

Stages of processing in face perception: an MEG study

Jia Liu¹, Alison Harris¹ and Nancy Kanwisher^{1,2,3}

¹ Department of Brain and Cognitive Sciences, NE20-443, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

² The McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

³ MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, Massachusetts 02129, USA

Correspondence should be addressed to J.L. (liujia@psyche.mit.edu)

Published online: 19 August 2002, doi:10.1038/nn909

Here we used magnetoencephalography (MEG) to investigate stages of processing in face perception in humans. We found a face-selective MEG response occurring only 100 ms after stimulus onset (the 'M100'), 70 ms earlier than previously reported. Further, the amplitude of this M100 response was correlated with successful categorization of stimuli as faces, but not with successful recognition of individual faces, whereas the previously-described face-selective 'M170' response was correlated with both processes. These data suggest that face processing proceeds through two stages: an initial stage of face categorization, and a later stage at which the identity of the individual face is extracted.

Face recognition is one of the most important problems the human visual system must solve. Here we used MEG to characterize the sequence of cognitive and neural processes underlying this remarkable ability.

Two candidate stages of face processing are the categorization of a stimulus as a face, and the identification of a specific individual. Several studies of object recognition suggest that objects are first categorized at a 'basic level' (dog, bird) before a finer 'subordinate level' identification is achieved (poodle, sparrow)^{1,2}. Evidence that a similar sequence may occur in face perception comes from single-unit recordings in macaques showing that the initial transient response of face-selective neurons in inferotemporal cortex reflects a rough categorization of face versus nonface, whereas subsequent firing of the same neural population represents finer information such as facial expression or individual identity³. It has been argued, however, that visual expertise in discriminating exemplars of a specific visual category may shift the point of initial contact with memory representations from the basic level to the subordinate level^{1,4}. Given that we are all experts at face recognition, this hypothesis predicts that people should be able to identify an individual face as fast as they can determine that it is a face at all. Although some behavioral evidence is consistent with this hypothesis⁵, other evidence is not⁶.

MEG is an ideal technique for addressing these questions, as its high temporal resolution enables us to tease apart processing stages that may occur within tens of milliseconds of each other. Prior work with MEG and event-related potentials (ERPs) has characterized a response component called the N170 (or M170 in MEG) that occurs around 170 ms after stimulus onset, and is about twice as large for face stimuli as for a variety of control nonface stimuli such as hands, houses or animals^{7–10}. This response is thought to reflect the structural encoding of a face^{7,11,12}, that is, the extraction of a perceptual representation of the face. Although several reports of even earlier category-

selective responses have been published^{13–18}, they are open to explanations invoking nonspecific repetition effects¹⁹ or differences in the low-level features of the stimuli²⁰.

When do the first truly face-selective responses occur? We recorded MEG responses while subjects passively viewed a sequence of photographs of faces and a variety of control stimuli (experiments 1 and 2). These experiments found a new face-selective response occurring only 100 ms after stimulus onset (the M100), generated from extrastriate cortex (experiment 3). Further, we tested whether the M100 and M170 amplitudes are correlated with success in face categorization and/or face identification (experiment 4). Finally, we tested for further qualitative differences in the processes underlying the M100 and the M170 by measuring the responses of each component to face configurations and face parts (experiment 5).

RESULTS

A face-selective response at a latency of 100 ms

An interpolated map of the *t*-value comparing the MEG response to faces versus houses for a typical subject (experiment 1) shows the previously-described¹⁰ face-selective M170 response occurring at a latency of about 160 ms (Fig. 1a). This result may correspond approximately to the N170 component in ERP studies⁷ and/or the N200 in intracranial ERP studies²¹, as discussed later.

In addition to the M170, we found a smaller response peaking at a mean latency of 105 ms (the 'M100'; range 84.5–130.5 ms, s.d. 16.1) that was significantly higher for faces than for houses. This result was seen with the same polarity in 13 of 15 subjects. The scalp distribution of the face-selective M100 response was slightly posterior to that of the M170, but the sensors showing the strongest face-selectivity for the M100 largely overlapped with those showing the strongest face-selectivity for the M170. The MEG response of a representative subject at a typical overlapping face-selective sensor in the right hemisphere is shown in Fig. 1b.

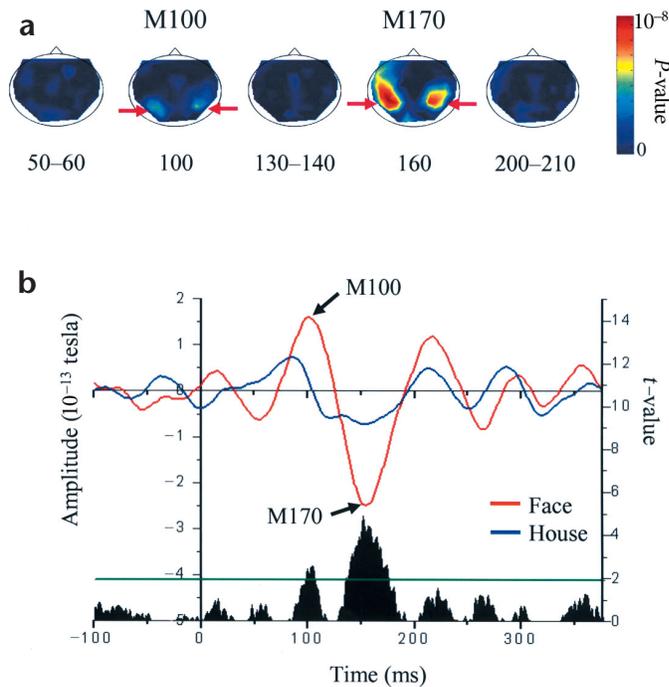


Fig. 1. MEG data from a typical subject. **(a)** Pairwise t-tests between the responses at each sensor reveal early (M100) and late (M170) significant differences in the MEG response to faces versus houses over occipitotemporal cortex. **(b)** The MEG waveforms are averaged across all face and house trials at a typical sensor of interest in the right hemisphere. Red, faces; blue, houses; black, t-values. The left vertical scale indicates the amplitude of the MEG response (10^{-13} tesla) whereas the right one shows the t-value. A value $t = 1.99$ (horizontal green line) corresponds to $P < 0.05$ (uncorrected for comparisons at multiple time points).

For a stronger test of face selectivity, we measured the magnitude of the M100 response to a variety of control stimuli (experiment 2). Subjects were asked to press a button whenever two consecutive images were identical, obligating them to attend to all stimuli regardless of inherent interest. Target trials containing such repeated stimuli were excluded from the analysis. Accuracy on this one-back matching task was high for all categories (>90% correct) except for hands (76%), which are visually very similar to each other. The MEG data from experiment 1 were first used to define sensors of interest (SOIs) in each subject—those sensors that produced significantly face-selective responses for both the M100 and the M170 (Methods). Both the amplitudes and latencies of peak responses to the new stimulus conditions in experiment 2 in these same SOIs were then quantified in the same subject in the same session. The M100 response to photographs of faces was greater (Fig. 2) than that to photographs of animals ($F_{1,12} = 10.2, P < 0.01$), human hands ($F_{1,12} = 9.0, P < 0.02$), houses ($F_{1,12} = 8.1, P < 0.02$) and nonface objects ($F_{1,12} = 10.3, P < 0.01$). Therefore, the M100 is not generally selective for anything animate, or for any human body part; instead, it seems to be selective for faces. However, both the magnitude and selectivity of the M100 response were weaker than those for the M170 response.

The earlier latency and somewhat more posterior scalp distribution of the M100 compared to the M170 suggest that the two components may not originate from the same anatomical source. To test whether the M100 might originate in retinotopically-organized visual cortex, we separately measured the

amplitude of the M100 in each hemisphere to faces presented in the contralateral versus ipsilateral visual field (2.8° off fixation) in experiment 3. This manipulation is known to affect responses in visual areas V1, V2, V3, VP, V3A and V4v in humans²². We found no difference in the amplitude of the M100 in each hemisphere for contralaterally versus ipsilaterally presented faces ($F_{1,10} < 1$), indicating that the source of this component must be beyond retinotopic cortex.

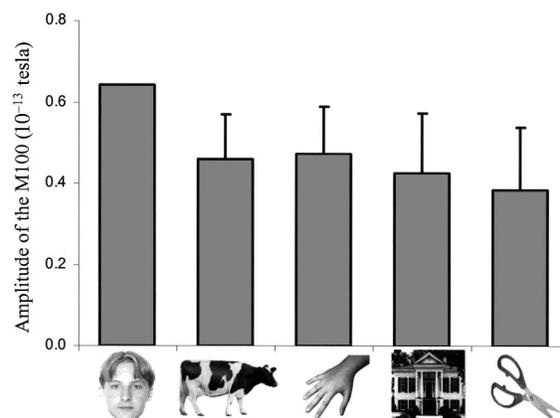
Decoupling categorization and identification

The results described so far implicate both the M100 and M170 in face processing, but do not indicate what aspect of face processing each component reflects. In experiment 4, we asked whether each component is involved in the categorization of a stimulus as a face, in the extraction of the individual identity of a face, or both. Subjects were instructed to make two judgments about each stimulus, determining both its category (face or house) and its individual identity.

In this experiment, ten subjects matched front-view test images of faces and houses to profile views (faces) or three-quarter views (houses) of the same stimulus set presented earlier in the same trial (Fig. 3a). There were three possible responses: ‘different category’ if the sample was a face and the test was a house or vice versa, ‘different individual but same category’ and ‘same individual’. Correct categorization required discrimination between ‘different category’ trials and either ‘different individual’ or ‘same individual’ trials; correct identification required distinguishing between ‘different individual’ and ‘same individual’ trials.

A set of face and house stimuli (five exemplars each) were constructed, each of which had identical spatial frequency, luminance and contrast spectra. Subjects were first trained to

Fig. 2. Amplitudes of the peak M100 response, averaged across subjects, to faces and a variety of nonface objects at predefined sensors of interest. The error bars show the standard deviation across subjects of the difference of the M100 amplitudes between faces and each category of nonface object.



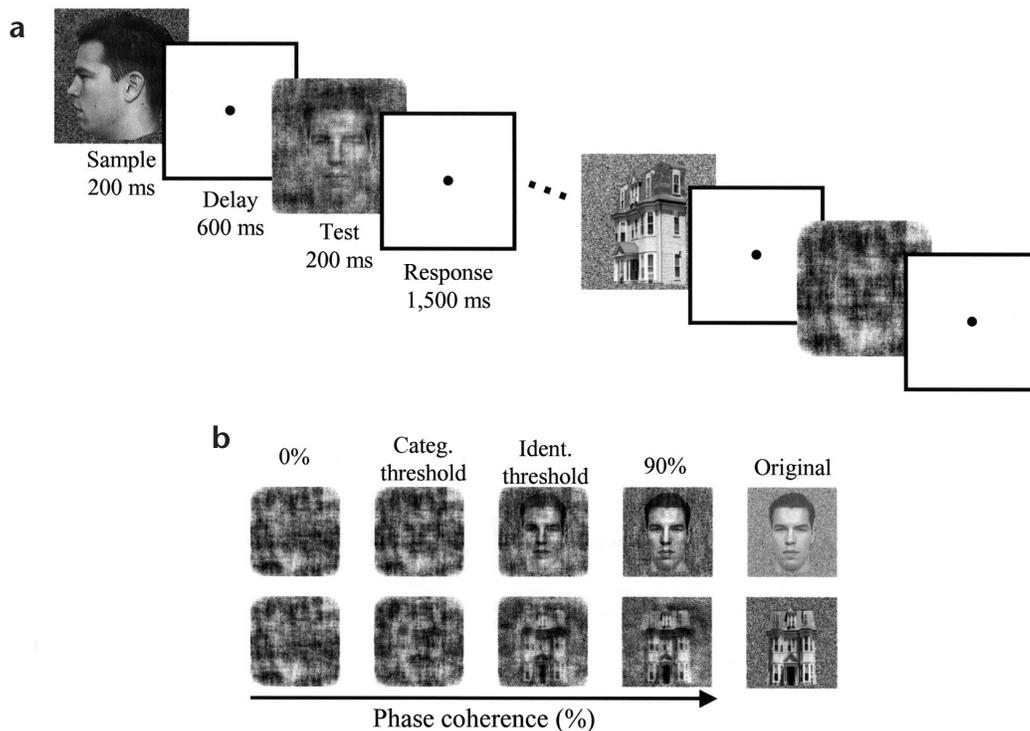


Fig. 3. Stimulus and task for experiment 4. **(a)** In each trial, a sample stimulus was followed (after a delay) by a test stimulus. The sample images (5 exemplars for each category) were profile-view faces or three-quarter-view houses. **(b)** Test stimuli were frontal views of the sample stimuli. The phase coherence of the test stimuli was varied from 0% (visual noise) to 90% in four levels; original images with 100% coherence were also included. Here we show the data for the stimuli presented at categorization and identification thresholds only.

match each face and house with its profile or three-quarter view, respectively, at 100% accuracy. Using a technique similar to the recently proposed RISE technique^{23,24} (Methods), each subject was then run on a psychophysical staircase procedure in which the percentage of phase coherence of each test stimulus was gradually reduced until the subject reached threshold performance on the matching task (75% correct, 20 steps by QUEST staircase²⁵). In this way, five threshold face stimuli and five threshold house stimuli were constructed for each subject for each of the two tasks (categorization and identification). Across all stimuli and subjects, the resulting threshold face and house stimuli had a mean percent phase coherence of 14% (face) and 18% (house) for the categorization task and 38% (face) and 51% (house) for the identification task, indicating that more stimulus information was necessary for the identification task than for the categorization task, as expected.

Next, each subject performed the same matching task (different category, different individual, or same individual) in the MEG scanner, using face and house stimuli with phase coherence varied across four levels: 0%, 90%, and the two previously-derived sets of thresholds for that subject—one for the categorization task, and the other for the identification task (Fig. 3b). In addition, the original version of each image with unmodified spatial frequencies was included to localize face-selective SOIs. By measuring both categorization and identification performance on each trial, the task allowed us to decouple the MEG correlates of successful categorization from those of successful identification. To obtain the MEG correlates of successful categorization, we compared the average

MEG response to the same test image when the subject correctly categorized but failed to identify it, versus when they categorized it incorrectly. For identification, we compared the response to the same test image when the subject correctly identified it versus when they incorrectly identified it but categorized it successfully.

MEG waveforms averaged across each subject's face-selective SOIs from the face categorization and identification tasks are shown in Fig. 4a. The magnitudes of both the M100 ($F_{1,9} = 9.5, P < 0.02$) and the M170 ($F_{1,9} = 5.8, P < 0.05$) were larger for successful than for unsuccessful categorization of faces (Fig. 4b, top left). However, only the M170 ($F_{1,9} = 43.3, P < 0.001$), but not the M100 ($F_{1,9} < 1$), was higher for correct than for incorrect identification of faces (interaction, $F_{1,9} = 8.7, P < 0.02$) (Fig. 4b, bottom left). For house stimuli, neither the M100 nor the M170 differed for correct versus incorrect trials in either task ($F < 1$ in all cases; Fig. 4b, top right and bottom right). The finding that the M170 is specific for face identification (not house identification) is further supported by a significant three-way interaction ($F_{1,9} = 6.73, P < 0.03$) of face versus house identification, success versus failure, and M100 versus M170 (Fig. 4b, bottom left and right). In addition, neither the main effect of hemisphere nor the interaction of task by hemisphere was significant ($F < 1$ in both cases).

Accuracy on categorization and identification tasks at the two levels of phase coherence (derived from the previous psychophysical measurements) is shown in Table 1. Pairwise *t*-tests revealed no significant difference ($P > 0.2$ in all cases) in accuracy as a function of stimulus category (faces versus houses) for

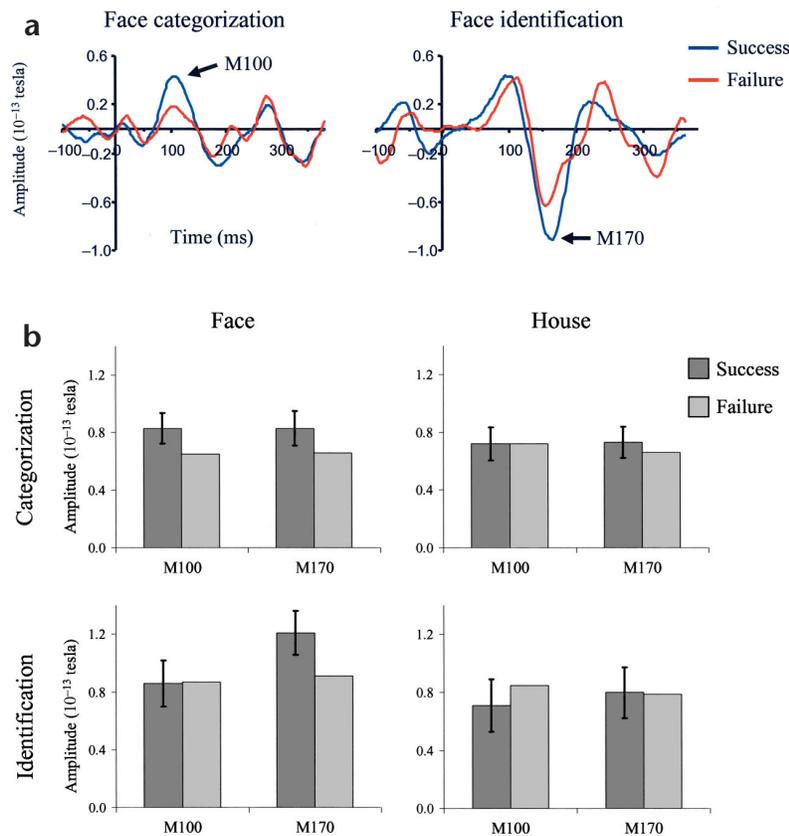


Fig. 4. Categorization versus identification. (a) MEG waveforms from the face categorization (left) and identification (right) tasks. Blue, success; red, failure. The waveforms were generated by averaging the selected raw data (Methods) from independently defined SOIs in ten subjects. (b) The amplitudes of the M100 and M170 at SOIs averaged across subjects. Successful categorization of faces elicited higher amplitudes at both M100 and M170 (top left), but no significant difference was found between successfully and unsuccessfully categorized houses at the predefined SOIs (top right). Correctly identified (compared to incorrectly identified) faces produced a significantly larger amplitude of the M170, but not of the M100 (bottom left). The amplitude elicited by houses was not affected by success or failure in the identification task (bottom right).

either task (categorization versus identification). Therefore, any difference between faces and houses seen in MEG responses cannot be explained in terms of differences in behavioral performance. Note that even when the stimuli were degraded to ‘categorization threshold level’, the subjects’ performance in the identification task was above chance ($P < 0.01$ in both cases).

In sum, both the M100 and M170 are correlated with successful face categorization, but only the later M170 component is correlated with successful face identification, indicating a dissociation between the two processes. One possible account of this dissociation is that the selectivity of the underlying neural population may continuously sharpen over time, permitting crude discriminations (for example, between a face and a house) earlier, and more fine-grained discriminations (for example, between two different faces) later. Indeed, the ratio of the response to faces versus houses was lower for the M100 (1.6) than for the M170 (1.8, interaction $P < 0.03$), showing that selectivity does indeed sharpen over time. However, this fact alone does not indicate whether the selectivity changes only in degree, or whether qualitatively different representations underlie the M100 and the M170. This question was addressed in experiment 5, in which we measured the M100 and M170 responses to information about face configurations and face parts.

In experiment 5, two face-like stimulus categories were constructed from veridical human faces by orthogonally eliminating or disrupting either face configuration or face parts (eyes, nose and mouth; Fig. 5a). Specifically, for ‘configuration’ stimuli, face parts in each stimulus were replaced by solid black ovals in their corresponding locations, preserving the face configuration but removing the contribution of face parts

(Fig. 5a, left). Conversely, for ‘parts’ stimuli, the face parts were kept intact but were rearranged into a novel nonface configuration (Fig. 5a, right). We measured MEG responses in 14 subjects who, while fixating, passively viewed these two sets of face-like stimuli (50 exemplars each) presented in a random order.

The responses to configuration and part stimuli recorded at independently defined face-selective sensors, averaged across subjects, are shown in Fig. 5b. Critically, we found a significant two-way interaction of M100 versus M170 by configuration versus parts ($F_{1,13} = 13.4, P < 0.005$). This interaction reflects the fact that the amplitude of the M100 was significantly larger for parts stimuli than for configuration stimuli ($F_{1,13} = 11.5, P < 0.005$), whereas a trend in the opposite direction was found for the M170 ($F_{1,13} = 3.35, P = 0.09$). Thus it is not merely the degree of selectivity, but the qualitative nature of the selectivity, that differs between the M100 and the M170. Again, neither the main effect of hemisphere nor the interaction of stimulus type by hemisphere was significant ($F < 1$ in both cases).

DISCUSSION

In experiments 1–3, we report a MEG response component, occurring over occipitotemporal cortex and peaking at a latency of ~100 ms, that is significantly larger for faces than for a variety of nonface objects. This result indicates that the categorization of a stimulus as a face begins within 100 ms after stimulus onset, substantially earlier than previously thought^{7,20,26}.

Unlike prior reports of very early category-selective ERP responses^{13–18}, the M100 reported here cannot be explained in terms of differences in the low-level features present in the stimuli. The M100 response was stronger when the same face stimulus was correctly perceived as a face than when it was wrongly categorized as a nonface. This result shows that the M100 reflects the subject’s percept, not simply the low-level properties of the stimulus.

It is possible that a correlate of the face-selective M100 could be obtained with ERPs. However, because MEG is sensitive to only a subset of the neural activity that can be detected with scalp ERPs²⁷, there is no simple correspondence between MEG responses and scalp ERP responses, and selectivities that are clear in the MEG response may be diluted with ERPs. Similarly, the



Table 1. Accuracy as a function of task and stimulus category (experiment 4, guessing corrected).

| | Categorization task | | Identification task | |
|--------------------------|---------------------|-------|---------------------|-------|
| | Face | House | Face | House |
| Categorization threshold | 74% | 72% | 26% | 19% |
| Identification threshold | 95% | 95% | 73% | 65% |

M170 response measured with MEG probably corresponds to only one of the two sources hypothesized to underlie the N170 response^{7,28}. On the other hand, direct correspondences may exist between the M100 and M170 and the more focal intracranial P150 and N200 ERPs²¹, respectively, assuming that the later latencies in the intracranial case arise from medication and/or histories of epilepsy typical of that subject population. Unfortunately, the limitations in current source localization techniques leave these correspondences only tentative at present.

The latency of the M100 response is not directly comparable to the category-selective response that occurs at a latency of 100 ms in inferotemporal (IT) neurons in macaques^{29,30} because all cortical response latencies are shorter in macaques than humans. For example, V1 responses occur 40–60 ms after stimulus presentation in macaques³¹—about 20 ms earlier than they do in humans^{26,32}.

Given that at least 60–80 ms are thought to be necessary for visual information to reach primary visual cortex in humans³², this leaves only an additional 20–40 ms for the first face-selective responses to be generated in cortex. Such latencies are hard to reconcile with models of visual categorization that rely heavily on iterative feedback loops and/or recurrent processing, and strengthen the claim that initial stimulus categorization is accomplished by largely feedforward mechanisms³³.

The second major finding of this study is that both the M100 and the M170 are correlated with successful face categorization, but only the later M170 component is correlated with successful face identification. This finding indicates that processes critical for the identification of a face begin at a substantially later latency than processes critical for the categorization of a stimulus as a face. Evidently, our expertise with faces has not led us to be able to identify individual faces as early as we can tell they are faces at all (as argued in ref. 5).

The dissociation we report here between the processes underlying face categorization and those underlying face identification do not simply reflect the greater difficulty of identification compared to categorization, because our results were obtained even when the stimuli were adjusted to produce identical performance in the categorization and identification tasks (experiment 4). Furthermore, the difference in the response for successful versus unsuccessful trials on face stimuli cannot be explained by general processes such as attention or association, because neither the M100 nor the M170 amplitude differed for correct versus incorrect trials on house stimuli. Thus, our data argue strongly that the categorization of a stimulus as a face begins substantially earlier than the identification of the particular face.

Are these two stages—categorization and identification—simply different points on a continuous spectrum of discrimination, with cruder discriminations occurring at earlier latencies and finer discriminations occurring later, perhaps as initially coarse neural population codes get sharpened over time^{34,35}? Consistent with this hypothesis, the M170 shows stronger face selectivity than the M100. However, this hypothesis predicts that the rank ordering of preferred stimuli must be the same for the M100

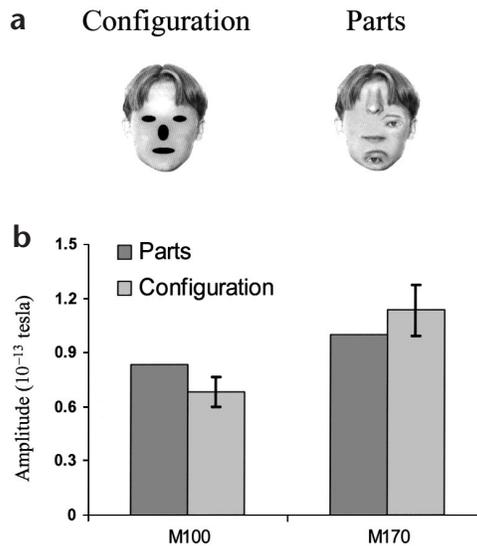


Fig. 5. Face configuration versus face parts. (a) Example stimuli. (b) Amplitudes of the M100 and the M170 response, averaged across subjects, to configuration and parts stimuli at predefined sensors of interest.

and the M170. Contrary to this prediction, the M100 showed a stronger response to stimuli depicting face parts than face configurations, whereas the M170 showed the opposite response profile (experiment 5). If neural populations simply sharpened the selectivity of their response over time, this preference reversal would not be seen. Instead, our data suggest that qualitatively different information is extracted from faces at 100 ms versus 170 ms after stimulus onset. Finally, the observed change in response profile cannot be easily explained in terms of a progression from coarse/global information to fine/local information or in terms of a progression from less to more clear face features. Instead, information about relatively local face parts is more important in determining the M100 response, whereas information about relatively global face configurations is more important in the later M170 response. Thus the most natural account of our data is that the M100 and the M170 reflect qualitatively distinct stages of face perception: an earlier stage that is critical for categorizing a stimulus as a face, which relies more on information about face parts, and a later stage that is critical for identifying individual faces, which relies more on information about face configurations.

Will the findings reported here hold for stimulus classes other than faces? Given the numerous sources of evidence that faces are ‘special’³⁶, we cannot simply assume that they will. Unfortunately, we cannot run experiments comparable to those reported here on other stimulus categories, because we have not found MEG markers selective for other categories. However, recent behavioral experiments suggest that the stages of processing reported here for face recognition will generalize to the visual recognition of nonface objects as well⁶.

METHODS

MEG recordings for experiments 1–3 were made using a 64-channel whole-head system with SQUID-based first-order gradiometer sensors (Kanazawa Institute of Technology MEG system at the KIT/MIT MEG Joint Research Lab at MIT); experiments 4 and 5 were run after the system was upgraded to 96 channels. Magnetic brain activity was digitized continuously at a sampling rate of 1,000 Hz (500 Hz for experiment 4) and was filtered with 1-Hz high pass and 200-Hz low-pass cutoff and a



60-Hz notch. Informed consent was obtained from all subjects, and the study was approved by the MIT Committee on the Use of Humans as Experimental Subjects (COUHES).

Experiments 1–3: the face-selective M100 response. Fifteen subjects (age range 18–40) passively viewed 100 intermixed trials of faces and houses (50 exemplars each) in experiment 1; two additional subjects' data were discarded due to self-reported sleepiness. The thirteen subjects who showed the early face preference over occipitotemporal cortex also performed a one-back task on faces and a variety of nonface objects (50 exemplars each) in experiment 2. Each image subtended $5.7 \times 5.7^\circ$ of visual angle and was presented at the center of gaze for 200 ms, followed by an ISI of 800 ms. The design for experiment 3 is described in Results.

For each subject in experiment 1, *t*-tests were conducted between the MEG responses to faces and houses at each time point (from –100 to 400 ms; 500 time points) and each sensor (64 channels) separately. SOIs were defined as those sensors where the magnetic fields evoked by faces were significantly larger than those by houses ($P < 0.05$) for at least five consecutive time points both within the time window centered at the latency of the M100 and within that of the M170. *P*-values for these SOI-defining statistics were not corrected for multiple sensors or multiple time point comparisons. All critical claims in this paper are based on analyses of the average responses over these sensors in independent data sets, and thus require no correction for multiple spatial hypotheses.

The peak amplitude of the M100 (maximum deflection) was determined for each stimulus type in each hemisphere within a specified time window (width >40 ms) for each subject individually. Because there was no main effect of hemisphere ($P > 0.05$) and no interaction of condition by hemisphere ($F < 1$), in subsequent analyses the data from the left and right hemisphere were averaged within each subject (after flipping the sign of the data from the right hemisphere to match the polarities).

Experiment 4: categorization versus identification. Ten subjects (age range 22–32) participated in experiment 4. The MEG recordings were preceded by a training session (<10 min) and then a psychophysical staircase adjustment session conducted in the MEG room. MEG data from two additional subjects were discarded, one because of performance that was more than two standard deviations below the mean, the other because of polarities of both M100 and M170 that were reversed compared to all other subjects (although including the data from this subject did not change the pattern or significance of the results).

Noise images were generated by inverse Fourier transformation of the mean amplitude spectra with randomized phase spectra^{23,24,37}. Intermediate images containing $x\%$ phase spectra of original images and $(100 - x)\%$ random phase spectra were generated using linear interpolation (phase spectra levels of 0% and 90% along with categorization and identification thresholds). This procedure ensured that all images were equated for spatial frequency, luminance and contrast.

Analyses were done on only the subset of data for which subjects responded both correctly and incorrectly to an identical stimulus. That is, for each stimulus, equal numbers of successful and unsuccessful trials were chosen, and the extra trials were omitted from the analysis from whichever condition had more trials. In particular, because there were more correct than incorrect trials, each incorrect trial was paired with the temporally closest correct trial. This analysis was conducted for each stimulus, guaranteeing that average MEG responses on successful and unsuccessful trials were derived from identical stimuli. Finally, the MEG recordings were averaged across stimuli separately for successful and unsuccessful trials. Note that the trials used to generate the waveform for face categorization were only selected from the MEG responses to those stimuli degraded to 'categorization thresholds', and the trials used to generate the waveform for face identification were only selected from the MEG responses to those stimuli degraded to 'identification thresholds'. The same held for house categorization and identification. The exact number of success and failure trials for each task varied across subjects, but ranged from 20 to 30 trials each for successful and unsuccessful categorization and from 15 to 20 trials for successful and unsuccessful identification.

Experiment 5: face configuration versus face parts. Two stimulus categories were constructed from veridical faces. In 'configuration' stimuli,

face parts in each veridical face were replaced by solid black ovals in their corresponding locations, whereas for 'parts' stimuli, the face parts were kept intact but were rearranged into a novel nonface configuration. The size of the black ovals in the configuration stimuli was matched to the actual size of corresponding face parts, and the arrangements of nonface configurations varied across all exemplars of parts stimuli. Photographs of these stimuli were presented at the center of gaze for 200 ms with an 800-ms ISI. Fourteen subjects (age range 18–41) passively viewed 100 intermixed trials of each stimulus category (50 exemplars each); three additional subjects' data were discarded due to lack of a face-selective MEG response in the independent localizer scan.

Acknowledgments

We thank J. Sadr and P. Sinha for helpful discussions on their RISE technique, M. Eimer, S. Hillyard, A. Marantz, M. Valdes-Sosa, P. Downing, W. Freiwald, K. Grill-Spector, Y. Jiang and the rest of the Kanwisher Lab for comments on the manuscript. Supported by the Reed Fund and National Eye Institute grant (EY13455) to N.K.

Competing interests statement

The authors declare that they have no competing financial interests.

RECEIVED 4 APRIL; ACCEPTED 30 JULY 2002

- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Braem, P. Basic objects in natural categories. *Cognit. Psychol.* 8, 382–349 (1976).
- Jolicoeur, P., Gluck, M. A. & Kosslyn, S. M. Pictures and names: making the connection. *Cognit. Psychol.* 16, 243–275 (1984).
- Sugase, Y., Yamane, S., Ueno, S. & Kawano, K. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* 400, 869–873 (1999).
- Tanaka, J. W. & Taylor, M. Object categories and expertise: is the basic level in the eye of the beholder? *Cognit. Psychol.* 23, 457–482 (1991).
- Tanaka, J. W. The entry point of face recognition: evidence for face expertise. *J. Exp. Psychol. Gen.* 130, 534–543 (2001).
- Grill-Spector, K. & Kanwisher, N. Common cortical mechanisms for different components of visual object recognition: a combined behavioral and fMRI study. *J. Vision* 1, 474a (2001).
- Bentin, S., Allison, T., Puce, A., Perez, E. & McCarthy, G. Electrophysiological studies of face perceptions in humans. *J. Cogn. Neurosci.* 8, 551–565 (1996).
- Jeffreys, D. A. Evoked potential studies of face and object processing. *Vision* 3, 1–38 (1996).
- Sams, M., Hietanen, J. K., Hari, R., Ilmoniemi, R. J. & Lounasmaa, O. V. Face-specific responses from the human inferior occipito-temporal cortex. *Neuroscience* 77, 49–55 (1997).
- Liu, J., Higuchi, M., Marantz, A. & Kanwisher, N. The selectivity of the occipitotemporal M170 for faces. *Neuroreport* 11, 337–341 (2000).
- Bruce, V. & Young, A. Understanding face recognition. *Br. J. Psychol.* 77 (Pt 3), 305–327 (1986).
- Eimer, M. The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport* 11, 2319–2324 (2000).
- Seeck, M. *et al.* Evidence for rapid face recognition from human scalp and intracranial electrodes. *Neuroreport* 8, 2749–2754 (1997).
- Schendan, H. E., Ganis, G. & Kutas, M. Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology* 35, 240–251 (1998).
- Mouchetant-Rostaing, Y., Giard, M. H., Bentin, S., Aguera, P. E. & Pernier, J. Neurophysiological correlates of face gender processing in humans. *Eur. J. Neurosci.* 12, 303–310 (2000).
- Kawasaki, H. *et al.* Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. *Nat. Neurosci.* 4, 15–16 (2001).
- Braetigam, S., Bailey, A. J. & Swithenby, S. J. Task-dependent early latency (30–60 ms) visual processing of human faces and other objects. *Neuroreport* 12, 1531–1536 (2001).
- Streit, M., Wolwer, W., Brinkmeyer, J., Ihl, R. & Gaebel, W. Electrophysiological correlates of emotional and structural face processing in humans. *Neurosci. Lett.* 278, 13–16 (2000).
- George, N., Jemel, B., Fiori, N. & Renault, B. Face and shape repetition effects in humans: a spatiotemporal ERP study. *Neuroreport* 8, 1417–1423 (1997).
- VanRullen, R. & Thorpe, S. J. The time course of visual processing: from early perception to decision-making. *J. Cogn. Neurosci.* 13, 454–461 (2001).
- Allison, T., Puce, A., Spencer, D. D. & McCarthy, G. Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* 9, 415–430 (1999).

22. Tootell, R. B., Mendola, J. D., Hadjikhani, N. K., Liu, A. K. & Dale, A. M. The representation of the ipsilateral visual field in human cerebral cortex. *Proc. Natl. Acad. Sci. USA* **95**, 818–824 (1998).
23. Sadr, J. & Sinha, P. Exploring object perception with Random Image Structure Evolution. *MIT Artif. Int. Lab. Memo*, 2001–6 (2001).
24. Sadr, J. & Sinha, P. Object recognition and random image structure evolution. *Cognit. Sci.* (in press).
25. Watson, A. B. & Pelli, D. G. Quest: a Bayesian adaptive psychometric method. *Percept. Psychophys.* **33**, 113–120 (1983).
26. Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *Nature* **381**, 520–522 (1996).
27. Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J. & Lounasmaa, O. V. Magnetoencephalography: theory, instrumentation and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* **65**, 413–497 (1993).
28. Bentin, S. in *Encyclopedia of Cognitive Science* (ed. Nadel, L.) Neural basis of face perception. (Macmillan, London, in press).
29. Oram, M. W. & Perrett, D. I. Time course of neural responses discriminating different views of the face and head. *J. Neurophysiol.* **68**, 70–84 (1992).
30. Rolls, E. T. Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Hum. Neurobiol.* **3**, 209–222 (1984).
31. Thorpe, S. J. & Fabre-Thorpe, M. Seeking categories in the brain. *Science* **291**, 260–263 (2001).
32. Gomez Gonzalez, C. M., Clark, V. P., Fan, S., Luck, S. J. & Hillyard, S. A. Sources of attention-sensitive visual event-related potentials. *Brain Topogr.* **7**, 41–51 (1994).
33. Thorpe, S. & Imbert, M. *Connectionism in Perspective* (eds. Pfeifer, R., Schreter, Z., Fogelman-Soulie, F. & Steels, L.) 63–92 (Elsevier, Amsterdam, 1989).
34. Kovacs, G., Vogels, R. & Orban, G. A. Cortical correlate of pattern backward masking. *Proc. Natl. Acad. Sci. USA* **92**, 5587–5591 (1995).
35. Keysers, C., Xiao, D. K., Foldiak, P. & Perrett, D. I. The speed of sight. *J. Cogn. Neurosci.* **13**, 90–101 (2001).
36. Farah, M. J. *Visual Cognition* (eds. Kosslyn, S. M. & Osherson, D. N.) 101–119 (MIT Press, Cambridge, Massachusetts, 1995).
37. Rainer, G., Augath, M., Trinath, T. & Logothetis, N. K. Nonmonotonic noise tuning of BOLD fMRI signal to natural images in the visual cortex of the anesthetized monkey. *Curr. Biol.* **11**, 846–854 (2001).

