Behavioral/Systems/Cognitive

Two Critical and Functionally Distinct Stages of Face and Body Perception

David Pitcher,[1,2] **Tanya Goldhaber,**[1] **Bradley Duchaine,**[3] **Vincent Walsh,**[2] **and Nancy Kanwisher** [1]

[1]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, [2]Institute of Cognitive Neuroscience, University College London, London, WC1N 3AR, United Kingdom, and [3]Department of Psychology and Brain Sciences, Dartmouth College, Hanover, New Hampshire 03755

Cortical regions that respond preferentially to particular object categories, such as faces and bodies, are essential for visual perception of these object categories. However, precisely when these regions play a causal role in recognition of their preferred categories is unclear. Here we addressed this question using transcranial magnetic stimulation (TMS). Across a series of experiments, TMS was delivered over the functionally localized right occipital face area (rOFA) or right extrastriate body area (rEBA) at different latencies, up to 150 ms, after stimulus onset while adult human participants performed delayed match-to-sample tasks on face and body stimuli. Results showed that TMS disrupted task performance during two temporally distinct time periods after stimulus onset, the first at 40/50 ms and the second at 100/110 ms. These two time periods exhibited functionally distinct patterns of impairment: TMS delivered during the early time period (at 40/50 ms) disrupted task performance for both preferred (faces at rOFA and bodies at rEBA) and nonpreferred (bodies at rOFA and faces at rEBA) categories. In contrast, TMS delivered during the later time period (at 100/110 ms) disrupted task performance for the preferred category only of each area (faces at rOFA and bodies at rEBA). These results indicate that category-selective cortical regions are critical for two functionally distinct stages of visual object recognition: an early, presumably preparatory stage that is not category selective occurring almost immediately after stimulus onset, followed by a later stage of category-specific perceptual processing.

Introduction

Humans can accurately identify visually presented objects within a fraction of a second (Potter and Faulconer, 1975; Grill-Spector and Kanwisher, 2005; Crouzet et al., 2010), but the cortical mechanisms that enable this ability are not well understood. Neuropsychological patients with focal brain damage and neuroimaging studies of the undamaged brain demonstrate that recognition of particular stimulus categories, such as faces and bodies, depends on dissociable regions in occipitotemporal cortex (Puce et al., 1996; Kanwisher et al., 1997; Moscovitch et al., 1997; Rossion et al., 2003; Grill-Spector et al., 2004; Moro et al., 2008). However, these techniques do not possess the necessary temporal resolution to determine when category-selective visual regions first contribute to the cognitive mechanisms responsible for face and body perception. Here we address this question by asking at what latencies, up to 150 ms after stimulus onset, the occipital face area (OFA) (Gauthier et al., 2000) and extrastriate body area (EBA) (Downing et al., 2001) contribute to the perception of faces and bodies.

Experimental techniques with high temporal resolution, including scalp electroencephalography, magnetoencephalography (MEG), and intracranial recording in neuropsychological patients, report the earliest cortical response recorded to any visually presented stimulus occurs at latencies ranging from 50 to 100 ms after stimulus onset (Martínez et al., 1999; Furey et al., 2006; Yoshor et al., 2007). Other studies examining high-level visual object recognition report a selective response for specific object categories, including faces and bodies, beginning 100–200 ms after stimulus onset (Bentin et al., 1996; Thorpe et al., 1996; Eimer, 1998; Allison et al., 1999; Liu et al., 2002, 2009; Furey et al., 2006; Thierry et al., 2006; Engell and McCarthy, 2010). However, these results do not tell us when category-selective cortical regions causally contribute to recognition of stimuli from their preferred categories.

Transcranial magnetic stimulation (TMS) can be delivered with precise temporal resolution, and the induced task disruption demonstrates that the stimulated region is critically engaged when performing the concurrent behavioral task. In the present study, TMS was delivered over the right OFA (rOFA) and right EBA (rEBA), two cortical regions implicated in the perception of faces and bodies (Haxby et al., 2000; Peelen and Downing, 2007). Previous studies have demonstrated that TMS has the necessary spatial specificity to selectively disrupt face, but not body, processing when delivered over the rOFA and body, but not face, processing when delivered over the rEBA (Urgesi et al., 2004; Pitcher et al., 2009). Moreover, double-pulse TMS (dTMS) de-

Received May 30, 2012; revised Aug. 3, 2012; accepted Sept. 6, 2012.

Author contributions: D.P. and N.K. designed research; D.P. and T.G. performed research; D.P. analyzed data; D.P., B.D., V.W., and N.K. wrote the paper.

This work was supported by National Institutes of Health Grant EY13455 (N.K.) and Biotechnology and Biological Sciences Research Council Grant BB/F022875/1 (B.D. and V.W.). Thanks to Rebecca Lawson, Danny Dilks, Julie Golomb, and Ed Vul.

The authors declare no competing financial interests.

Correspondence should be addressed to David Pitcher, Laboratory of Brain and Cognition, National Institute for Mental Health, Bethesda, MD 20892. E-mail: david.pitcher@nih.gov.

T. Goldhaber's present address: Department of Engineering, University of Cambridge, Cambridge, UK.

DOI:10.1523/JNEUROSCI.2624-12.2012

Copyright © 2012 the authors 0270-6474/12/3215877-09$15.00/0

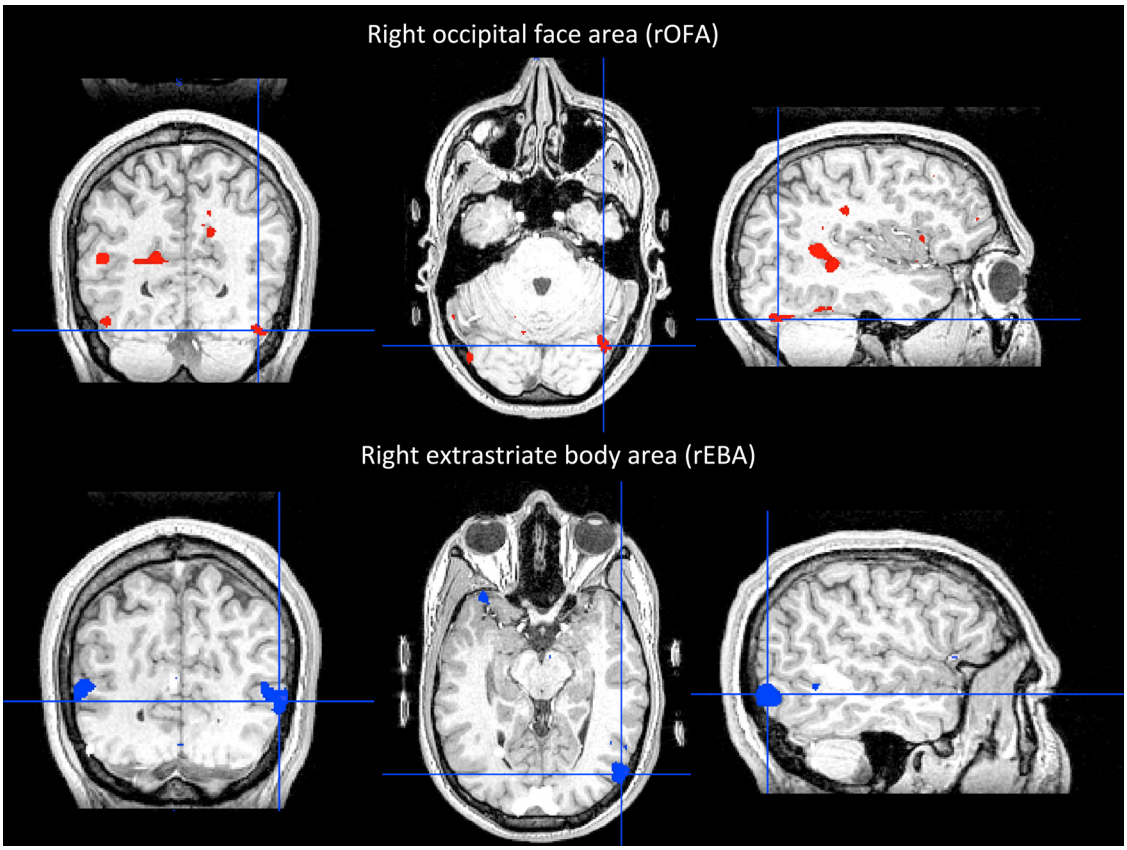

Figure 1. Locations in one participant of the rOFA in red (faces − objects) and the rEBA in blue (bodies − objects).

livered over the rOFA at 60 and 100 ms (but not at other latencies up to 290 ms) disrupted face processing (Pitcher et al., 2007, 2008), suggesting that the rOFA operated during a single discrete time window. In the present study, we sampled at a temporally finer grain by delivering pairs of TMS pulses, separated by 10 ms, over the functionally localized rOFA or rEBA at latencies ranging from 20 to 140 ms after stimulus onset while participants performed delayed match-to-sample tasks on face and body stimuli and an eye-gaze discrimination task.

Materials and Methods

Participants

Sixteen right-handed participants (7 males, 9 females, 18–34 years old) with normal or corrected-to-normal vision gave informed consent as directed by the Massachusetts Institute of Technology Institutional Review Board committee. Twelve participants participated in experiments 1 and 2. Ten participants from experiments 1 and 2 returned to participate in experiments 3–5. Eight participants participated in experiment 6, four of whom had taken part in experiments 1–5, four of whom had not taken part in experiments 1–5.

Brain imaging

An fMRI localizer using dynamic stimuli was used to individually identify the TMS target sites (rOFA and rEBA) in each participant (Pitcher et al., 2011a). Functional data were acquired over four blocked-design functional runs each lasting 234 s. Scanning was performed in a 3.0 T Siemens Trio scanner at the A. A. Martinos Imaging Center at the McGovern Institute for Brain Research at the Massachusetts Institute of Technology. Functional images were acquired with a Siemens 32-channel phased array head coil and a gradient-echo EPI sequence (32 slices; repetition time, 2 s; echo time, 30 ms; voxel size, 3 × 3 × 3 mm; and 0.6 mm interslice gap) providing whole-brain coverage (slices were aligned with the anterior/posterior commissure). In addition, a high-resolution T1-weighted MPRAGE anatomical scan was acquired for anatomically localizing the functional activations. Each functional run contained two sets of five consecutive dynamic stimulus blocks (faces, bodies, scenes, objects, or scrambled objects) sandwiched between rest blocks to make two blocks per stimulus category per run. Each block lasted 18 s and contained stimuli from one of the five stimulus categories.

Data were analyzed with FS-FAST, Freesurfer (http://surfer.nmr.mgh.harvard.edu/) (Dale et al., 1999; Fischl et al., 1999). Before statistical analysis, images were motion corrected (Cox and Jesmanowicz, 1999), smoothed (3 mm FWHM Gaussian kernel), detrended, and fit using a gamma function (delta = 2.25 and tau = 1.25). The preprocessing did not involve any spatial normalization of subjects in a common reference space (e.g., transformations into Talairach or MNI space). The functional data of each subject were coregistered with that subject's anatomical image.

Significance maps of the brain were computed using the same statistical threshold for both TMS target sites ($p = 10^{-4}$, uncorrected). The rOFA was identified using a contrast of dynamic faces greater than dynamic objects and was always located on the lateral surface of the occipital lobe (group mean MNI coordinates = 42, −79, −10). The rEBA was identified using a contrast of dynamic bodies greater than dynamic objects and was always located on the lateral surface of the occipital cortex and superior to the rOFA (group mean MNI coordinates = 51, −70, 2). The coordinates and strength of the peak responses varied across participants, but rOFA and rEBA were identified in each participant. The location of the rOFA and rEBA in a typical participant is shown in Figure 1.

TMS experiments

Experimental stimuli. All stimuli were presented centrally on an super video graphics array 20-inch monitor set at 1024 × 768 resolution. Computer-generated face and body stimuli were used to create a delayed match-to-sample discrimination task. FantaMorph software (Abrosoft; http://www.abrosoft.com/) was used to make a morph series between 10 pairs of faces and 10 pairs of bodies, as described next. This stimulus set

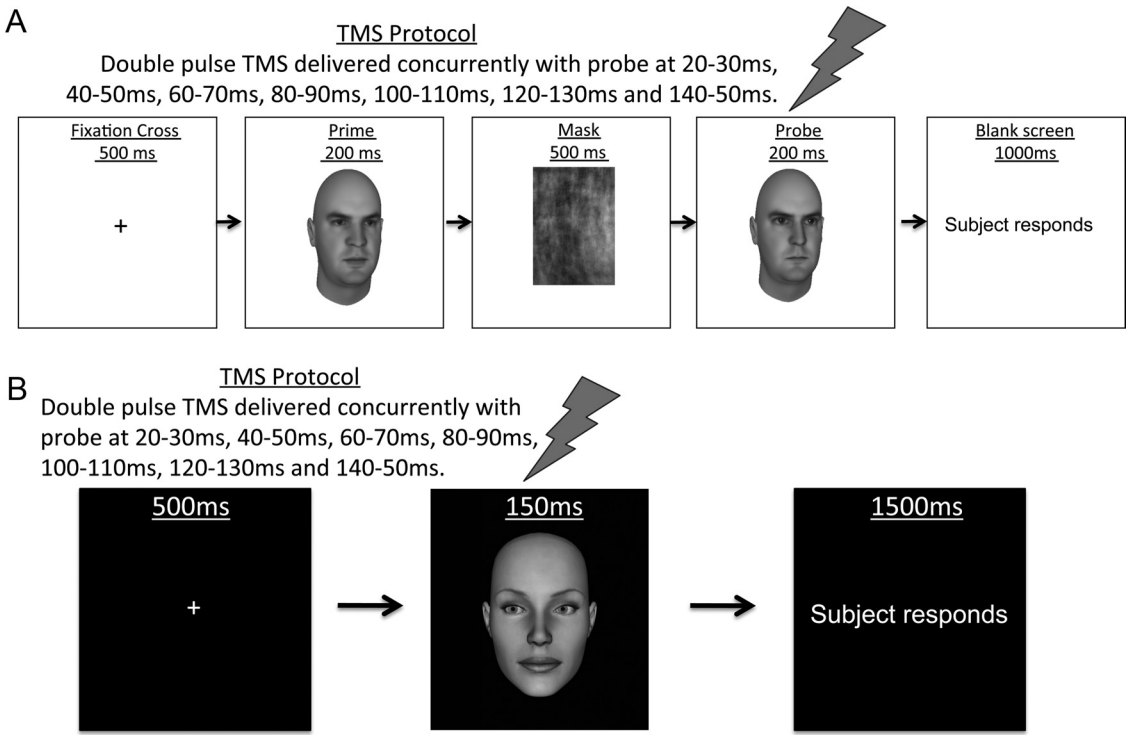

Figure 2. Timeline of the experimental trial procedure from experiments 1 and 6. In experiments 1–5, participants judged whether the probe stimulus was the same or different from the prime stimulus. In experiment 6, participants judged the gaze direction of the stimulus.

was used in a previous TMS study of the rOFA and rEBA (Pitcher et al., 2009).

Faces. Ten faces (varied in gender and viewing angle) were created using FaceGen software (Singular Inversions), and the component parts of these faces (eyes, mouth, and nose) were then individually altered to create a second face. Each face pair was then used to create a morph series. Each morph series was composed of 10 images with a 10% difference along the morph spectrum between each successive image. These morph–series images were then used to create 40 unique experimental trials (20 same, 20 different) comprising four trials per morph–series pair. For the different trials, the percentage morph difference between the two images was 80% (10 trials) or 100% (10 trials), based on pilot data indicating that these morph level distances would yield ~80% correct performance.

Bodies. Ten pairs of male bodies (varied in corpulence and muscle tone) wearing white shorts and seen from different viewing angles were created using Poser software (Smith Micro). Adobe Photoshop was used to remove the head. Body pose was the same for both images in each trial. For different trials, the percentage difference between the two images was 50% (five trials), 80% (10 trials), or 100% (five trials), again based on pilot data indicating that these morph level distances would yield ~80% correct performance.

Gaze faces. Eight faces (four male, four female) displaying five different gaze directions (right 10°, right 5°, direct gaze, left 5°, left 10°), used in experiment 6, were created using Poser 6 software.

TMS stimulation and site localization

dTMS, in which the two pulses were separated by 10 ms, was delivered at 60% of maximal stimulator output, using a Magstim Super Rapid Stimulator and a 70 mm figure-of-eight coil. Sixty percent of stimulator output using the Magstim Super Rapid system typically equates to ~100% of active motor threshold in the majority of participants. The coil was held with the handle pointing upward and parallel to the midline. A single intensity was used on the basis of previous TMS studies (O'Shea et al., 2004; Silvanto et al., 2005) and for ease of comparison with previous studies of the rOFA and rEBA (Pitcher et al., 2009, 2011b; Sadeh et al., 2011).

In experiments 1 and 2, test stimuli were presented while dTMS was delivered over the experimental site of interest (rOFA, rEBA, or vertex). dTMS was delivered at seven different time windows after stimulus onset: either at 20 and 30 ms, 40 and 50 ms, 60 and 70 ms, 80 and 90 ms, 100 and 110 ms, 120 and 130 ms, or 140 and 150 ms. These time windows were chosen to cover the most likely times that the rOFA and rEBA would be critically involved in the task based on previous TMS studies of the rOFA (Pitcher et al., 2007, 2008). To control for the nonspecific effects of TMS (both the somatic sensation on the scalp and the loud click that accompanies each pulse), we also included blocks in which dTMS was delivered over the vertex and blocks in which no TMS was delivered.

Experiments 1 and 2 found two latency windows at which TMS disrupted performance, in 10 of the 12 subjects. This finding was followed up in experiment 3, in which we retested 10 of the participants from experiments 1 and 2 who exhibited both the early and the late impairment windows. dTMS separated by 10 ms was delivered over the rOFA, rEBA, and vertex at three different TMS time windows: early, middle, and late. The delivery time from stimulus onset of each TMS time window was determined individually for each participant based on the results for that individual in experiments 1 and 2. For each subject, the early and late TMS time windows corresponded to the mean peak impairment latencies seen for face discrimination at rOFA in experiment 1 and body discrimination at rEBA in experiment 2 relative to the vertex control condition. The middle TMS time window was set at the time precisely between these two impairment windows and acted as a temporal condition in which dTMS was not predicted to impair task performance. The range of the three time windows across all participants was as follows: early window, 30–50 ms; middle window, 50–80 ms; and late window, 70–110 ms.

Experiments 4–6 used the same TMS stimulation parameters as experiments 1 and 2.

Procedure

In experiment 1, participants performed a delayed match-to-sample task using face stimuli (Fig. 2*a*). Participants were required to judge whether the prime stimulus was the same as the probe stimulus. Both the prime and probe stimuli were presented for 200 ms each. Subjects made key-

board responses using the right hand while seated with their heads stabilized on a chinrest. They were instructed to respond accurately and as quickly as possible.

There were 40 unique trials (20 same, 20 different) using 10 different faces that were repeated within each of the seven TMS latency conditions (20/30, 40/50, 60/70, 80/90, 100/110, 120/130, and 140/150 ms), producing a total of 280 trials per TMS site (rOFA and vertex). Trial order within each of the seven time windows was randomized. Each subject completed four TMS blocks consisting of 140 trials each (two rOFA and two vertex). Within each stimulation block, the dTMS delivery times were interleaved and randomized. Each subject also completed two no-TMS blocks consisting of 40 trials each, which acted as a baseline condition. Block order was alternated (TMS target site, TMS control site, no TMS) and balanced across subjects.

The scalp locations of the TMS target sites (rOFA and rEBA) were individually located in each participant using the Brainsight TMS–MRI coregistration system (Rogue Research). In each participant, the rOFA and rEBA were identified by overlaying individual activation maps for the faces greater than objects and bodies greater than objects contrasts from the fMRI localizer task for that participant. The TMS target site was marked as the voxel exhibiting the peak activation in the center of each ROI. The proper coil locations above these peak voxels were then marked on each subject's scalp, and the position of the TMS coil over the category-selective experimental sites was monitored over the course of the experiment using the online tracking software in the Brainsight system to ensure accurate coil placement over the TMS target site. The vertex control site, a point at the center of the top of the head, was defined as the point midway between the inion and the nasion and equidistant from the left and right intertragal notches.

Experiment 2 followed the same procedure as experiment 1, except that participants performed a delayed match-to-sample task on body stimuli while dTMS was delivered over rEBA and vertex. Subjects also completed two no-TMS blocks.

In experiment 3, subjects repeated the face task from experiment 1 and the body task from experiment 2 while dTMS was delivered over rOFA, rEBA, and vertex. A no-TMS block for each task was also included to act as additional control. dTMS separated by 10 ms was delivered in three different TMS time windows: early, middle, and late. Participants performed a total of eight blocks (four face blocks and four body blocks) while dTMS was delivered over rOFA, rEBA, and vertex or no TMS was delivered. Block order for task (face or body) and TMS site (rOFA, rEBA, vertex, no TMS) was alternated during each testing session and balanced across subjects. Within each block, the TMS delivery time (early, middle, or late) was randomized. During experiment 3, we closely monitored more than half of the participants to check that there was no systematic pattern of blinking during the task. We observed no evidence that TMS delivery was inducing a stereotyped eye movement.

In experiments 4 and 5, we tested whether adding temporal jitter to the delayed match-to-sample task would influence the latencies of the early and late windows. We hypothesized that jittering the duration of the interval between the sample and probe stimuli would have no effect on the latency at which information is transmitted from the retina to extrastriate cortex but would reduce the participant's ability to predict the latency of the probe stimulus. We could thus investigate whether TMS would disrupt the participant's expectation of an upcoming stimulus versus their perception of it. In experiment 4, participants performed the face delayed match-to-sample task while dTMS was delivered over rOFA and vertex, whereas in experiment 5, participants performed the body delayed match-to-sample task while dTMS was again delivered over rOFA and vertex. Experiments 4 and 5 followed the same trial procedure as experiment 1 except that the timing of the interstimulus interval between the presentation of the prime and probe stimuli was varied. After the prime stimulus, the mask image was presented for 500 ms and was followed by a blank screen that was present for 500, 750, 1000, or 1250 ms. Each of these four blank screen presentation times was used 10 times during each group of 40 trials in each TMS time window, and the order was randomized to produce a jittered interstimulus interval.

In experiment 6, we tested whether the existence of two discrete windows of processing might arise only when a working memory load is imposed, as in the delayed match-to-sample task. Thus, in experiment 6, subjects performed a gaze discrimination task on a single stimulus, a task that does not require holding of stimulus information over a delay. dTMS separated by 10 ms was delivered over the rOFA and vertex at different latencies after stimulus onset to test whether the early and late windows were task dependent (Fig. 2*b*). The gaze task consisted of 40 unique trials, eight trials each in which gaze was directed 10° right, 5° right, direct, 5° left, and 10° left. These 40 trials were repeated across the seven different dTMS time windows (20–30, 40–50, 60–70, 80–90, 100–110, 120–130, and 140–150 ms) producing a total of 280 trials per TMS site (rOFA and vertex). Each subject completed four TMS blocks consisting of 140 trials each (two rOFA and two vertex), and block order was alternated between TMS sites and balanced across subjects. Within each block, the dTMS delivery times were interleaved and randomized. Each subject also completed two no-TMS blocks consisting of 40 trials each. Gaze stimuli were presented for 150 ms.

Results

Over six experiments, we measured performance while participants performed delayed match-to-sample tasks on face and body stimuli and an eye-gaze discrimination task while dTMS (pulses separated by 10 ms) was delivered at different latencies after stimulus onset over the rOFA or rEBA. In experiments 1–5, performance was measured with d' (Green and Swets, 1966), an unbiased measure of performance accuracy. In experiment 6, participants performed a three-choice forced-discrimination task so performance was measured with percentage correct. Analysis of the reaction time data showed that TMS produced no significant main effects ($p > 0.15$) or significant interactions of TMS delivery time and TMS site ($p > 0.3$) across all six experiments.

Experiment 1: face processing in the rOFA

Participants performed a delayed match-to-sample task on faces (Fig. 1). To determine precisely when category-selective visual cortex first contributes to face perception, dTMS was delivered over the functionally localized rOFA and a vertex control site at different latencies ranging from 20 to 140 ms after stimulus onset. Face processing was impaired when dTMS was delivered over the rOFA relative to the vertex at two temporally distinct time windows, the first at 40/50 ms and the second at 100/110 ms but at no other times (Fig. 3*a*). A 2×7 repeated-measures ANOVA with TMS site and TMS time window as independent factors showed a main effect of TMS site ($F_{(1,11)} = 8.5$, $p = 0.016$) but not of TMS time window ($F_{(6,66)} = 1.4$, $p = 0.22$). TMS site and TMS time window combined in a significant two-way interaction ($F_{(6,66)} = 3.2$, $p = 0.009$). Bonferroni's corrected *post hoc* tests showed that task performance was impaired when dTMS was delivered over the rOFA relative to the vertex during the 40/50 ms time window ($p = 0.004$) and the 100/110 ms time window ($p < 0.001$). No other *post hoc* tests approached significance ($p > 0.18$).

The results of experiment 1 demonstrate that the rOFA is causally engaged in face processing during two temporally distinct latencies, an early window occurring at 40/50 ms and a later window occurring at 100/110 ms, with no disruption occurring during the intervening interval from 60 to 90 ms.

Experiment 2: body processing in the rEBA

In experiment 2, we tested whether the two TMS-induced impairment windows we observed in experiment 1 would generalize to another category-selective region in extrastriate cortex. Participants performed a delayed match-to-sample task on body stimuli while dTMS was delivered over the functionally localized rEBA and the vertex control site at different latencies ranging

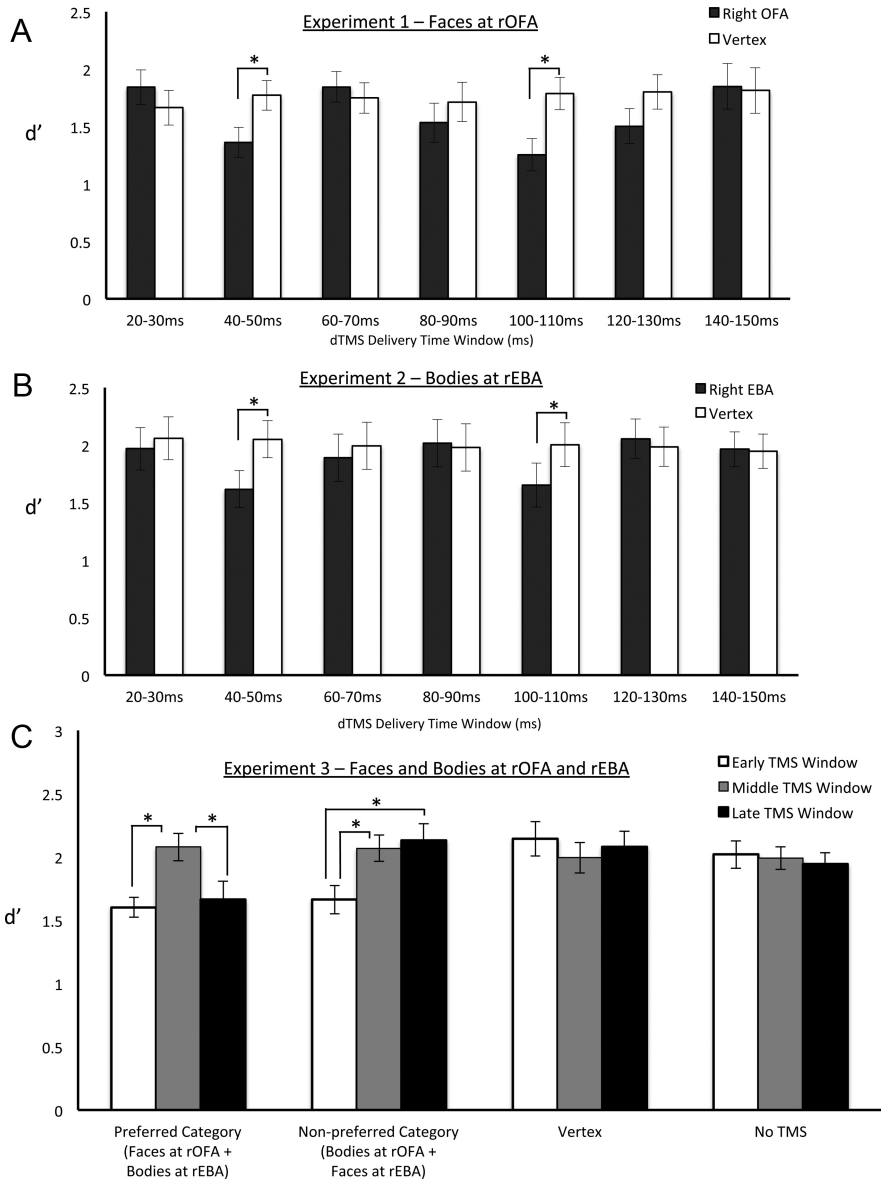

Figure 3. Results from experiments 1–3 (error bars denote SEs). An asterisk denotes a significant difference in Bonferroni's corrected tests. ***A***, dTMS delivered over rOFA disrupted face recognition at 40/50 and 100/110 ms. ***B***, dTMS delivered over rEBA disrupted body recognition at 40/50 and 100/110 ms. ***C***, dTMS delivered over the rOFA and rEBA during the early window disrupted face and body recognition. dTMS delivered during the late window disrupted recognition of faces over the rOFA and bodies over the rEBA only.

from 20 to 140 ms after stimulus onset. Body processing was impaired when dTMS was delivered over the rEBA at two temporally distinct time windows, the first at 40/50 ms and the second at 100/110 ms (Fig. 3*b*). A 2 × 7 repeated-measures ANOVA with TMS site and TMS time window as independent factors failed to show a main effect of TMS site ($F_{(1,11)} = 2.8, p = 0.13$) or of TMS time window ($F_{(6,66)} = 0.5, p = 0.8$). However TMS site and TMS time window combined in a significant two-way interaction ($F_{(6,66)} = 2.3, p = 0.045$). Bonferroni's corrected *post hoc* tests revealed that task performance was impaired when dTMS was delivered over the rEBA relative to the vertex control condition during the 40/50 ms time window ($p = 0.009$) and the 100/110 ms time window ($p = 0.038$). No other *post hoc* tests approached significance ($p > 0.4$).

The results of experiment 2 replicated the timing of the impairment windows we observed for face processing in experiment 1 by demonstrating that the rEBA is causally engaged in body processing during an early time window occurring at 40/50 ms and a later time window at 100/110 ms.

Experiment 3: face and body processing at rOFA and rEBA

In experiment 3, we asked whether the two phases of disruption we observed in experiments 1 and 2 affected only stimuli preferred for the stimulated cortical region (i.e., faces for the OFA and bodies for the EBA) or whether they played a more general role in visual processing. Participants performed face and body delayed match-to-sample tasks (in different conditions) while dTMS was delivered over the functionally localized rOFA, rEBA, and the vertex control site at three different latencies (early, middle, and late). We also included blocks in which no TMS was delivered to act as an additional control condition. The latencies of these three dTMS delivery times were individually calculated for each participant based on their performance during experiments 1 and 2 (for more details, see Materials and Methods). This was done to account for the variability in the latency of the early and late impairments across participants. We wanted to precisely target the two impairment windows on a participant by participant basis to improve the temporal accuracy of the experiment.

We first collapsed performance accuracy into preferred category (face accuracy when dTMS was delivered over rOFA and body accuracy when dTMS was delivered over rEBA) and nonpreferred category (body accuracy when dTMS was delivered over rOFA and face accuracy when dTMS was delivered over rEBA) conditions. Results, shown in Figure 3*c*, demonstrate that dTMS delivered during the early time window impaired processing performance for both the preferred category (faces at rOFA and bodies at rEBA) and nonpreferred category (bodies at rOFA and faces at rEBA) tasks (Fig. 3*c*). In contrast, dTMS delivered during the late time window only impaired processing of the preferred category (faces at rOFA and bodies at rEBA). dTMS delivered during the middle time window had no effect on accuracy for either preferred or nonpreferred categories. Note that we have collapsed the data for faces and bodies together into the category preferred and noncategory preferred conditions for ease of data presentation only. Both the face and the body results showed the same pattern in separate analyses.

These findings were supported by a 3 × 4 repeated-measures ANOVA with TMS time window (early, middle, and late) and TMS site (preferred category, nonpreferred category, vertex, and no TMS) and as independent factors. This analysis found significant main effects of TMS site ($F_{(3,27)} = 4.5, p = 0.011$), TMS time window ($F_{(2,18)} = 3.8, p = 0.04$), and a significant interaction of

the two ($F_{(6,54)} = 7.2, p < 0.001$). Planned Bonferroni's corrected *post hoc* tests showed that dTMS disrupted task performance on the preferred category task (faces at rOFA and bodies at rEBA) in the early time window ($p = 0.001$) and the late time window $p = 0.023$) relative to the middle control window. For the nonpreferred category (bodies at rOFA and faces at rOFA), dTMS delivered in the early time window disrupted task performance relative to the middle ($p = 0.003$) and late ($p = 0.004$) time windows. There were no significant effects between the time windows in the vertex or no-TMS control condition ($p > 0.24$).

Thus, the results of experiment 3 functionally dissociated the early and late impairment windows by demonstrating that dTMS delivered during the early window disrupted perception of both preferred and nonpreferred categories, whereas dTMS was delivered during the late window disrupted perception of preferred categories only.

Experiment 4: face processing at rOFA with temporal jitter

In experiment 4, we tested whether one or both of the impairment windows reflected task preparation rather than bottom-up stimulus processing. To do this, we added jitter to the duration of the interval between the prime and probe face stimuli on the same delayed match-to-sample task that participants had performed in experiment 1. Adding jitter reduced the participants' ability to predict the latency at which the probe stimulus would appear and thus tested whether either of the two impairment windows reflected a disruption of preparatory rather than perceptual processing in the rOFA. dTMS was delivered over the rOFA and vertex at latencies ranging from 20 to 140 ms after probe stimulus onset. We hypothesized that reducing the participant's ability to predict the timing of the probe stimulus onset would not affect the latency of bottom-up neural responses into extrastriate cortex but could affect how they prepared to process that stimulus. Task performance was impaired when dTMS was delivered over the rOFA relative to the vertex control site during three time windows: 40/50, 60/70, and 100/110 ms (Fig. 4*a*). A 2 × 7 repeated-measures ANOVA with TMS site and TMS time window as independent factors showed a main effect of TMS site ($F_{(1,9)} = 7.5, p = 0.023$) but not of TMS time window ($F_{(6,54)} = 0.9, p = 0.52$). TMS site and TMS time window combined in a significant two-way interaction ($F_{(6,54)} = 2.3, p = 0.041$). Planned Bonferroni's corrected *post hoc* tests showed that performance was impaired when dTMS was delivered over the rOFA relative to the vertex control condition in the 40/50 ms ($p = 0.016$), 60/70 ms ($p = 0.037$), and 100/110 ms ($p = 0.026$) time windows. No other *post hoc* tests were significant ($p > 0.13$). Because it is unlikely that a reduced ability to predict the timing of the probe stimulus onset would affect the latency of bottom-up neural responses into extrastriate cortex, the lengthening of the early impairment window favors the hypothesis that the early phase reflects the disruption of preparatory processing. In contrast, the late window remained temporally discrete so is more likely to reflect the disruption of the bottom-up neural response into the rOFA.

Experiment 5: body processing at rOFA with temporal jitter

In experiment 5, we tested whether adding temporal jitter to the task would affect the pattern of category selectivity we observed in experiment 3. Participants performed the same delayed match-to-sample task from experiment 2 with temporal jitter added between the prime and probe stimuli. dTMS was delivered over the rOFA and vertex at latencies ranging from 20 to 140 ms after stimulus onset. Results showed body processing was impaired

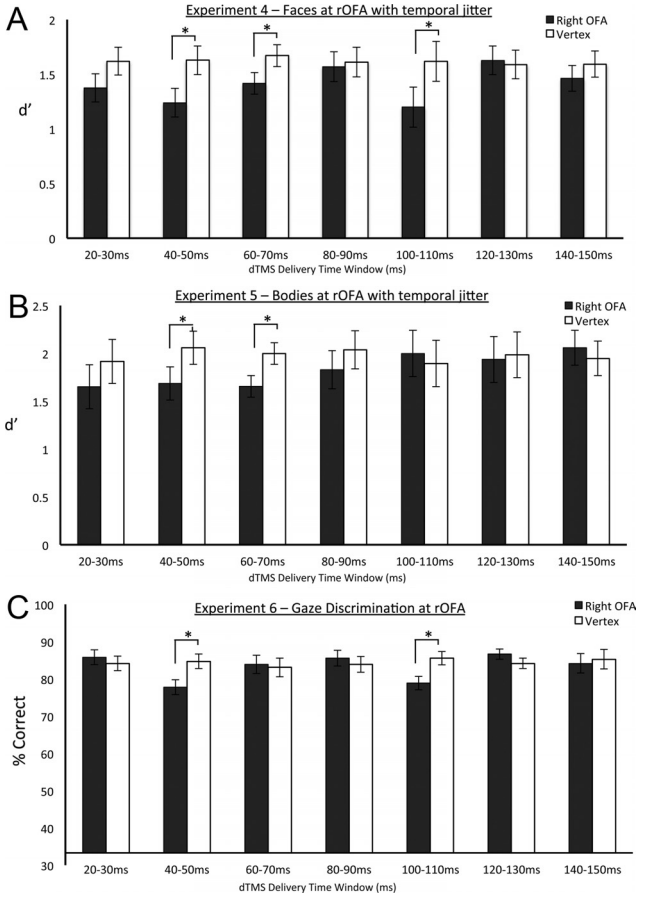

Figure 4. Results from experiments 4–6 (error bars denote SEs). An asterisk denotes a significant difference in Bonferroni's corrected tests. ***A***, dTMS delivered over the rOFA disrupted face recognition at 40/50, 60/70, and 100/110 ms when we added temporal jitter. ***B***, dTMS delivered over the rOFA disrupted body recognition at 40/50 and 60/70 ms when we added temporal jitter. ***C***, dTMS delivered over the rOFA disrupted gaze discrimination at 40/50 and 100/110 ms.

when dTMS was delivered over the rOFA relative to the vertex control site at two time windows, 40/50 and 60/70 ms, but at no other time windows (Fig. 4*b*). A 2 × 7 repeated-measures ANOVA with TMS site and TMS time window as independent factors failed to show main effects of TMS site ($F_{(1,9)} = 3.4, p = 0.095$) or of TMS time window ($F_{(6,54)} = 1.0, p = 0.44$). However, TMS site and TMS time window combined in a significant two-way interaction ($F_{(6,54)} = 2.9, p = 0.015$). Planned Bonferroni's corrected *post hoc* showed that body recognition performance was impaired when dTMS was delivered over the rOFA relative to the vertex control condition in the 40/50 ms ($p = 0.034$) and the 60/70 ms ($p = 0.005$) time windows. No other *post hoc* tests were significant ($p > 0.12$). The results of experiment 5 extend the pattern of results from experiments 3 and 4 by demonstrating that the early window is not category selective and spreads over multiple time windows in the presence of temporal uncertainty, whereas the absence of the late impairment window again suggests that it reflects a category-selective stage of visual processing.

Experiment 6: gaze discrimination at rOFA

In experiment 6, we tested whether the early disruption window could reflect memory recall of the prime stimulus rather than perceptual processing of the probe stimulus in the delayed match-to-sample task. To test this hypothesis, we asked participants to perform a perceptual discrimination task with no mem-

ory component in which they had to judge the direction of eye gaze (left, right, or direct) of a single face stimulus in each trial, presented for 150 ms. dTMS was delivered over the rOFA and vertex at different latencies ranging from 20 to 140 ms after stimulus onset. Gaze discrimination was impaired when dTMS was delivered over the rOFA at two temporally distinct time windows, the first at 40/50 ms and the second at 100/110 ms (Fig. 4*c*). A 2 × 7 repeated-measures ANOVA with TMS site and TMS time window as independent factors did not show a main effect of TMS site ($F_{(1,7)} = 0.7$, $p = 0.6$) or of TMS time window ($F_{(6,36)} = 1.8$, $p = 0.14$). However, TMS site and TMS time window did combine in a significant two-way interaction ($F_{(6,36)} = 5.6$, $p = 0.001$). Planned Bonferroni's corrected *post hoc* tests showed that gaze discrimination performance was impaired when dTMS was delivered over the rOFA relative to the vertex control condition in the 40/50 ms time window ($p = 0.03$) and the 100/110 ms time window ($p = 0.034$). No other *post hoc* tests approached significance ($p > 0.32$). These results demonstrate that the early and late impairment windows are also present in a basic perceptual face discrimination task and thus that the early impairment window does not arise only when the prime stimulus is being held in memory.

Discussion

We exploited the temporal precision of TMS to determine the latency at which the face-selective rOFA and body-selective rEBA causally contribute to face and body processing during the first 150 ms after stimulus onset. TMS induced two temporally distinct impairment windows: an early window at 40/50 ms and a late window at 100/110 ms. The two impairment windows were functionally distinct. TMS delivered over either region during the early window disrupted perception of both faces and bodies. However, during the late window, we observed a category-specific disruption, in which TMS delivered over rOFA impaired face but not body perception and TMS delivered over rEBA impaired body but not face perception. In two follow-up experiments, we further established the functional profiles of the two windows by varying the interstimulus interval between the prime and the probe stimuli. We reasoned that adding temporal jitter would reduce the participants' ability to predict when the probe stimulus would appear and thus jitter should disrupt preparatory but not bottom-up perceptual processing. Indeed, the length of the early impairment window was extended (40/50 and 60/70 ms) when temporal jitter was added, whereas the late time window was unaffected, implicating preparatory processing in the first window but not the second. Finally, we established that both impairment windows were present in a simple discrimination task with no stimulus recall component, suggesting that neither impairment window reflected disruption of working memory processes for faces in the rOFA. Together, our six experiments demonstrate that two functionally distinct phases of processing occur in extrastriate category-selective regions during the first 150 ms after stimulus onset.

Although it has been suggested that a single feedforward pass of information into visual cortex is sufficient for object categorization (Thorpe et al., 1996; VanRullen and Thorpe, 2001), the evidence we report here for two functionally distinct phases of visual perception is not without precedent. Previous studies have reported an early nonspecific phase of stimulus processing followed by a later specific phase, although the precise latency of each phase differs across these studies, which have used different experimental techniques. For example, an ERP visual attention study identified two functionally distinct stages in response to visually presented checkerboard patterns (Martínez et al., 1999): an early phase at 50–80 ms, unmodulated by attention, and a later phase at ~75–135 ms that was modulated by attention. An MEG study also reported an early nonselective response to face and house stimuli peaking at 85 ms, followed by a face-specific response peaking later at 160 ms (Furey et al., 2006; but see Liu et al., 2002). In addition, a study that recorded local field potentials (LFPs) from the middle face-selective patch in macaques reported an initial nonselective response to multiple categories (peaking at 100 ms), followed by a face-selective response (peaking at 130 ms) (Tsao et al., 2006, their supplemental Fig. 4). Our results extend the findings of these studies by demonstrating that both processing phases perform a causal role in perception, in the sense that TMS delivered during either phase disrupted task performance.

What perceptual functions might be performed during these two phases of visual stimulus processing reported here? The late window was category selective, namely TMS delivered over rOFA impaired face but not body recognition, whereas TMS delivered over rEBA impaired body but not face recognition. This result is consistent with previous TMS studies of these areas (Urgesi et al., 2004; Pitcher et al., 2009) and with the category-selective impairments in neuropsychological patients (Rossion et al., 2003; Moro et al., 2008). The latency of the late window (100–110 ms) is also consistent with the earliest face-selective response reported in ERP and MEG studies of face perception (Eimer, 1998; Liu et al., 2002; Sadeh et al., 2010). Moreover, Liu et al. (2009) were able to successfully decode object category information from the LFPs recorded intracranially in the visual cortex of preoperative epileptic patients at latencies beginning 115 ms after stimulus onset. It is also important to note that the latency of the late impairment window was not affected by varying the interstimulus interval in experiments 4 and 5. It is unlikely that adding temporal jitter to the task would influence the latency at which incoming visual information is transmitted from the retina to extrastriate cortex. Thus, the temporal consistency of the late TMS window (in contrast to the temporal smearing of the early window) further suggests that this latency (100 ms after stimulus onset) reflects the earliest latency at which category-selective responses contribute to within-category perceptual processing of faces and bodies.

The functional significance of the early impairment window is less obvious. Task disruption during this early time window was not restricted to the preferred category of the stimulated region. Furthermore, the precise latency of the disruption window was not fixed but spread out over a broader range of latencies when we added jitter to the interstimulus interval in experiments 4 and 5. Finally, the results of the gaze discrimination task in experiment 6 demonstrated that the early window is unlikely to result from disruption of working memory for the prime stimulus, because similar disruption was found in the a task requiring little or no working memory. We offer two hypotheses to account for what functions might be performed in extrastriate cortex ~40–70 ms after stimulus onset.

The first possible account of the early impairment window is that it reflects the disruption of preparatory mechanisms that orient the participant to the upcoming presentation of the expected visual stimulus. The strongest evidence that TMS may be disrupting preparation, rather than perception, is the latency of the early window: 40/50 ms seems too short a latency for visual information to be transmitted from the retina to extrastriate cortex. This hypothesis, that the early window may result from disruption of the participant's expectation of the upcoming stimulus is also consistent with fMRI studies reporting an in-

crease in the neural response in visual cortex preceding the onset of an expected visual stimulus (Kastner et al., 1998, 1999; Ress et al., 2000). Moreover, the results of experiments 4 and 5, in which adding temporal jitter between the prime and probe stimulus smeared the latency of the early window across two TMS delivery times (40/50 and 60/70 ms), further supports this hypothesis: adding jitter to the task can be expected to affect preparatory processing but is unlikely to change the latency at which information is transmitted from the retina to extrastriate cortex. These considerations suggest that TMS delivered during the early window disrupted preparation in the stimulated region for the upcoming stimulus rather the processing of incoming visual input. An open question on this account is why the early impairment window does not encompass the initial TMS delivery time window at 20/30 ms.

Another possible account of the induced impairments is that both the early and late impairment windows result from disruptions of the incoming visual input. For example, it is possible that the early window corresponds to the first sweep of visual information into visual cortex and that the late window corresponds to re-entrant processing from remote cortical areas that enhance this initial neural representation. This hypothesis is consistent with models of visual awareness that propose an early phase of neural activity beginning at ~30 ms in which the participant is unaware of the incoming visual input (Lamme and Roelfsema, 2000; Lamme, 2003). Awareness of visual stimuli then arises via re-entrant feedback processing from higher cortical areas to early visual cortex at ~100 ms. It should be noted that the latencies in this model are based on latencies recorded from nonhuman primates that may be shorter than those that would operate in human cortex.

However, Lamme's model would, at least qualitatively, be consistent with the pattern of results we observed here. For example, TMS delivered over early visual cortex has been shown to disrupt task performance as early as 30 ms after stimulus onset (Corthout et al., 2007), and it is possible that the disruptive effect of TMS delivered over rOFA and rEBA in the present study could have spread to early visual areas (but see Pitcher et al., 2009). If this were the case, then the nonselective early dip could arise from disruption of both the stimulated region (rOFA or rEBA) and early visual cortex, demonstrating that both regions are essential. In contrast, the category selectivity of the later dip would suggest that, by 100 ms, only the stimulated region (rOFA or rEBA), but not early visual cortex, was essential to task performance.

One of the most striking aspects of human vision is the speed of object recognition. We have demonstrated here that category-selective extrastriate regions contribute to visual recognition in two temporally and functionally distinct processing phases: an early and domain-general phase that may reflect preparatory signals, or perhaps instead a very rapid initial pass of feedforward information into the visual system and a later phase in which each region is selectively engaged in processing its preferred stimulus category.

References

Allison T, Puce A, Spencer DD, McCarthy G (1999) Electrophysiological studies of human face perception. I. Potentials generated in occipitotemporal cortex by face and non-face stimuli. Cereb Cortex 9:415–430.

Bentin S, Allison T, Puce A, Perez E, McCarthy G (1996) Electrophysiological studies of face perception in humans. J Cogn Neurosci 8:551–565.

Corthout E, Hallett M, Cowey A (2007) TMS-induced scotomata: time-based neglect. Clin Neurophysiol 118:1895–1898.

Cox RW, Jesmanowicz A (1999) Real-time 3D image registration for functional MRI. Magn Reson Med 42:1014–1018.

Crouzet SM, Kirchner H, Thorpe SJ (2010) Fast saccades toward faces: face detection in just 100 ms. J Vis 10:16.1–16.17.

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9:179–194.

Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. Science 293:2470–2473.

Eimer M (1998) Does the face-specific N170 component reflect the activity of a specialized eye processor? Neuroreport 9:2945–2948.

Engell AD, McCarthy G (2010) Selective attention modulates face-specific induced gamma oscillations recorded from ventral occipitotemporal cortex. J Neurosci 30:8780–8786.

Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis. II. Inflation, flattening, and a surface-based coordinate system. Neuroimage 9:195–207.

Furey ML, Tanskanen T, Beauchamp MS, Avikainen S, Uutela K, Hari R, Haxby JV (2006) Dissociation of face-selective cortical responses by attention. Proc Natl Acad Sci U S A 103:1065–1070.

Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW (2000) The fusiform "face area" is part of a network that processes faces at the individual level. J Cogn Neurosci 12:495–504.

Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.

Grill-Spector K, Kanwisher N (2005) Visual recognition: as soon as you know it is there, you know what it is. Psychol Sci 16:152–160.

Grill-Spector K, Knouf N, Kanwisher N (2004) The fusiform face area subserves face perception, not generic within-category identification. Nat Neurosci 7:555–562.

Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. Trends Cogn Sci 4:223–233.

Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 17:4302–4311.

Kastner S, De Weerd P, Desimone R, Ungerleider LG (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. Science 282:108–111.

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron 22:751–761.

Lamme VA (2003) Why visual attention and awareness are different. Trends Cogn Sci 7:12–18.

Lamme VA, Roelfsema PR (2000) The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci 23:571–579.

Liu H, Agam Y, Madsen JR, Kreiman G (2009) Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. Neuron 62:281–290.

Liu J, Harris A, Kanwisher N (2002) Stages of processing in face perception: an MEG study. Nat Neurosci 5:910–916.

Martínez A, Anllo-Vento L, Sereno MI, Frank LR, Buxton RB, Dubowitz DJ, Wong EC, Hinrichs H, Heinze HJ, Hillyard SA (1999) Involvement of striate and extrastriate visual cortical areas in spatial attention. Nat Neurosci 2:364–369.

Moro V, Urgesi C, Pernigo S, Lanteri P, Pazzaglia M, Aglioti SM (2008) The neural basis of body form and body action agnosia. Neuron 60:235–246.

Moscovitch M, Winocur G, Behrmann M (1997) What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. J Cog Neurosci 9:555–604.

O'Shea J, Muggleton NG, Cowey A, Walsh V (2004) Timing of target discrimination in human frontal eye fields. J Cogn Neurosci 16:1060–1067.

Peelen MV, Downing PE (2007) The neural basis of visual body perception. Nat Rev Neurosci 8:636–648.

Pitcher D, Walsh V, Yovel G, Duchaine B (2007) TMS evidence for the involvement of the right occipital face area in early face processing. Curr Biol 17:1568–1573.

Pitcher D, Garrido L, Walsh V, Duchaine BC (2008) TMS disrupts the perception and embodiment of facial expressions. J Neurosci 28:8929–8933.

Pitcher D, Charles L, Devlin JT, Walsh V, Duchaine B (2009) Triple dissociation of faces, bodies, and objects in extrastriate cortex. Curr Biol 19:319–324.

Pitcher D, Dilks DD, Saxe RR, Triantafyllou C, Kanwisher N (2011a) Differential selectivity for dynamic versus static information in face-selective cortical regions. Neuroimage 56:2356–2363.

Pitcher D, Duchaine B, Walsh V, Yovel G, Kanwisher N (2011b) The role of the occipital face area and lateral occipital area in the face inversion effect. Neuropsychologia 49:3448–3453.

Potter MC, Faulconer BA (1975) Time to understand pictures and words. Nature 253:437–438.

Puce A, Allison T, Asgari M, Gore JC, McCarthy G (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. J Neurosci 16:5205–5215.

Ress D, Backus BT, Heeger DJ (2000) Activity in primary visual cortex predicts performance in a visual detection task. Nat Neurosci 3:940–945.

Rossion B, Caldara R, Seghier M, Schuller AM, Lazeyras F, Mayer E (2003) A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. Brain 126: 2381–2395.

Sadeh B, Podlipsky I, Zhdanov A, Yovel G (2010) Face-selective fMRI and event-related potential responses are highly correlated: evidence from simultaneous ERP-fMRI investigation. Hum Brain Mapp 31:1490–1501.

Sadeh B, Pitcher D, Brandman T, Eisen A, Thaler A, Yovel G (2011) Stimulation of object-category selective areas modulates ERP to their preferred categories. Curr Biol 21:1894–1899.

Silvanto J, Lavie N, Walsh V (2005) Double dissociation of V1 and V5/MT activity in visual awareness. Cereb Cortex 15:1736–1741.

Thierry G, Pegna AJ, Dodds C, Roberts M, Basan S, Downing P (2006) An event-related potential component sensitive to images of the human body. Neuroimage 32:871–879.

Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. Nature 381:520–522.

Tsao DY, Freiwald WA, Tootell RB, Livingstone MS (2006) A cortical region consisting entirely of face cells. Science 311:670–674.

Urgesi C, Berlucchi G, Aglioti SM (2004) Magnetic stimulation of extrastriate body area impairs visual processing of nonfacial body parts. Curr Biol 14:2130–2134.

VanRullen R, Thorpe SJ (2001) The time course of visual processing: from early perception to decision-making. J Cogn Neurosci 13:454–461.

Yoshor D, Bosking WH, Ghose GM, Maunsell JH (2007) Receptive fields in human visual cortex mapped with surface electrodes. Cereb Cortex 17: 2293–2302.