2011; Duarte et al., 2005; MacPherson et al., 2008), whereas others (Wheeler and Stuss, 2003) report that dorsolateral PFC (dlPFC) lesions don't impair either, in contrast to frontopolar lesions which only impaired recollection. Evidence from human neuroimaging studies also provides mixed evidence as to the relative contributions of PFC sub-regions to familiarity/recollection (R. N.A. Henson et al., 1999; Richard N.A. Henson et al., 1999; Horner et al., 2015; Johnson et al., 2013; Kafkas and Montaldi, 2012; Skinner and Fernandes, 2007; Solstad et al., 2006), as comprehensively reviewed by Scalici et al. (2017).

Beyond spiking activities, the study of frequency-specific neuronal oscillations has now emerged as a key area of systems neuroscience research; within and between-area oscillatory coherence is increasingly considered to help mediate cognition, including learning and memory (Benn et al., 2016; Buzsáki and Schomburg, 2015; Fell and Axmacher, 2011; Fries, 2015; Gregoriou et al., 2009; Herweg et al., 2016; Köster et al., 2014; Pipa, 2009; Womelsdorf et al., 2007). In this context the local field potential (LFP) is thought to reflect a summation of local transmembrane currents (Buzsáki et al., 2012)
so LFP amplitude may reflect local synchrony within an area. Although studies have analysed LFP activity in macaque PFC during recognition memory they facilitate comparison between PFC regions and even cortical layers (Bastos et al., 2018). One study indicated dynamic modulation of frequency-specific oscillations in lateral PFC during working memory encoding wherein beta oscillations may reflect a default network state periodically interrupted by gamma oscillations when encoding or decoding stimuli (Lundqvist et al., 2016); another implicated low-frequency oscillations in top-down and higher-frequency oscillations in bottom-up processing (Kornblith et al., 2016).

Transcranial magnetic stimulation (TMS) in humans may assess causality of accessible PFC regions; single or repetitive pulses before or during task performance modulate ongoing neural oscillations (Huang et al., 2005; Rossi and Rossini, 2004), disrupting or enhancing behaviour in a frequency-specific manner (Albouy et al., 2017; Chanes et al., 2013; Elkin-Frankston et al., 2011; Thut et al., 2011); for example, perceptual discrimination was enhanced by high-beta frequency TMS over frontal eye fields, whereas response criterion was lowered by gamma-frequency TMS (Chanes et al., 2013).

Here we first observed increased beta power LFP in macaque dlPFC during sample presentation in object recognition memory. According to desynchronization hypotheses (Simon Hanslmayr et al., 2012; Holmes et al., 2018) PFC neurons desynchronize to retain stable memories. Therefore, to assess causality and generalizability across primate species we

hypothesized that targeting the homologous human region with beta TMS during sample presentation in a similar task may lessen desynchronization and disrupt memory for those samples. As human activation peaks related to familiarity were generally found more caudally, corresponding to dorsomedial PFC (dmPFC), in meta-analyses (Scalici et al., 2017), we further hypothesized that targeting dlPFC would primarily affect recollection indices, while targeting dmPFC (our chosen comparison target for this reason) would primarily affect familiarity indices. We included vertex as a control region wherein we hypothesized no effect of TMS, and we included non-beta frequency (random) stimulation as a control frequency unlikely to prevent beta desynchronization.

Materials and Methods

Experiment 1: NHP electrophysiological study.

Subject

Neural data was recorded in one young adult female macaque monkey (*Macaca mulatta, age 8 years, weight 10-13 kg*). All animals in our lab are socially housed (or socially housed for as long as possible if later precluded, for example, by repeated fighting with cage-mates despite multiple regrouping attempts) and all are housed in enriched environments (e.g. swings and ropes and objects, all within large pens with multiple wooden ledges at many levels) with a 12hr light/dark cycle. The NHP always had *ad libitum* water access 7 days/week. Most of its daily food ration of wet mash and fruit and nuts and other treats was delivered in the automated testing/lunch-box at the end of each behavioral session (this provided 'jack-pot' motivation for quickly completing successful session performance; supplemented by trial-by-trial rewards for correct choices in the form of drops of smoothie delivered via a sipping tube) and this was supplemented with fruit and foraging mix in the home enclosure. All animal training, array implantation surgery, and experimental procedures were performed in accordance with the

guidelines of the UK Animals (Scientific Procedures) Act of 1986, licensed by the UK Home Office, and approved by Oxford's Committee on Animal Care and Ethical Review.

Surgical procedures

After the basic behavioural and initial task training phase was complete the animal first received a titanium head-post, implanted posteriorly and secured to the cranium with titanium cranial screws through the legs of the post (the legs of the head-post were pre-shaped to fit the precise morphology of the skull in the region according to a 3D printed skull model based on pre-operative structural MRI scans); the reflected skin and galea was sutured and the wound closed around the base of the head-post.

Later, after more behavioural training with head-fixation was complete and task-performance satisfactory, the NHP received surgical implantation of microelectrode arrays (Utah arrays, Blackrock Microsystems). A bone flap was raised over the left anterior prefrontal cortex, the dura mater was cut and reflected, and Utah arrays implanted directly into the cortex; the dura mater was sewn back over the arrays where possible and the bone flap was replaced. Two reference wires connecting to the pedestals were left under the dura away from the site of the array and these and the wire bundle connecting to each electrode in the Utah array exited the cranium through a gap at the edge of the craniotomy and ran from there to the pedestal which was secured to the cranium (located away from the edge of the craniotomy) with titanium cranial screws through its base; to complete the procedure the wound was then closed in layers. A pedestal cap was screwed on to the top of the pedestal to protect the gold contacts (the cap was subsequently removed each day, while the NHP in the recording session, to connect the digital head-stage direct to the pedestal's connectors at which point recording commenced).

The operations to implant the head-post and later to implant the microelectrode arrays were performed in aseptic conditions with the aid of an operating microscope for intra-cranial surgeries. The monkeys were first sedated with ketamine (5 mg/kg), medetomidine (20mcg/kg) and midazolam (0.1 mg/kg) given i.m., intubated, and then artificially respirated using a mixture of carrier gases (oxygen/medical air) and volatile anaesthetic. Surgical depth of anaesthesia was maintained throughout the surgery with sevoflurane (1.0-2.0% to effect) and injectable adjuncts (fentanyl 5 mcg/kg/hr i.v., dexmedetomidine 0.5 mcg/kg/hr i.v.). On average three doses of steroids (methylprednisolone 20 mg/kg i.v. every 4-6 hrs) and in the case of intra-cranial surgery a bolus of mannitol (1000 mg/kg i.v.) were given on the day of surgery to protect against intraoperative brain edema and postoperative inflammation. Steroids were continued in the postoperative phase (dexamethasone 0.2 mg/kg s.c., daily for 5 days). The animals were given an antibiotic (30 mg/kg of amoxicillin intraoperatively every 2 hours, and 17.5 mg/kg daily postoperatively) for prophylaxis of infection. Additional intraoperative medication included atropine (10 mcg/kg i.v. as required), an H2 receptor antagonist ranitidine (1 mg/kg i.v.), meloxicam (0.2 mg/kg i.v.) and crystalloid fluids (Hartmann's solution 3-5 ml/kg/hr). Heart rate, oxygen saturation of hemoglobin, mean arterial blood pressure, end- tidal CO2, body temperature, and respiration rate were monitored continuously throughout surgery. Postoperative analgesia was provided via opioids (methadone 0.3 mg/kg i.m. or buprenorphine 10 mcg/kg i.m.) and nonsteroidal anti-inflammatory agents (0.1 mg/kg of meloxicam, p.o./i.m., 10 mg/kg acetaminophen p.o.). A proton pump inhibitor (Omeprazole 0.5 mg/kg) was given daily to protect against gastric ulceration as a side effect of the combination of steroid and nonsteroidal anti-inflammatory treatment.

This study records data from one 32-electrode Utah array inserted into dlPFC (area 9/46, just dorsal to the principal sulcus, see Fig. 3A); the electrodes were arrange on a 6 × 6

grid embedded in silicone; the electrode length was 1.0 mm with interelectrode spacing of 0.4 mm) All the microelectrodes were made of platinum (Pt).

Task stimuli and apparatus

The object recognition memory task was programmed using Turbo Pascal (Borland), run under DOS on a desktop PC. Visual stimuli used in the task were clip-art images in colour, which were presented on a 20.1" colour touch-sensitive screen (TFT LCD TS200H GNR). Those clip-art images used in the task were from a large pool of several thousand unique images. Each image subtended 5° of visual angle in width and 5° in height to the subject when presented on the screen. The sample image was always presented on the right top of the screen, positioned +9 ° horizontal and -9 ° vertical from the center of the screen. The test images were presented on the right bottom of the screen: one was positioned 0 ° horizontal and +3 ° vertical from the center of the screen; and the other one was positioned +17 ° horizontal and +3 ° vertical from the center of the screen. The background colour to the screen was white. In each session of recordings, images were randomly chosen from the pool without replacement and were not re-used on the other recording days.

The animal was seated in a primate chair (Rogue Research Inc.) in front of the touch screen with its head-fixated and whilst it performed the recognition memory task in a magnetic-shielded, and partially sound-attended, testing-box. A window in the front of the chair provided its access, both to the touch-screen itself and also to a touch-sensitive knob which we refer to as a 'key-touch' which was positioned low down in front of the touch screen; the animal had to steadily hold the keytouch at various times in each trial (to control for arm movement/position whilst it waited for stimuli, looked at stimuli, and waiting for a visual cue to release keytouch and touch the screen to make a choice). The distance between the monkey and touch screen was fixed at 50 cm enabling the animal to touch the screen easily. An infrared camera was used to monitor the general status of the monkey in the box.

A peristaltic pump device located on top of the box fed smoothie reward through a tube and to a spout positioned in the vicinity of the animal's mouth. Below the screen was also an automated lunch-box which contained the majority of the animal's daily meal (wet mash and fruits and nuts etc.) and which opened immediately at the end of the task.

Behavioural task

The task is a variant of the well established delayed-matching-to-sample (i.e. recognition-memory) paradigm in which a stimulus in a sample phase has to be judged as familiar or not in a subsequent choice phase after a short delay; the form of the task used is similar to that used by Basile and Hampton (2013) in that in the choice phase, to allow separation of hits, misses, false alarms, and correct rejections, one choice image is presented with one non-match button (black circle) such that the animal selects the choice image if it considers the choice image a match and selects black circle if it considers it is a non-match (see Fig.1). In each trial, the animal initiated the task by holding the key-touch when cued to do so by a small red circular keytouch cue (located towards upper centre of screen; Fig. 1) presented on the screen. The NHP was required to keep holding the keytouch device through a variable delay of 1000-1500 ms (if the keytouch hold was broken the trial aborted) after which a central sample (object image) appeared behind the red key-touch cue. The NHP was required to keep holding the keytouch device through another variable delay of 1000-1500 ms (again, if the keytouch hold was broken the trial aborted) after which the red keytouch cue disappeared which was the cue to the animal it could release its hand from the keytouch at which point the delay (1000ms) between sample and choice-phases began (the maximal time for releasing key-touch was 5000 ms else the trial aborted and the animal received a time-out for 10 s). Then after the 1000 ms delay period another red keytouch cue appeared, this time in the bottom of the screen (close to, and equidistant, from where the two choice items would appear) and animal was required to hold the key-touch again. The NHP was again required to

keep holding the keytouch device through a variable delay of 1000-1500 ms (if the keytouch hold was broken the trial aborted) after which two choice stimuli appeared (one an object image and the other a black circle, left-right randomized between trials). The NHP was required to keep holding the keytouch device through another variable delay of 1000-1500 ms (again, if the keytouch hold was broken the trial aborted) after which the red keytouch cue disappeared which was the cue to the animal that it could release holding the keytouch and now make a choice to the touchscreen to either the object test-image stimulus or to the black circle stimulus. The object stimulus was either the identical stimulus to the sample seen earlier in the trial or it was not identical to the sample. The animal was rewarded by delivery of 10 ml of smoothie for touching the test-item image if it matched the sample image (these we refer to as 'match trials'), or it was rewarded for selecting the standard 'non-match button' (i.e. the black circle) if the test-item was a non-match (these we refer to an 'non-match trials'). After a correct response, the intertrial interval was 3 s. However, following any error trial (including both incorrect response and aborted trials), the intertrial interval was 10 s Accordingly, on *match trials* the animal could either make a correct response ('hit') or an incorrect response ('miss') whereas on *non-match trials* the animal could either make a correct response ('correct rejection') or an incorrect response ('false alarm'); in this way the paradigm is similar to one previously used by Basile and Hampton (2013) but a key difference is that we did not restrict the stimulus set to just two stimuli as we used larger sets moreover we also varied the degree to which stimuli were either familiar or novel in the session. Specifically, in any given session 50% of the trials used trial-unique stimuli and 50% were 'repeat' stimuli used previously in the session (but not used in any previous session); in each session there were six repeat stimuli sorted into 3 pairs such that in each trial with repeat stimuli one pair was chosen at random and one member of the pair was randomly chosen to be the sample. Each session typically consisted of 150 trials so the six repeat stimuli got steadily more familiar across the session.

------ Insert Figure 1 about here ------

Electrophysiological Recordings

Data were recorded simultaneously from the 32-electrodes in the microelectrode array implanted in dlPFC throughout the recognition memory task. We only considered the trial-unique stimuli in these analyses as we used the trial-unique stimuli in human TMS study. The NHP performed, on average, 75 trials (with trial-unique stimuli) per day; the total average session duration was 45 minutes. In this study we analysed data recorded over 29 sessions approximately 8-12 weeks after the array implantation. The NHP had learnt the task in full prior to array implantation. Neural signals from each microelectrode were amplified, digitized at 30 kHz using a 256-channel Cerebus™ Neural Signal Processor (Blackrock Microsystems). The local field potential (LFP) was analyzed off-line using the FieldTrip toolbox (*http://www.fieldtriptoolbox.org*, details see Oostenveld et al., 2011) for MATLAB (R2017b, MathWorks), as well as custom scripts (available upon request from corresponding author).

Signal analysis

We investigated the power of the induced LFP responses during the encoding (i.e. sample presentation) phase of the NHP recognition memory task. A notch filter, the second-order infinite impulse response (IIR), was applied to the raw data signals, to remove the power from line noise (i.e. 50 Hz) and its harmonics (i.e. 100, 150 Hz). Then the data signals were down sampled from 30 kHz to 1 kHz. To obtain the power spectrums of LFPs for each of the 32 channels, a time-frequency decomposition analysis based on Morlet wavelet transform was performed. The frequency range in LFP spectral analysis was chosen to start at 4 Hz (to

exclude slow-wave LFP components) and terminate at 140 Hz, in steps of 1 Hz.

For the 32 channels in dlPFC, the induced power spectra from LFPs were limited to data segments which contained completed correct trials and inter-trial intervals (ITIs). A 1000 ms segment during the ITI was picked as the baseline for all the power spectra, and their root mean square (RMS) was divided from the raw data before the calculation of the spectrum using a wavelet transform. In the recognition memory task, LFPs in both the pre- and post-sample presentation periods (both within a time window of 1000 ms while NHPs holding the keytouch) from successfully completed trials were selected and analyzed.

Statistical methods

To test for statistical significance of differences of the induced LFP power spectra between pre- and post-sample presentations during the recognition memory task, we performed a nonparametric permutation test, with the median difference between the above two conditions (i.e. pre and post-sample) as our test statistic. The non-parametric permutation test is an assumption-free method without prescribing underlying distributions (Maris and Oostenveld, 2007). The null hypothesis of this test was assumed to be that there was no significant difference of modulation of LFP power when comparing the observed difference between the two conditions against a reference distribution of differences between the two randomly assigned conditions. The reference distribution was obtained by performing 10,000 permutations on trial labels to randomly distribute the two conditions at at each frequency and at each time point. The observed median differences between the two conditions were selected whose values were larger than the 97.5th percentile of the maximum or smaller than the 2.5th percentile of the minimal in this reference distribution ($p < 0.05$, two-tailed). Then the selected values were clustered in connected sets on the basis of temporal adjacency. We calculated cluster-level statistics by taking the sum of the median-difference values within a cluster and took the largest of the cluster-level statistics as the significant differences between

the two conditions. This procedure is a two-tailed test with a global false positive rate of 5% and correction for the multiple comparisons across frequencies and time points (Maris and Oostenveld, 2007).

Experiment 2: Human TMS study

Participants

27 participants (18 male, 9 female, age range 18-30 years) took part in the TMS study. Participants were fluent English speakers, right-handed, and had normal or corrected-to-normal vision. Prior to the study, all the participants provided written consent and went through safety screening check to make sure they had no history of previous or current neurological or psychiatric conditions and were not taking any psychoactive medication. All the participants received monetary compensation for their participation at a standard rate for volunteers for TMS studies in Oxford. This study was carried out with the approval of Medical Sciences Interdivisional Research Ethics Committee, University of Oxford.

Stimulation Brain Sites

In this study, we investigated the differential effects of rTMS to dlPFC and dmPFC in recollection and familiarity in a recognition memory task designed to be similar to that used with NHP in Experiment 1. We chose left dlPFC which approximately corresponded to area BA46 or BA9/46 (referred to as BA 9/46) in light of Experiment 1, and we chose dmPFC along the midline as an area the literature led us to expect to provide a contrast with respect to consideration of contributions to recollection and familiarity, this more medial area corresponded to area BA 8 or BA9 (referred to as BA 8/9). A further control brain site was chosen to be vertex, which we had no reason to expect to be important for either memory process, and which was localized at center point of the head. The three targeted sites were measured using Beam F3 localization system (Beam et al., 2009). This system allows the

measurement of the location of the F3 electrode position in the 10-20 EEG coordinate system, and takes into account individual skull variations. Based on this system, BA 9/46 was localized 1 cm caudal to F3. Both vertex and BA8/9 sit on the midline, in which vertex was localized at a site corresponding to location of electrode Cz (measured as half the distance between inion and nasion and intersecting with half the distance between the two aural tragus); and BA8/9 was localized 7cm rostral to the vertex and 6 cm medial to F3. The relative positions of three stimulation sites are illustrated in Fig. 4A. The participants were assigned to one of the three brain site groups to equalize numbers per group: dlPFC group (BA 9/46, 9 participants), dmPFC group (BA8/9, 9 participants) and control group (9 participants).

Defining Participant's Motor Threshold

Prior to the experiment, all the participants were invited to a taster session, in which the individual's Resting Motor Threshold (RMT) was obtained for each participant. To measure RMT, stimulation was applied over the site of left primary motor cortex (M1) (localized 5 cm laterally and 5 cm rostrally to the vertex), where the largest twitch in right index finger of the participant was found. Site search was initially started at the lowest stimulator output. By gradually increasing output, the RMT of the participant was determined when 5 out of 10 TMS pulses caused a twitch in the right index finger (Rossini et al., 1994).

A stimulation intensity of 90% of RMT was used in the study. This intensity was within an appropriate range taking into account the average scalp-cortical surface distance between M1 and stimulation areas and a 2.8% reduction in stimulator output for every mm closer to the skull (Stokes, 2005; Stokes et al., 2013, 2007). The mean RMT of dlPFC group was 44.33 (SE = 3.61), and of dmPFC group was 51.33 (SE = 5.52), and of vertex group was 49.11 (SE = 7.39).

Task Stimuli and Apparatus

The task was an object recognition memory task similar to the one used in NHPs, similarly programmed using Turbo Pascal (Borland), run under DOS on a desktop PC and presented on a 20.1" colour touchscreen (TFT LCD TS200H GNR). The objects images used in the task were clip-art images as in NHP study, but in order to increase difficulty (in light of a pilot study wherein performance was close to ceiling) the images were all converted to grey-scale and the contrast toned down in an attempt to make them harder to discriminate from each other. Additionally, the samples were presented in lists followed by lists of choice trials, again to reduce ceiling effects in the human version of the NHP task. Each image subtended 10° of visual angle in width and 10° in height to the participant sitting facing the screen. The sample image was always presented on the right top of the screen, positioned +12 ° horizontal and -12 ° vertical from the center of the screen. The test images were presented on the right bottom of the screen: one was positioned 0 ° horizontal and +5 ° vertical from the center of the screen; and the other one was positioned +23 ° horizontal and +5 ° vertical from the center of the screen. The background colour to the screen was white.

Participants sat with their eyes a distance of 25 cm from the screen, wearing earplugs in order to avoid audio disturbance of TMS pulses, resting their chins on a chin-rest and their foreheads on a head holder to stabilize their head position throughout the experiment. They were instructed to respond to items by touching them on the screen and gestured their confidence ratings using their right hands. For example, they indicate by raising fingers (1, 2 or 3) whether their opinion corresponding to their being somewhat confident (1), moderately confident (2), or absolutely confident (3) in their judgment as to whether they considered the test-item to be old (i.e. presented before in the preceding list as a sample) or new (not seen before in the preceding list as sample). None of the samples in this task were used in more than one list so all stimuli were trial-unique (and hence compared to the novel/trial-unique stimuli in the NHP task).

Transcranial Magnetic Stimulation (TMS) was carried out using a biphasic Magstim Super Rapid^2 magnetic stimulator (Magstim, Dyfed, UK) with a double 70-mm figure-of-eight coil. The TMS coil was clamped in line with participants' assigned brain sites and localized at a 45° angle off the midline with the handle pointing to the posterior direction throughout the experiment.

Behavioural task

Prior to the experiment, participants were given instructions on how to perform the task, and how to indicate their confidence judgments (see above), and then they were introduced to the behavioural task in a short training/practice session (comprised of one list of 12 samples and then 12 test-item trials) of the task without stimulation to provide an opportunity to become familiarized with task procedure. We used lists in the human version to avoid ceiling effects as pilot investigations revealed the NHP task with single samples and single test-trials to be too easy for participants. The experimental session itself contained 3 sub-sessions, with each sub-session containing 15 blocks of trials. Participants took a 10-15 min break between each sub-session. Each block contained an encoding phase (i.e. a list of 12 sample images), and then a short delay, and then a test phase (i.e. a list of 12 test-item trials). The task structure is depicted, for one block, in Figure 2; during the encoding phase, 12 sample images were presented sequentially on the screen (individually for 480 ms), with inter-stimulus-interval as 350 ms. Participants were instructed to view and try to remember each sample image. After the sample phase (all 12 sample items) was completed, a blank screen was presented for 1 s (delay), and then the test phase commenced (all 12 test trials). Test trials were either 'match trials' or 'non-match trials'. In each test trial, either an identical stimulus to one of the preceding12 samples (i.e. 'match trials') was shown as a test image, or a novel and previously unseen image (i.e. 'non-match trials') was shown as a test image, and that test image appeared on the screen together with a black circle (the left/right position of the black circle and the test-trial image were randomized between trials). The match and non-match

trials were also counter balanced (6 of each per test phase of 12 test-trials) and were put in a random order. Just as in the NHP version of the task (Experiment 1) participants were instructed to touch the test image if they thought it matched one of the 12 sample images, or touch the standard 'non-match button' (i.e. the black circle) if they thought the test image was not a match. After responding to the test image, participants were instructed to rate their confidence as to whether the test item was new or old using a scale of 1-3 by making three different movements with their fingers, corresponding to somewhat, moderate, and absolute confidence. Participants were further instructed to try to use the entire range of confidence responses as best they could and not simply select the extremes of confidence as defaults.

------ Insert Figure 2 about here ------

Repetitive TMS Protocol

As our local field potential (LFP) investigations in the NHP (Experiment 1) revealed heightened beta power (i.e. increased beta frequency oscillations) in dlPFC when macaques viewed and encoded sample images in a similar recognition memory task, rTMS stimulation in this human study was targeted during the encoding phase of the memory task so as to assess whether that activity in that region at that time may be necessary for successful encoding and subsequent recognition in humans.

In line with the safety guidelines for TMS by Rossi et al. (2009) to avoid seizure risks caused by too much sequential high-frequency stimulation, a restriction was implemented that there were at least 2 non-stimulated images in between every stimulated image in each sample list of 12 items. Therefore, we determined all the possible combinations wherein 4 out of 12 sample images might be targeted, with at least 2 non-targeted intervening samples, and each block's sample phase took one of those schedules selected randomly each time. During

beta-stimulation condition blocks, each of the four rTMS targeted samples in the list (each with a duration of 480 ms as detailed above) were targeted with the delivery of 6 pulses every 80 ms (hence at 12.5 Hz and within the range of beta frequency); while in the random-frequency stimulation blocks, the same total number of pulses (6) were delivered less regularly over the same period of presentation time (specifically, the 480ms sample presentation epoch was divided into 30ms intervals and 6 of those, with the constraint that no two consecutive 30ms intervals could be chosen, were chosen at random to trigger a TMS pulse delivery). All these considerations of TMS frequency ensured we always kept well within the known safe ranges of TMS stimulation at high frequency and over extended durations of stimulation in line with the safety guidance of Rossi et al. (2009). Each sub-session contained 15 blocks, with each of the three 'types' of stimulation (i.e. beta-stimulation, random-stimulation, and no-stimulation) occurring 5 times per sub-session, randomly ordered. Each participant received the same sets of stimuli and same stimulation order. As there were 3 sub-sessions in total each stimulation condition was repeated 15 times per participant and so participants performed a total of 540 trials (i.e. 3 sub-sessions x 15 blocks per sub-session x 12 trials per block).

Data analysis

Behavioural effects of TMS were evaluated based on dual-process signal detection (DPSD) theory model. According to this model, both familiarity and recollection contribute to recognition memory (Yonelinas, 2002, 2001). A recollection index (R) and familiarity index (F) were extracted by fitting the model to data using a standard approach that minimized the squared difference between the observed and predicted data in each confidence rating bin. Specifically, participants were asked to express their confidence in each of their decision using a 6-point scale, where 1 implies absolutely certain old, 2 for moderately certain old, 3 for slightly certain old, 4 for slightly certain new, 5 for moderately certain new, and 6 for absolutely certain new. The cumulative hit rate was then plotted against the cumulative false

alarm rate to create the ROC plot of each participant. To be more specific, the first point at the left-hand side of the plot was the hit rate and false alarm rate at confidence level 6. After that, the second point was for the combination of confidence levels 6 and 5, the third point for confidence levels 6 to 4 and so on. The cumulative hit rate and false alarm rate of all confidence levels was constrained to 1.0. Accordingly, we had five coordinate points for each ROC plot.

The ROC data was analyzed using the DPSD model. This model assumes recognition memory to be contributed towards by two distinct processes referred to as recollection and familiarity, and the cumulative proportions of the target-item and lure-item trials in each confidence level (CL_i) are given by (Koen et al., 2017):

$$\mathrm{p(Lure|CL_i)} = (1 - R_n) * \Phi(0, 1)$$

$$\mathrm{p(Target|CL_i)} = R_0 + (1 - R_0) * \Phi(d'_F, \sigma_F)$$

In the target and lure distribution (Φ is the cumulative response function), R_0 is a target threshold parameter being labeled recollection of old stimuli, and R_n is a lure threshold parameter being labeled recollection of newness stimuli. In the lure item distribution, R_n is a threshold parameter. If the strength of the lure item is beyond this threshold, it will always be classified as new. We have set R_n to be 0 so the classification of lure item is always set as a standard normal Gaussian function with a mean of 0 and a unit standard deviation. Similarly, in the target item distribution, the R_0 parameter is the threshold for the classification of the target, which is an index of recollection, R. If the strength of a target item falls below R_0, its classification would be governed by the familiarity component of the model which is a Gaussian distribution with the mean of d'_F and standard deviation of σ_F. The standard deviation σ_F is commonly fixed to 1 in DPSD model, and the mean d'_F is the index of

familiarity, F. Based on the above assumptions, the algorithm of DPSD model is re-set as below:

$$\mathrm{p(Lure|CL_i)} = N(0,1)$$

$$\mathrm{p(Target|CL_i)} = R_0 + (1 - R_0) * \Phi(d'_F, 1)$$

Each pair of the free parameters, d'_F (F) and R_0 (R), defines a ROC curve. These parameters were obtained by minimizing the sum of squared residuals in both the abscissa and the ordinate of all the five points (that is, the sum of squared errors of prediction, or SSE, method for both axes). Based on the assumptions of the DPSD model, the range of R is between 0 and 1 while the range of F is between 0 and infinity (a theoretical maximum that is; practically it is much less, around 3-4 as a maximum). If the best fit of the *R* value turns out to be negative in the initial fitting of ROC and thus meaningless, we fixed it to be 0 and refit the model to get F. In each stimulation type and in each stimulation site condition, a ROC is plotted cumulatively for each confidence level by the proportion of correct 'old' judgements against the proportion of incorrect 'old' judgements, and both R and F in each stimulation type and stimulation site condition were extracted from DPSD model.

Statistical methods

All the statistical analysis on human behavioral data was carried out using SPSS software (IBM). The above calculated indices (i.e. R and F) were subjected to 2×3×3 mixed model repeated measures ANOVAs with 3 levels of the within-group factor stimulation-type (beta-frequency, random-frequency, and no-stimulation) and with 3 levels of the between-subject factor stimulation-site (dlPFC, dmPFC, and vertex).

Accuracy was assessed for each human participant in each stimulation condition by calculating the area under each ROC curve (AUC). AUCs were calculated using the parameters from a fit of the DPSD model to the individual ROCs. Response time (RT) for

correct responses were also evaluated for each participant under each stimulation condition. Then performance (i.e. accuracy or response time) was subjected to a 3×3 mixed model repeated measures ANOVAs with 3 levels of the within-subject factor stimulation-type (beta-frequency, random-frequency, and no- stimulation) and with 3 levels of the between-subject factor stimulation-site (dlPFC, dmPFC, and vertex).

Results

Experiment 1: NHP electrophysiological study.

During the NHP recognition memory task, we focused on the encoding (i.e. sample presentation) phase. A nonparametric permutation test as described earlier revealed a significant cluster ranging from 7-24 Hz starting immediately after the sample image onset in macaque dlPFC (two-tailed permutation test, $p < 0.05$). This post-sample alpha/beta power lasted roughly 300 ms (Fig. 3B). Thus, a stimulation frequency that was intermediate within this range (we chose 12.5 Hz, i.e. low beta) and that was well within the safety guidelines (with respect to duration and frequency considerations for repetitive stimulation) for humans (Rossi et al., 2009) was selected to be used in Experiment 2, our repetitive TMS (rTMS) study examining the effect of intervention at this frequency in the corresponding encoding phase of a homologous recognition memory task in humans.

------ Insert Figure 3 about here ------

Experiment 2: Human TMS study.

To summarise the results (statistical analyses described in detail below) we observed that only rTMS at beta frequency to dlPFC drove down the recollection index, and only rTMS at beta frequency to dmPFC drove up the familiarity index.

Impact of TMS on accuracy and response time in the human recognition memory task

Participants were highly accurate in their recognition memory task; mean performance (calculated as AUCs) is shown in Table 1. A mixed-model repeated measures ANOVA, with stimulation-site as the between-subjects factor and stimulation-type as the within-subject factors, showed that neither the main effects (i.e. stimulation-site, stimulation-type) nor their interaction was significant (main effect of stimulation-site: $F_{(2,24)} = 2.035$, $p = 0.153$, $\eta^2 = 0.145$; main effect of stimulation-type: $F_{(2,48)} = 1.252$, $p = 0.295$, $\eta^2 = 0.050$; interaction: $F_{(4,48)} = 0.529$, $p = 0.715$, $\eta^2 = 0.042$) with respect to AUC.

Response times (RTs) were also calculated for each participant group in beta frequency, random frequency, and no-stimulation TMS conditions (Table 1). The aforementioned mixed-model repeated measures ANOVA applied to RT also showed neither the main effects (i.e. stimulation-site, stimulation-type) nor their interaction was significant (main effect of stimulation-site: $F_{(2,24)} = 2.255$, $p = 0.127$, $\eta^2 = 0.158$; main effect of stimulation-type: $F_{(2,48)} = 1.184$, $p = 0.315$, $\eta^2 = 0.047$; interaction: $F_{(4,48)} = 1.387$, $p = 0.253$, $\eta^2 = 0.104$).

------ Insert Table 1 about here ------

Impact of TMS on recollection and familiarity in the human recognition memory task

The recollection index (R) and familiarity index (F) were calculated for each participant in each of the three stimulation types. A mixed-model repeated measures ANOVA was conducted on the indices data, with stimulation-site (3 levels: dlPFC, dmPFC, and vertex) as

the between-subjects factor and stimulation-type (3 levels: beta-frequency, random-frequency, and no- stimulation) and index-type (2 levels: recollection-index, familiarity-index) as the two within-subject factors. A significant main effect of index type ($F_{(1,24)}$ = 65.809, p = 0.00000002, η^2 = 0.733) was found, in addition to a significant two-way interaction between index-type and stimulation-type ($F_{(2,48)}$ = 6.112, p = 0.008, η^2 = 0.203) with Greenhouse-Geisser correction due to a significant Mauchly's test of sphericity. Although the three-way interaction between index type, stimulation type and stimulation sites did not attain significance ($F_{(4,48)}$ = 2.312, p = 0.071, η^2 = 0.162), the tests of within-subjects contrasts showed a significant linear effect of index type and a linear effect of stimulation type ($F_{(2,24)}$ = 5.190, p = 0.013, η^2 = 0.302) prompting further analyses.

To scrutinize the effects further, repeated-measures ANOVAs were conducted on each group separately; firstly, on the dlPFC group, with aforementioned index-type and stimulation-type as within-subjects factors. A significant main effect of index-type ($F_{(1,8)}$ = 82.060, p = 0.000018, η^2 = .911), and a significant two-way interaction between index-type and stimulation-type ($F_{(2,16)}$ = 3.984, p = 0.039, η^2 = 0.332) were found. A further repeated-measures ANOVA on R, with stimulation type as the within-subjects factor, showed that although the main effect of stimulation was marginally significant ($F_{(2,16)}$ = 4.432, p = 0.055, η^2 = 0.356) with Greenhouse-Geisser correction due to a significant Mauchly's test of sphericity, the tests of within-subjects contrasts showed a significant linear effect of stimulation type ($F_{(1,8)}$ = 21.547, p = 0.002, η^2 = 0.729) which prompted further scrutiny. Post-hoc pairwise t-tests with bonferroni correction revealed a significant suppression of R under beta stimulation compared with R under no-stimulation (p = 0.005). Another repeated-measures ANOVA on F, with stimulation type as the within-subjects factor, showed the main effect of stimulation type was not significant ($F_{(2,16)}$ = 2.380, p = 0.125, η^2 = 0.229).

Secondly, another repeated-measures ANOVA was conducted on dmPFC group, with aforementioned index-type and stimulation-type as within-subjects factors. A significant

main effect of index-type ($F_{(1,8)}$ = 15.964, p = 0.004, η^2 = 0.666), a significant main effect of stimulation-type ($F_{(2,16)}$ = 5.430, p = 0.016, η^2 = 0.404), and a significant two-way interaction between index-type and stimulation-type ($F_{(2,16)}$ = 5.326, p = 0.017, η^2 = 0.400) were found. A further repeated-measures ANOVA on R, with stimulation-type as the within-subjects factor, showed the main effect of stimulation type was not significant ($F_{(2,16)}$ = 2.185, p = 0.145, η^2 = 0.215). A further repeated-measures ANOVA on F, with stimulation-type as the within-subjects factor, showed that a significant main effect of stimulation was found ($F_{(2,16)}$ = 6.144, p = 0.010, η^2 = 0.434). Post-hoc pairwise t-tests with bonferroni correction showed a significant enhancement of F under beta stimulation compared with F under no-stimulation (p = 0.012).

Finally, the same analyses were conducted on the vertex group, with accuracy-type and stimulation-type as within-subject factors. A significant main effect of index-type was found ($F_{(1,8)}$ = 22.931, p = 0.001, η^2 = 0.741), but neither the main effect of stimulation-type nor the two-way interaction between index-type and stimulation-type were significant (main effect of stimulation-type: $F_{(2,16)}$ = 0.160, p = 0.854, η^2 = 0.020; two-way interaction: $F_{(2,16)}$ = 0.094, p = 0.911, η^2 = 0.012). Separate repeated-measures ANOVAs on R and F showed no significant main effects of stimulation type (main effect of R: $F_{(2,16)}$ = 0.691, p = 0.516, η^2 = 0.079; main effect of F: $F_{(2,16)}$ = 0.087, p = 0.917, η^2 = 0.011).

------ Insert Figure 4 about here ------

Discussion

In this study, we first investigated in the macaque what changes in frequency-specific LFP power occurred, and when, in dlPFC during our object recognition memory task. We observed a significantly increased beta power (12.5-24 Hz) in dlPFC during the sample

encoding phase of the task, indicating that beta oscillations in primate dlPFC may play a key role in encoding visual object stimuli. Then, to test causality and generalizability to humans, we targeted rTMS stimulation at a beta frequency (12.5 Hz) to the homologous brain region (dlPFC) in humans while they performed a similar task and while some samples were presented. We also targeted the same rTMS stimulation to both an active comparison site (dmPFC) and a passive control site (vertex, which we had no reason to think was implicated in recognition memory). As our hypothesis was that preventing beta desynchronization would be deleterious to successful encoding we also employed a control frequency (random-frequency rTMS) stimulation unlikely to prevent de-synchoronization. We found a functional double dissociation between the effects of short bursts of beta frequency rTMS stimulation to left dlPFC and left dmPFC; only 12.5 Hz rTMS to dlPFC during sample presentation reduced behavioural indices of recollection whereas only 12.5 Hz rTMS to dmPFC enhanced behavioural indices of familiarity. Short bursts of random frequency rTMS (i.e., with same total number of pulses over same total duration, but with each burst occurring pseudorandomly in time) had no effect on either recollection or familiarity over both prefrontal regions, nor over the vertex. Hence our hypothesis as to the behavioural effect of rTMS beta-frequency stimulation in humans, itself based on our area-specific, frequency-specific, and task-epoch specific macaque LFP observations and on the neural de-synchronization hypothesis, was supported. Importantly, our study helps further bridge the species divide and confirms that human and NHP PFC show similarity in function and in underlying neural mechanisms supporting recognition memory performance.

The neural de-synchronization hypothesis, based on information theory, was proposed (Simon Hanslmayr et al., 2012) to account for a *decrease* of alpha/beta LFP power in frontal and parietal cortex correlating with *successful* memory encoding and retrieval processes across some human studies (Hanslmayr et al., 2011, 2009; Khader and Rösler, 2011; Waldhauser et al., 2012). Specifically, power decreases in the alpha/beta frequency band are

presumed to reflect a de-synchronization of local neural assembles. Based on information theory, synchronization of neural firings reduces the richness of information transfer which results in information redundancy. In contrast, a de-synchronization of neural firing patterns leads to more potential for carrying information to improve the efficiency of neural communications (Simon Hanslmayr et al., 2012). This de-synchronization hypothesis, applied to human recognition memory, maintains that a decrease of alpha/beta power is associated with more efficient encoding, which helps contribute to successful memory retrieval subsequently (Simon Hanslmayr et al., 2012). In our study, stimulation over dlPFC in a beta frequency (12.5 Hz) using rTMS, at the times some specific sample object stimuli were viewed, induced a reduction in subsequent recollection of those specific samples. In another combined EEG-TMS study, stimulation over left dlPFC (BA 9) using rTMS at a particular beta frequency (18.7 Hz), but not at other frequencies (6.8 Hz and 10.7 Hz), also impaired recall processes (Hanslmayr et al., 2014). Our TMS study and the study from Hanslmayr et al. (2014) both support the de-synchronization hypothesis which suggests that intrinsic beta oscillations in dlPFC and their dynamic synchronization and desynchronization during memory formation, if impeded will impair recognition memory; hence our external source of beta frequency entrainment from rTMS over the same region may have impaired recognition memory performance by lessening natural processes of de-synchronization as this hypothesis predicts. Holmes et al. (2018) also showed a dissociation between LFP beta power and mnemonic tuning and whilst in a spatial task this is also consistent with de-synchronization aiding memory stability (Holmes et al., 2018). Neural attractor network modeling (Compte et al., 2003), in the context of delayed response task performance in NHPs, also suggests mechanisms by which a drop in LFP power may relate to memory processing; asynchronous networks can maintain memory coding for longer durations whereas synchronization leads to an overall instabilities incompatible with stable memory-

coding states. According to these models too, frontal memory networks desynchronize to preserve memory states.

The neural communication through coherence hypothesis (Fries, 2005) suggests that synchronized neural oscillations within a particular brain region or between distant brain areas provides important mechanisms for mediating cognition. In a review focusing on frontal-temporal interactions in recollection/familiarity, a "multi-effect multi-nuclei" model was proposed, in which inputs to PFC directly from hippocampus, and indirectly from hippocampus *via* the anterior and midline thalamic nuclei, are related to recollection process, while connections to PFC from perirhinal cortex and mediodorsal thalamic nucleus are related to familiarity process (Aggleton et al., 2011). Indeed, a three-circuit model of neural oscillations between PFC and MTL *via* thalamus has been proposed (Ketz et al., 2015). The first theta-circuit model emphasized interaction between PFC and hippocampus through theta oscillations (4-8 Hz) to support core functions of the hippocampus in successful memory encoding and retrieval processes (Ketz et al., 2015). The human neuropsychology and neuroimaging literature highlight the involvement of the hippocampus (Bowles et al., 2010; Eichenbaum et al., 2007; Ranganath et al., 2004; Skinner and Fernandes, 2007; Yonelinas et al., 2002) and PFC (Wheeler and Stuss, 2003) in recollection, hence theta oscillations within a PFC-hippocampus network may be essential for recollection process. In our TMS study then, the external source of beta frequency from rTMS over dlPFC may disrupt the communication between PFC and hippocampus mediated by theta-oscillation, which may be a plausible mechanism by which our rTMS effected a decrease of recollection process. The theta-circuit model of dlPFC and hippocampus interactions in recollection processes will require further investigations and will shed light on the systems-wide oscillatory mechanisms underlying recognition memory. Recording simultaneously from multiple areas and multiple cortical layers in NHPs will further elucidate these issues.

Other studies have also applied the information theory by de-synchronization hypothesis to investigate voluntary memory suppression. A combined fMRI-EEG-TMS study has shown that an increased BOLD signal in left dlPFC (BA 9) was associated with decreased neural synchrony at beta frequency (11-18 Hz) in a directed forgetting task; moreover, stimulation over dlPFC using a slow train of 1 Hz rTMS pulses in the same task boosted cued forgetting-behaviour and induced a further reduction in neural synchrony at 13 Hz (S. Hanslmayr et al., 2012). This study also suggests a causal role for beta desynchrony in dlPFC, this time in a directed forgetting paradigm and also suggests a key role of dlPFC in memory control operates *via* downregulating neural synchrony at beta frequency (S. Hanslmayr et al., 2012). Another fMRI study has shown that voluntary memory suppression was associated with an increased activation of dlPFC and a decreased activation of hippocampus (Anderson et al., 2004), which suggests an interaction between dlPFC and hippocampus in regulating memory suppression.

In our study, familiarity was enhanced in association with rTMS delivered at beta frequency over dmPFC. This result is in line with human neuroimaging meta-analyses suggesting that more activation peaks related to familiarity processing in PFC were found in dmPFC (BA area 8 and 9) (Scalici et al., 2017). Our study not only shows a causal involvement of dmPFC in familiarity-based recognition, but suggests that the mechanism driving familiarity may operate through a mechanism involving neural synchrony at beta frequency. A thalamo-cortical circuit (involving entorhinal, parahippocampal, and perirhinal cortical areas connected via the medial dorsal thalamic nucleus to PFC) may synchronise within the beta frequency range and it has been proposed that this beta-circuit model may underlie familiarity judgment processes (Ketz et al., 2015). Indeed, the functional role of rhinal cortex, especially perirhinal cortex, has long been associated more with familiarity-based recognition than recollection (Aggleton and Brown, 1999; Bowles et al., 2010; Norman and O'Reilly, 2003; Ranganath et al., 2004; Skinner and Fernandes, 2007). TMS was a viable

methodology in our study due to dlPFC and dmPFC being accessible, but this approach is not viable for deep structures in the medial temporal lobe. Future work in NHPs will of course be able to target microstimulation in via the same electrodes that record LFPs in the medial temporal lobe and record, simultaneously, the effects on multiple other brain regions.

The current project has helped bridge the species divide in understanding the neural mechanisms supporting object recognition memory between macaques and humans. The site of stimulation and the frequency we used for intervention in our human rTMS study of object recognition memory was chosen based on our observations of induced LFP power at beta frequency in dlPFC during the sample (encoding) phase of a similar recognition memory task for the macaque. Moreover, in our rTMS study, we found evidence for dissociable functional roles of beta frequency oscillations in different prefrontal regions; the rTMS study provides causal evidence for a functional double dissociation between left dorsolateral and left dorsomedial prefrontal cortex in recollection and familiarity and the involvement of different patterns of underlying beta frequency oscillations in each case.

Reference

Aggleton JP, Brown MW (1999) Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. Behav Brain Sci 22:425–444.

Aggleton JP, Dumont JR, Warburton EC (2011) Unraveling the contributions of the diencephalon to recognition memory: A review. Learn Mem 18:384–400.

Albouy P, Weiss A, Baillet S, Zatorre RJ (2017) Selective Entrainment of Theta Oscillations in the Dorsal Stream Causally Enhances Auditory Working Memory Performance. Neuron 1–14.

Aly M, Yonelinas AP, Kishiyama MM, Knight RT (2011) Damage to the lateral prefrontal cortex impairs familiarity but not recollection. Behav Brain Res 225:297–304.

Anderson MC, Ochsner KN, Kuhl B, Cooper J, Robertson E, Gabrieli SW, Glover GH, Gabrieli JDE (2004) Neural Systems Underlying the Suppression of Unwanted Memories. Science (80-) 303:232–235.

Bachevalier J, Mishkin M (1986) Visual recognition impairment follows ventromedial but not dorsolateral prefrontal lesions in monkeys. Behav Brain Res 20:249–261.

Basile BM, Hampton RR (2013) Recognition errors suggest fast familiarity and slow recollection in rhesus monkeys. Learn Mem 20:431–437.

Bastos AM, Loonis R, Kornblith S, Lundqvist M, Miller EK (2018) Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. Proc Natl Acad Sci 201710323.

Beam W, Borckardt JJ, Reeves ST, George MS (2009) An efficient and accurate new method for locating the F3 position for prefrontal TMS applications. Brain Stimul 2:50–54.

Benn A, Barker GRI, Stuart SA, Roloff EVL, Teschemacher AG, Clea Warburton E, Robinson ESJ (2016) Optogenetic stimulation of prefrontal glutamatergic neurons enhances recognition memory. J Neurosci 36:4930–4939.

Bowles B, Crupi C, Pigott S, Parrent A, Wiebe S, Janzen L, Köhler S (2010) Double

dissociation of selective recollection and familiarity impairments following two different surgical treatments for temporal-lobe epilepsy. Neuropsychologia 48:2640–2647.

Buzsáki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents-EEG, ECoG, LFP and spikes. Nat Rev Neurosci 13:407–420.

Buzsáki G, Schomburg EW (2015) What does gamma coherence tell us about inter-regional neural communication? Nat Neurosci 18:484–489.

Chanes L, Quentin R, Tallon-Baudry C, Valero-Cabre A (2013) Causal Frequency-Specific Contributions of Frontal Spatiotemporal Patterns Induced by Non-Invasive Neurostimulation to Human Visual Performance. J Neurosci 33:5000–5005.

Compte A, Constantinidis C, Tegnér J, Raghavachari S, Chafee M V., Goldman-Rakic PS, Wang XJ (2003) Temporally Irregular Mnemonic Persistent Activity in Prefrontal Neurons of Monkeys during a Delayed Response Task. J Neurophysiol 90:3441–3454.

Duarte A, Ranganath C, Knight RT (2005) Effects of unilateral prefrontal lesions on familiarity, recollection, and source memory. J Neurosci 25:8333–8337.

Eichenbaum H, Yonelinas AP, Ranganath C (2007) The Medial Temporal Lobe and Recognition Memory.

Elkin-Frankston S, Fried P, Rushmore RJ, Valero-Cabré A (2011) From qualia to quantia: A system to document and quantify phosphene percepts elicited by non-invasive neurostimulation of the human occipital cortex. J Neurosci Methods 198:149–157.

Fell J, Axmacher N (2011) The role of phase synchronization in memory processes. Nat Rev Neurosci 12:105–118.

Fries P (2015) Rhythms for Cognition: Communication through Coherence. Neuron 88:220–235.

Fries P (2005) A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. Trends Cogn Sci 9:474–480.

Fuster JM, Alexander GE (1971) Neuron Activity Related to Short-Term Memory. Science

(80-) 173:652–654.

Gregoriou GG, Gotts SJ, Zhou H, Desimone R (2009) High-Frequency, Long-Range Coupling Between Prefrontal and Visual Cortex During Attention. Science (80-) 324:1207–1210.

Hanslmayr S, Matuschek J, Fellner MC (2014) Entrainment of prefrontal beta oscillations induces an endogenous echo and impairs memory formation. Curr Biol 24:904–909.

Hanslmayr S, Spitzer B, Bäuml KH (2009) Brain oscillations dissociate between semantic and nonsemantic encoding of episodic memories. Cereb Cortex 19:1631–1640.

Hanslmayr Simon, Staudigl T, Fellner M-C (2012) Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. Front Hum Neurosci 6:1–12.

Hanslmayr S., Volberg G, Wimber M, Oehler N, Staudigl T, Hartmann T, Raabe M, Greenlee MW, Bauml K-HT (2012) Prefrontally Driven Downregulation of Neural Synchrony Mediates Goal-Directed Forgetting. J Neurosci 32:14742–14751.

Hanslmayr S, Volberg G, Wimber M, Raabe M, Greenlee MW, Bauml K-HT (2011) The Relationship between Brain Oscillations and BOLD Signal during Memory Formation: A Combined EEG-fMRI Study. J Neurosci 31:15674–15680.

Henson Richard N.A., Rugg MD, Shallice T, Josephs O, Dolan RJ (1999) Recollection and familiarity in recognition memory: An event-related functional magnetic resonance imaging study. J Neurosci 19:3962–3972.

Henson R. N.A., Shallice T, Dolan RJ (1999) Right prefrontal cortex and episodic memory retrieval: A functional MRI test of the monitoring hypothesis. Brain 122:1367–1381.

Herweg NA, Apitz T, Leicht G, Mulert C, Fuentemilla L, Bunzeck N (2016) Theta-alpha oscillations bind the hippocampus, prefrontal cortex, and striatum during recollection: Evidence from simultaneous EEG–fMRI. J Neurosci 36:3579–3587.

Holmes CD, Papadimitriou C, Snyder LH (2018) Dissociation of LFP power and tuning in

the frontal cortex during memory. J Neurosci 38:8177–8186.

Horner AJ, Bisby JA, Bush D, Lin WJ, Burgess N (2015) Evidence for holistic episodic recollection via hippocampal pattern completion. Nat Commun 6:1–11.

Huang YZ, Edwards MJ, Rounis E, Bhatia KP, Rothwell JC (2005) Theta burst stimulation of the human motor cortex. Neuron 45:201–206.

Johnson JD, Suzuki M, Rugg MD (2013) Recollection, familiarity, and content-sensitivity in lateral parietal cortex: a high-resolution fMRI study. Front Hum Neurosci 7:1–15.

Kafkas A, Montaldi D (2012) Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. Neuropsychologia 50:3080–3093.

Ketz NA, Jensen O, Reilly RCO (2015) Thalamic pathways underlying prefrontal cortex – medial temporal lobe oscillatory interactions. Trends Neurosci 38:3–12.

Khader PH, Rösler F (2011) EEG power changes reflect distinct mechanisms during long-term memory retrieval. Psychophysiology 48:362–369.

Koen JD, Barrett FS, Harlow IM, Yonelinas AP (2017) The ROC Toolbox: A toolbox for analyzing receiver-operating characteristics derived from confidence ratings. Behav Res Methods 49:1399–1406.

Kornblith S, Buschman TJ, Miller EK (2016) Stimulus load and oscillatory activity in higher cortex. Cereb Cortex 26:3772–3784.

Köster M, Friese U, Schöne B, Trujillo-Barreto N, Gruber T (2014) Theta-gamma coupling during episodic retrieval in the human EEG. Brain Res 1577:57–68.

Kowalska DM, Bachevalier J, Mishkin M (1991) The role of the inferior prefrontal convexity in performance of delayed nonmatching-to-sample. Neuropsychologia 29:583–600.

Levy R, Goldman-Rakic PS (1999) Association of storage and processing functions in the dorsolateral prefrontal cortex of the nonhuman primate. J Neurosci 19:5149–5158.

Lundqvist M, Rose J, Herman P, Brincat SL, Buschman TJ, Miller EK (2016) Gamma and

Beta Bursts Underlie Working Memory. Neuron 90:152–164.

MacPherson SE, Bozzali M, Cipolotti L, Dolan RJ, Rees JH, Shallice T (2008) Effect of frontal lobe lesions on the recollection and familiarity components of recognition memory. Neuropsychologia 46:3124–3132.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190.

Meunier M, Bachevalier J, Mishkin M (1997) Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. Neuropsychologia 35:999–1015.

Mishkin M, Manning FJ (1978) Non-spatial memory after selective prefrontal lesions in monkeys. Brain Res 143:313–323.

Norman KA, O'Reilly RC (2003) Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. Psychol Rev 110:611–646.

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011.

Parker A (1998) Interaction of frontal and perirhinal cortices in visual object recognition memory in monkeys. Eur J Neurosci 10:3044–3057.

Petrides M (2000) Dissociable roles of mid-dorsolateral prefrontal and anterior inferotemporal cortex visual working memory. J Neurosci 20:7496–7503.

Petrides M (1995) Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. J Neurosci 15:359–375.

Pipa G (2009) Performance- and stimulus-dependent oscillations in monkey prefrontal cortex during short-term memory. Front Integr Neurosci 3:1–7.

Ranganath C, Yonelinas AP, Cohen MX, Dy CJ, Tom SM, D'Esposito M (2004) Dissociable correlates of recollection and familiarity within the medial temporal lobes. Neuropsychologia 42:2–13.

Rao SC, Rainer G, Miller EK (1997) Integration of what and where in the primate prefrontal cortex. Science (80-) 276:821–824.

Rossi S et al. (2009) Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clin Neurophysiol 120:2008–2039.

Rossi S, Rossini PM (2004) TMS in cognitive plasticity and the potential for rehabilitation. Trends Cogn Sci 8:273–279.

Rossini PM, Barker AT, Berardelli A, Caramia MD, Caruso G, Cracco RQ, Dimitrijević MR, Hallett M, Katayama Y, Lücking CH, Maertens de Noordhout AL, Marsden CD, Murray NMF, Rothwell JC, Swash M, Tomberg C (1994) Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. Report of an IFCN committee. Electroencephalogr Clin Neurophysiol 91:79–92.

Scalici F, Caltagirone C, Carlesimo GA (2017) The contribution of different prefrontal cortex regions to recollection and familiarity. A review of fMRI data. Neurosci Biobehav Rev 83:240–251.

Skinner EI, Fernandes MA (2007) Neural correlates of recollection and familiarity: A review of neuroimaging and patient data. Neuropsychologia 45:2163–2179.

Solstad T, Moser EI, Einevoll GT (2006) RAPID COMMUNICATION From Grid Cells to Place Cells: A Mathematical Model. Hippocampus 1031:1026–1031.

Stokes MG (2015) "Activity-silent" working memory in prefrontal cortex: A dynamic coding framework. Trends Cogn Sci 19:394–405.

Stokes MG (2005) Simple Metric For Scaling Motor Threshold Based on Scalp-Cortex

Distance: Application to Studies Using Transcranial Magnetic Stimulation. J Neurophysiol 94:4520–4527.

Stokes MG, Barker AT, Dervinis M, Verbruggen F, Maizey L, Adams RC, Chambers CD (2013) Biophysical determinants of transcranial magnetic stimulation: Effects of excitability and depth of targeted area. J Neurophysiol 109:437–444.

Stokes MG, Chambers CD, Gould IC, English T, McNaught E, McDonald O, Mattingley JB (2007) Distance-adjusted motor threshold for transcranial magnetic stimulation. Clin Neurophysiol 118:1617–1625.

Thut G, Veniero D, Romei V, Miniussi C, Schyns P, Gross J (2011) Rhythmic TMS causes local entrainment of natural oscillatory signatures. Curr Biol 21:1176–1185.

Tulving E, Stuss DT, Tulving E (1995) Frontal Lobe Damage Produces Episodic Memory Impairment. J Int Neuropsychol Soc 1:525–536.

Waldhauser GT, Johansson M, Hanslmayr S (2012) Alpha/beta oscillations indicate inhibition of interfering visual memories. J Neurosci 32:1953–1961.

Wheeler MA, Stuss DT (2003) Remembering and knowing in patients with frontal lobe injuries. Cortex 39:827–846.

Wilson FAW, Scalaidhe SP., Goldman-Rakic PS (1993) Dissociation of object and spatial processing domains in primate prefrontal cortex. Science (80-) 260:1955–1958.

Wixted JT (2007) Dual-process theory and signal-detection theory of recognition memory. Psychol Rev 114:152–176.

Womelsdorf T, Schoffelen JM, Oostenveld R, Singer W, Desimone R, Engel AK, Fries P (2007) Modulation of neuronal interactions through neuronal synchronization. Science (80-) 316:1609–1612.

Yonelinas AP (2002) The nature of recollection and familiarity: A review of 30 years of research. J Mem Lang 46:441–517.

Yonelinas AP (2001) Components of episodic memory: The contribution of recollection and

familiarity. Philos Trans R Soc B Biol Sci 356:1363–1374.

Yonelinas AP (1994) Receiver-Operating Characteristics in Recognition Memory: Evidence for a Dual-Process Model. J Exp Psychol Learn Mem Cogn 20:1341–1354.

Yonelinas AP, Kroll NEA, Dobbins I, Lazzara M, Knight RT (1998) Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. Neuropsychology 12:323–339.

Yonelinas AP, Kroll NEA, Quamme JR, Lazzara MM, Sauvé MJ, Widaman KF, Knight RT (2002) Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. Nat Neurosci 5:1236–1241.

Yonelinas AP, Parks CM (2007) Receiver Operating Characteristics (ROCs) in Recognition Memory: A Review. Psychol Bull 133:800–832.

Figure and Figure legends

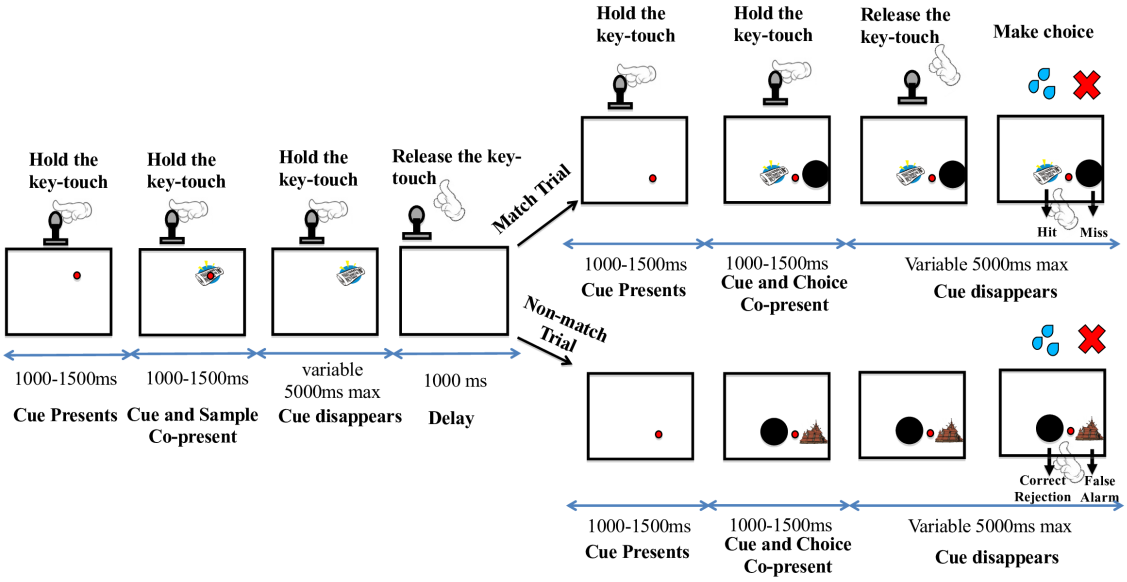

Figure 1. NHP recognition memory task structure. *The figure depicts one trial in the electrophysiological study. The NHP initiated the task by holding the key-touch device once the red dot cue appeared through a variable time of 1000-1500 ms, after which a sample object image was shown on the screen behind the red dot cue (encoding phase). The NHP was required to keep holding the keytouch device through another variable delay of 1000-1500 after which the red keytouch cue disappeared then the NHP was trained to release its hand from the keytouch prompting the commence of the delay (1000ms) between sample and choice-phases. Then after the 1000 ms delay period another key-touch red dot cue, this time in the bottom of the screen, and the animal was required to hold the key-touch again. The NHP was again required to keep holding the key-touch device through a variable delay of 1000-1500 ms after which two choice stimuli (one an object image and the other a black circle, left-right randomized between trials) appeared. The NHP was required to keep holding the keytouch device through another variable delay of 1000-1500 ms after which the key-touch red dot cue disappeared cueing the animal that it could release holding the key-touch and make a choice to the touchscreen to either the object test-image stimulus or to the black circle. Just as in the human task in Figure 2, in 'Match Trials' the black circle was*

presented with an identical image to one of the preceding sample; in 'Non-match Trials' the black circle was presented with a novel stimulus not seen before. Animals were trained to touch the test-item if they remembered the test-item was a match but to touch the black circle if they thought the test-item was a non-match; accordingly the responses could be separated into hits, misses, correct rejections, and false alarms as indicted.

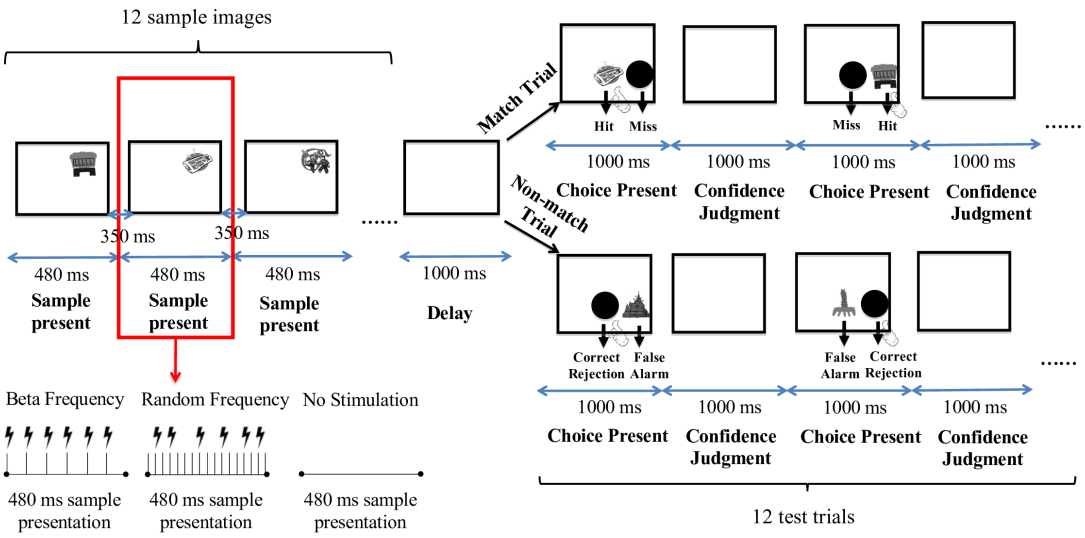

Figure 2. Human recognition memory task structure. *The figure depicts one block in the TMS study. Twelve sample object images were sequentially presented for 480 ms with an inter-stimuli-interval of 350 ms (the figure shows the procedure for one of these 12). Six rTMS pulses (at beta frequency or at random frequency) or no stimulation were given during the 480ms sample presentation for 4 out of the 12 samples in the list. Then after a delay of 1000 ms, 12 test trials followed (the figure shows the procedure for one of these 12). Just as in the NHP task in Figure 1, in 'Match Trials' the black circle was presented with an identical image to one of the preceding sample; in 'Non-match Trials' the black circle was presented with a novel stimulus not seen before. Participants were told to touch the test-item if they remembered the test-item was a match but to touch the black circle if they thought the test-item was a non-match; accordingly the responses could be separated into hits, misses, correct rejections, and false alarms as indicted. Each test-trial in this human version of the task ended with a confidence judgment indicated by the participant by their making a finger gesture as described in the main text.*

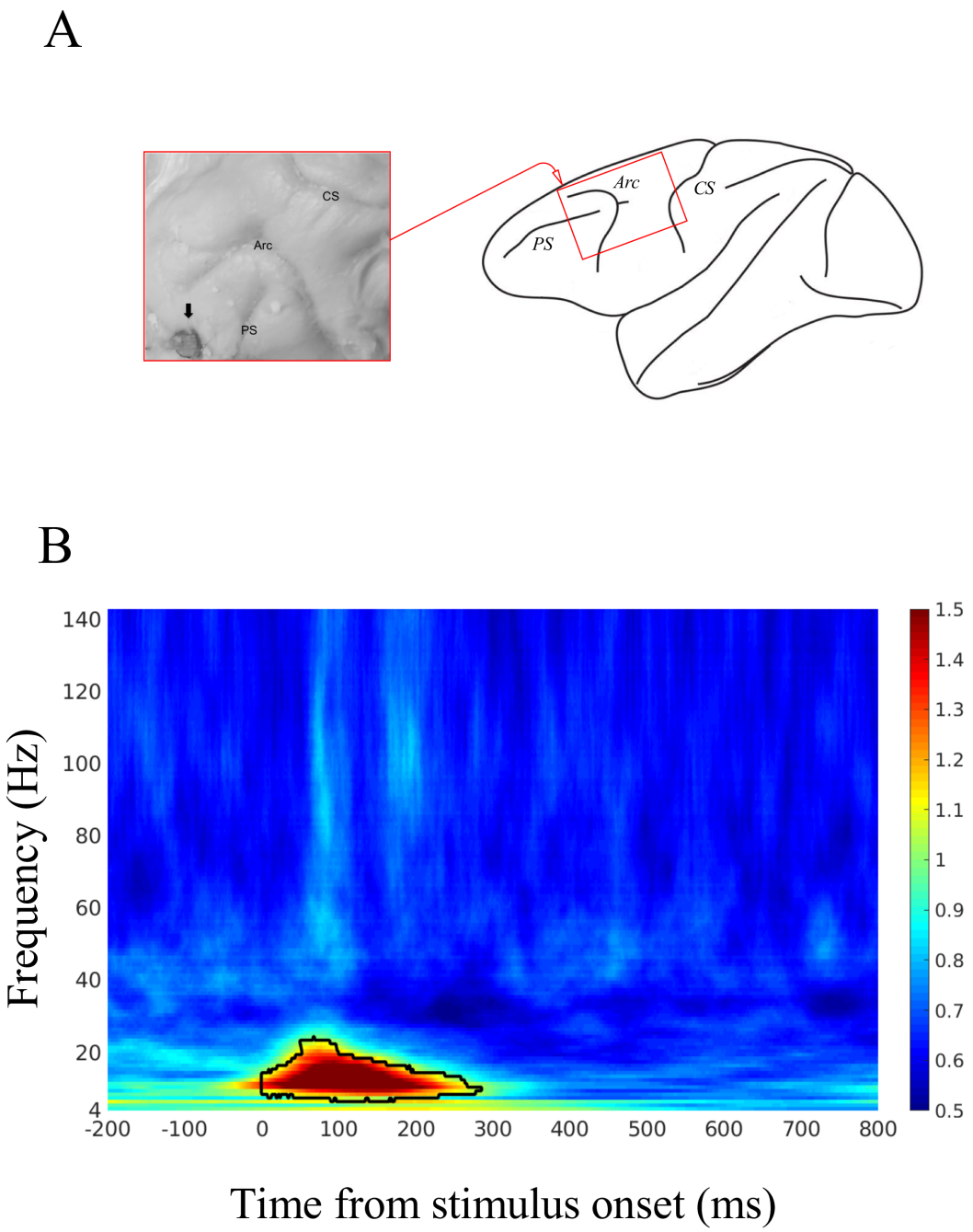

***Figure 3*. Recording sites and time-frequency spectrogram for induced local field potential (LFP) power from dlPFC array in the sample presentation epoch in NHP in Expt.1.** *Panel A: drawing of a lateral view of the NHP brain; the Utah array was placed in the region of the red square and the enlarged insert shows photographic verification that the array was positioned in the cortex above the principle sulcus (PS) and anterior to the arcuate sulcus (Arc). Panel B: The area of statistically significant differences of LFPs between pre- and*

post-sample periods in macaque dlPFC is highlighted by a black rectangle and includes beta frequencies for around 300ms from sample onset (indicated by zero ms on this figure). The data depicted in this figure is averaged across 32 electrodes in dlPFC, and across 29 daily sessions, and across approx. 75 trials per session.

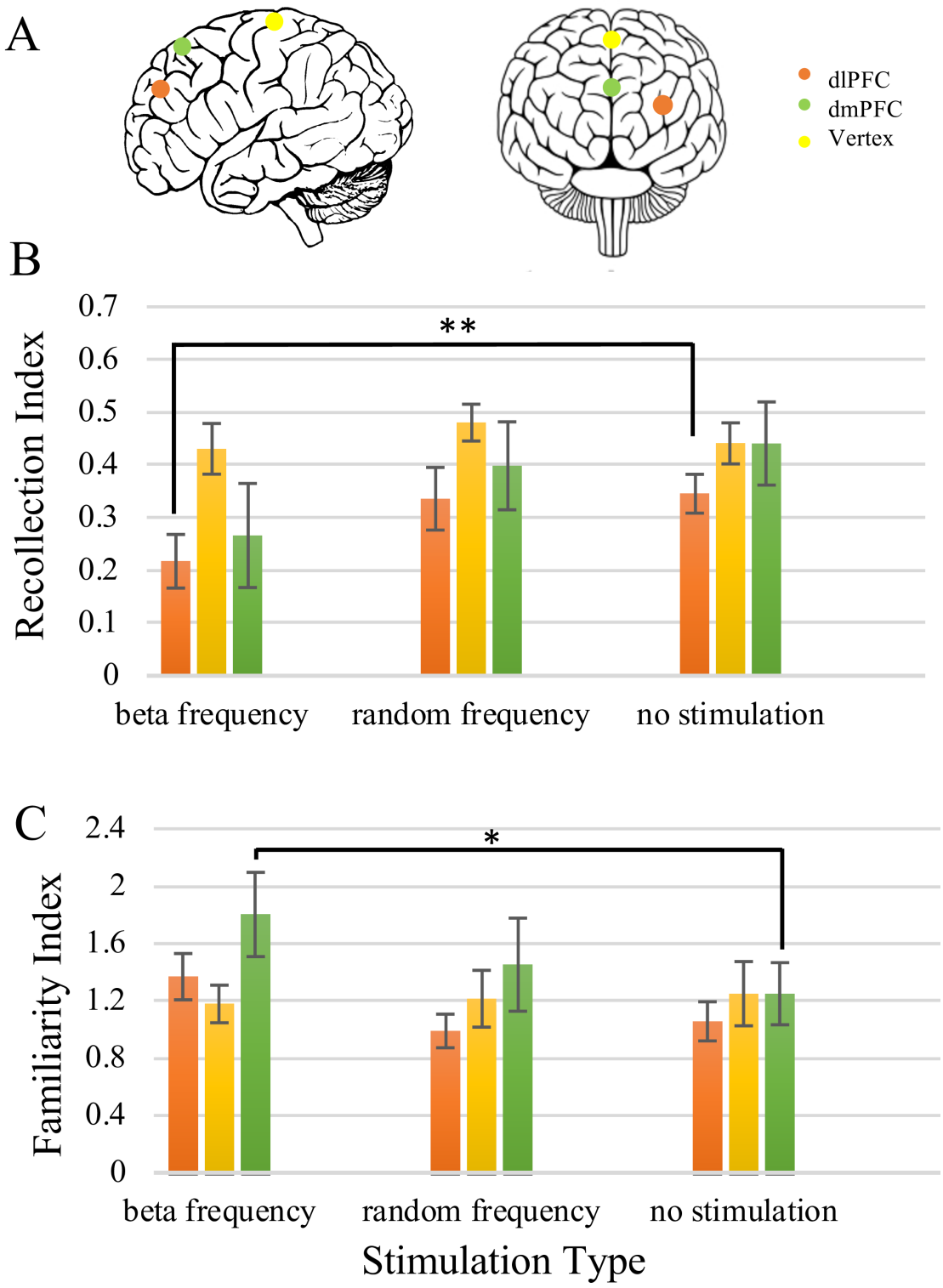

Figure 4. Stimulation sites for each group of human participants and effects of TMS upon recollection and familiarity indices in Expt. 2. *Panel A: lateral and frontal view of the human brain. Human stimulation sites for rTMS are shown as dlPFC (orange), dmPFC (green) and vertex (yellow). Panel B-C: Bar graphs presenting the mean recollection index*

*(Panel B) and familiarity index (Panel C) in dlPFC (orange bar), dmPFC (yellow bar) and vertex (green bar) under beta-frequency stimulation, random-frequency stimulation and no stimulation. Note that beta-frequency stimulation (but neither random-frequency stimulation nor no stimulation) suppressed recollection index in dlPFC. * indicates $p < 0,05$, and ** indicates $p < 0.01$.*

Table and Table legends

Table 1. Descriptive statistics: accuracy (AUCs) and response time of correct responses for each participant group across each stimulation condition

Brain Area	Stimulation	Accuracy	Response Time (ms)
dlPFC	Beta frequency	0.85 ± 0.02	1726.62 ± 146.90
(n = 9)	Random frequency	0.84 ± 0.01	1741.79 ± 157.21
	No Stimulation	0.84 ± 0.02	1737.81 ± 153.63
dmPFC	Beta frequency	0.92 ± 0.02	1899.95 ± 112.99
(n = 9)	Random frequency	0.89 ± 0.03	1856.12 ± 98.14
	No Stimulation	0.88 ± 0.03	1836.22 ± 95.24
Vertex	Beta frequency	0.88 ± 0.02	1545.63 ± 66.20
(n = 9)	Random frequency	0.89 ± 0.02	1502.88 ± 72.49
	No Stimulation	0.88 ± 0.02	1537.93 ± 75.49

Legend: Accuracy and response time of correct responses for each participant group across each stimulation condition; neither accuracy nor response time differed significantly across different stimulation-sites and stimulation-types in each group of human participants.