Levels of categorization in visual recognition studied using functional magnetic resonance imaging

Isabel Gauthier*, Adam W. Anderson†, Michael J. Tarr‡, Pawel Skudlarski† and John C. Gore†

Background: Recent functional neuroimaging results implicate part of the ventral temporal lobe of the brain in face recognition, and have, together with neurophysiological findings, been used as evidence for a face-specific neural module in the brain. Experimental designs, however, have often failed to distinguish between the class of the object used as the stimulus (face or non-face) and the level of categorization at which the stimulus is recognized (the 'basic' level, such as 'bird', at which familiar objects are first recognized, or more subordinate levels – 'sparrow', for example – which require additional perceptual processing). We have used echo-planar functional magnetic resonance imaging to compare brain activation for the matching of non-face objects with subordinate-level and basic-level descriptors.

Results: The additional visual processing required to verify the subordinate level of a picture over its basic level was associated with activation of the fusiform and inferior temporal gyri (FIT) as well as the temporal poles. These areas correspond closely to those previously implicated in the processing of facial images.

Conclusions: Our results indicate that areas of the ventral visual pathway that have been associated with face recognition are sensitive to manipulations of the categorization level of non-face objects. This idea offers an alternative to the dominant view that FIT may be organized according to conceptual categories, and our results establish the importance of manipulating task requirements when evaluating a 'neural module' hypothesis.

Addresses: *Department of Psychology, Yale University, PO Box 208205, New Haven, Connecticut 06520-8205, USA. †Department of Diagnostic Radiology, Yale University School of Medicine, PO Box 208042, New Haven, Connecticut 06520-8042, USA. ‡Department of Cognitive and Linguistic Sciences, Brown University, Box 1978, Providence, Rhode Island 02912, USA.

Correspondence: Isabel Gauthier
E-mail: isabel.gauthier@yale.edu

Background

The neural processes that underlie recognition of a face, rather than another object, could be special in at least two ways: they may require unique perceptual processing and/or engage a specific region of the brain [1,2]. Several behavioral studies suggest, however, that a common mechanism is used to process faces and non-face objects when experimental conditions are carefully matched [3–5]. In contrast, neuroimaging [6–11] and neuropsychological [11–14] studies point to an area of the ventral temporal lobe, including parts of the fusiform and inferior temporal gyri (FIT), as more specialized for face than non-face object recognition. The experimental designs used in such studies, however, often failed to control for the level of categorization at which a stimulus is recognized (basic, such as 'bird', or subordinate such as 'sparrow'), a factor that can potentially be confounded with the effects of the class of the stimulus (such as face or non-face). We argue that investigating the effects of independently manipulating the categorization level for non-face objects is crucial in assessing whether there is a face-specific neural module in the brain.

The study of face recognition has been closely associated with theories of modular specialization in the brain. In particular, the postulated existence of a brain area that is specialized for face recognition would provide a clear example of domain specificity, one of the defining features of cognitive modules [15]. Whether any brain area is or is not domain-specific is, however, an empirical question. Recent neuroimaging results [6–11] appear to offer compelling evidence for a face-specific module.

Upon closer examination, however, these neuroimaging studies reveal a bias towards experimental methods that can only lead to conclusions favoring domain-specific modules. For example, Kanwisher and colleagues [7,8] have performed several neuroimaging studies comparing passive viewing of faces with viewing of non-face objects. This comparison between different stimulus classes was argued to be preferable to task manipulation within a class, because the recognition of faces at the individual level is thought to be highly automatic. Indeed, facial recognition of individuals is a task that we need to perform often, so there are reasons why the processing of facial identity should be automatized (see [16] for behavioral evidence). In contrast, common objects are typically recognized first at a more general level (termed the 'basic' level) that is more appropriate to everyday functional

needs [17–19] (e.g., chair or dog). Access to more subordinate levels of categorization (e.g., folding chair or Dalmatian) may require additional perceptual processing above and beyond that required for basic-level access [4,17–19]. Consequently, common non-face objects may be automatically recognized at a more general categorical level than faces, unless subjects are required to perform subordinate-level judgments (as in the present study) or unless they are experts with a particular class of object [19]. If subjects have developed expertise at processing faces at the subordinate level, they may not perform equivalent processing of non-face objects, especially if given no particular instructions or task (or, as in some studies, when animals are tested under anesthetic [20]).

The results of recent imaging experiments that compared the processing of faces to that of various non-face object categories [8] have been described as evidence for 'special-purpose cortical machinery' for face recognition (evidence from single-cell recording and other approaches will be considered in the Discussion). It is our contention that such studies failed to consider experimental designs that might refute this conclusion. For instance, no attempts were made to engage the face-specific module using only non-face objects. Faces and non-face objects differ along many dimensions (e.g., shared configuration of parts, number of known exemplars, social importance, expertise of subjects). Attempting to equate two sets of stimuli on any single dimension can never ensure that such sets do not differ in other ways. We suggest, therefore, that no experiment restricted to a comparison between different stimulus classes can provide definitive evidence for a face-specific neural module (a criticism that extends to all techniques, including single-cell recording).

The present study examines the hypothesis that there is a face-specific neural substrate, that is, an area of the ventral temporal lobe, including parts of the FIT, that appears to respond more strongly to faces than to other objects, because they are faces [6–11]. It is worth noting that we are testing the face-specificity hypothesis at the scale typically used in neuroimaging and neuropsychological studies, not at the finer anatomical resolutions available using single-cell recording or optical imaging. At the scale used in neuroimaging and neuropsychological studies, there is already some evidence arguing against a face-specific module: patients with brain lesions of the ventral temporal cortex who demonstrate a face-recognition deficit (prosopagnosia) often have more general problems with subordinate-level processing of non-face objects [12,21–24]. Thus, it is somewhat surprising that previous results using positron emission tomography (PET) have failed to find inferior temporal lobe activation for subordinate processing of non-face objects (over and above that observed for basic-level processing [25]; similar results

have been obtained using event-related potentials (Jim Tanaka, personal communication).

Using functional magnetic resonance imaging (fMRI) and improving on several aspects of the stimuli and design used in a recent study [25], we tested whether the putative face-specific module could be engaged by manipulating categorization level, using only non-face objects. A recent PET study [25] examined object verification at the subordinate and basic levels but failed to reveal inferior temporal activation when basic processing was subtracted from subordinate processing. The study was different in several ways from the current experiment and potential reasons for its failure to activate FIT include the use of line drawings that may spare subjects from the fine-grained processing necessary for subordinate judgments on more complex pictures, a larger number of basic than subordinate judgments over the same time course, and priming for words used in both the baseline and experimental tasks. Evidence that the putative face-specific module is engaged when only the categorization level is manipulated would suggest, first, that the face-specific area as defined by a group-average method may not be specific to faces and, second, that the results of any study that confounds comparison of faces and non-face objects with the level of categorization cannot be unequivocally interpreted as evidence for a face-specific module.

Results

Eight subjects performed two tasks — a visual task and a semantic task — at the basic and subordinate levels. In the visual task, subjects judged whether a picture matched a simultaneously presented word, and in the semantic task, subjects decided if the object described could move by its own power (Figure 1). All tasks required lexical access, either for basic-level or subordinate-level words. We were primarily interested in visual recognition processes, so we restricted our analyses to the two slices of the brain that contain the ventral part of the temporal lobe. Different regions of the inferior temporal lobe (particularly on the left) have been found to be important for lexical retrieval for persons, animals and tools [26]. To minimize the effect of lexical retrieval, two different lists of words were generated for the visual and semantic tasks: words were matched in average word length, word frequency and response times in the semantic task (in a separate pilot study), and both lists contained similar numbers of animals (17 for the visual task, 21 for the semantic task). The other items were common objects such as vehicles and pieces of furniture. Crucially, our design compared subordinate-level with basic-level judgments for the same items (in particular, in the visual task, the same picture was associated with both a basic-level and a subordinate-level label), such that the effects could not be explained by stimulus class differences (although the results could potentially differ between conceptual categories).
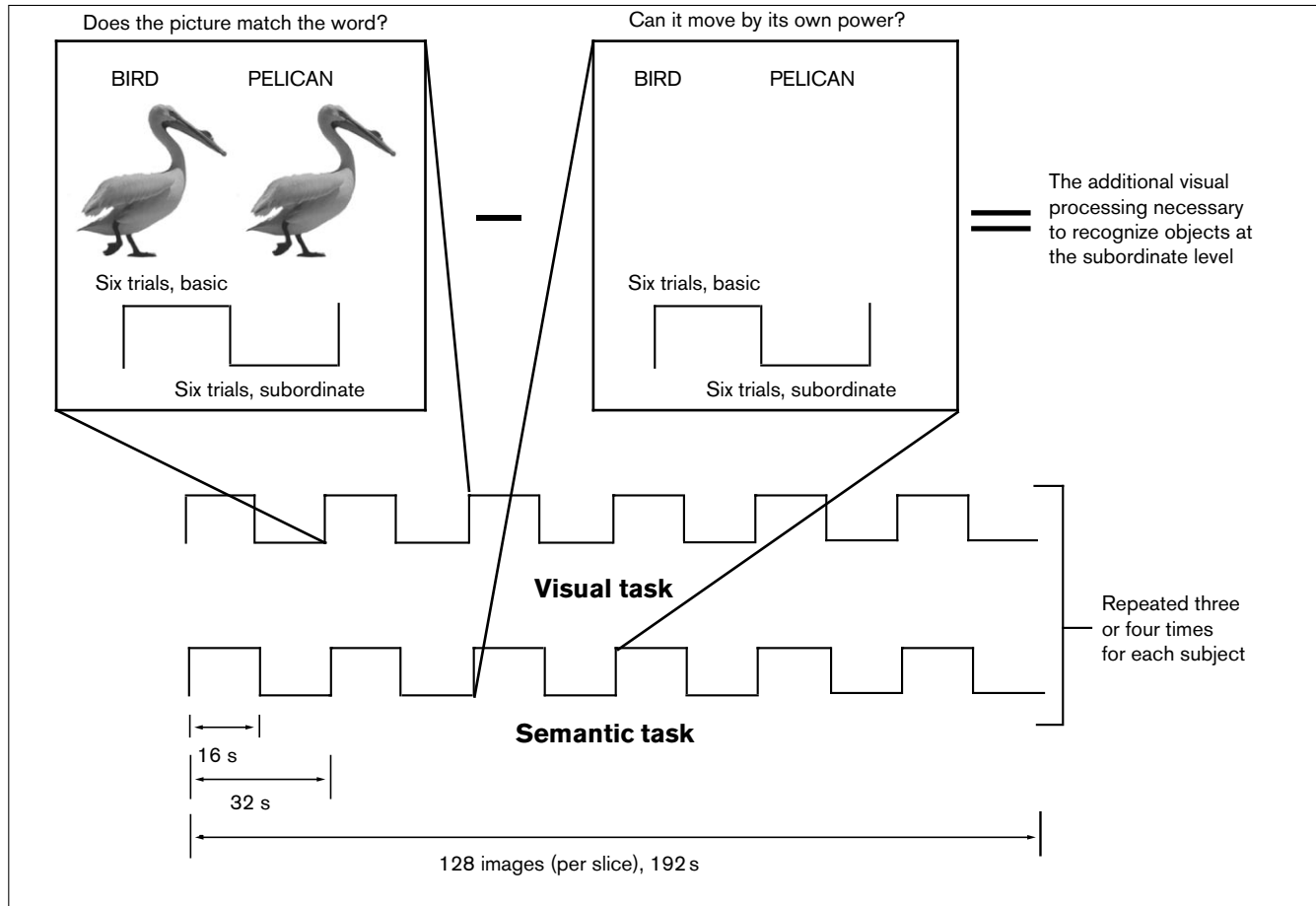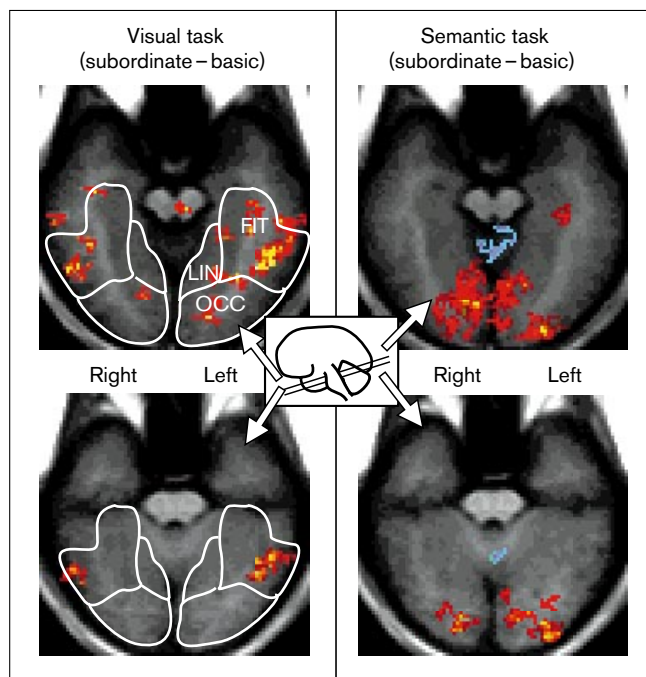
Figure 1

Illustration of task procedure. For both tasks (visual and semantic), the stimuli (a word or a word and a picture) remained on the screen for 2,000 ms with an intertrial interval of 750 ms. Both tasks shared the requirements of reading and accessing the meaning of the word and the response components (responding 'yes' or 'no' by pressing one of two buttons). Trials were randomized for each subject. In the visual task, 72 pictures were used, each being repeated four times (twice at each level with matching and non-matching labels).

By subtracting the basic from the subordinate neural activation in subjects performing the visual task, we isolated the activation associated with subordinate-level visual recognition of objects plus any additional visual, lexical, or semantic processing of the words engaged by the subordinate level (as compared to the basic level). This revealed activation in FIT for seven out of eight subjects (six had bilateral activation and one had activation only on the left side; Figure 2). The second most activated area was the occipital lobe (OCC), although it was activated in only four subjects (all on the left side). By subtracting basic from subordinate activation for the semantic task, we were able to isolate the activation associated with any visual, lexical or semantic processing of the words, as well as possible visual imagery used to perform the semantic task, for subordinate-level over basic-level words. This revealed activity in OCC for seven subjects (two bilaterally and five only on the left). The second most activated area was FIT,

which was activated in six subjects (two bilaterally, two on the left and two on the right).
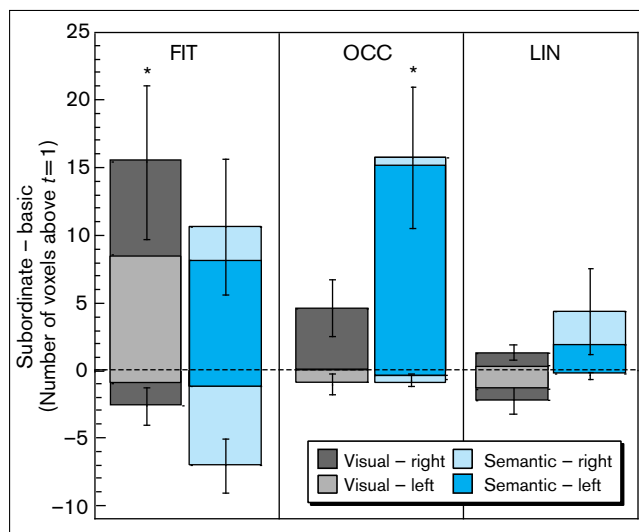
A statistical analysis of variance (ANOVA) conducted on the volume (measured in voxels) activated above threshold in each region of interest (ROI) across subjects (see Materials and methods for design) revealed statistically significant evidence for a two-way interaction between the ROI (OCC and FIT) and the task ($F(1,7) = 6.2$, $p < 0.05$; Figure 3). Post hoc tests indicated that OCC showed a larger volume of activation for the semantic than the visual task ($p < 0.05$) whereas FIT showed no task difference. A potential explanation for the greater OCC activation in the semantic task is that subjects may have used a visual imagery strategy. In that case, subordinate-level words could lead to relatively more early visual cortex activation [27] because they specify more detailed visual information. We also obtained a 'main effect of hemisphere'

Figure 2

Distribution, averaged across eight subjects, of positive and negative neural activation during the visual and semantic tasks for the 'basic subtracted from subordinate' comparison ($t = \pm 0.2$; cluster size = 11). The figure includes a schematic representation of the regions of interest: FIT, lingual gyri (LIN) and occipital cortex (OCC). Yellow corresponds to the highest activation satisfying the activation criteria. A lower threshold is used here than for the double subtraction (although the region of interest analysis is based on the same threshold for all comparisons) in order to illustrate the absence of negative FIT activation in the semantic task, which could have produced some positive FIT activation in the double subtraction.

Figure 3

Mean activation level (number of voxels above threshold; $t = \pm 1$; cluster size = 3) in the three regions of interest – FIT, OCC and lingual gyri (LIN) – for the visual and semantic tasks. Asterisks indicate the regions of interest (FIT and OCC) in which the volume of activation above threshold was significantly different from zero across subjects (by t tests, $p < 0.05$). Note that the experimental design predicted only positive activation (because subordinate-level access is thought to be inclusive of basic-level access). Indeed, volumes of negative activation were small and not significantly activated across subjects.

($F(1,7) = 7.1$, $p < 0.05$), qualified by an interaction with the ROI ($F(1,7) = 5.89$, $p < 0.05$). Scheffé tests ($p < 0.05$) indicated larger volumes of activation in the left than the right hemisphere, no difference between the two regions of interest in the left hemisphere and a larger volume of activation in FIT than OCC in the right hemisphere. We will not attempt to interpret hemispheric differences since our subject sample was highly heterogeneous in terms of both hand preference and sex. It should be noted, however, that no particular pattern was found in FIT and OCC for left-handed subjects.

Although the volume of FIT activation was not significantly larger in the visual task than in the semantic task, the particular voxels activated in each task might be different. To test whether the additional visual processing required in accessing the subordinate level of a picture activates the area of the brain determined in earlier studies to be face-specific [6–11,28], a double subtraction [visual (subordinate – basic) – semantic (subordinate – basic)] was performed. This analysis was designed to

eliminate from the results of the visual task any difference due to the visual, lexical or semantic processing resulting from the presence of the words in the visual task. As shown in Figure 4, the FIT region that was engaged by subordinate object processing in all of our eight subjects, when averaged over all subjects, is remarkably similar to the face-specific area described in previous studies [2]. In particular, it matches the region from which N200 face-specific scalp potentials have been reported in humans [28]. In addition, the averaged results revealed foci of activation in the temporal poles, a region that has previously been implicated in the processing of facial images [11,26]. An ANOVA conducted on the number of voxels in four regions of interest (OCC, FIT, lingual gyri (LIN) and temporal poles) revealed a statistically significant effect of ROI ($F(3,21) = 4.5$, $p < 0.02$) qualified by an interaction with a particular hemisphere ($F(3,21) = 4.5$, $p < 0.02$). Scheffé tests ($p < 0.05$) indicated larger volumes of activation bilaterally in FIT compared to OCC and temporal poles, while FIT was larger than LIN only in the right hemisphere. Other regions of interest did not otherwise differ within hemispheres.

Counting activated voxels in regions of interest may not be ideal for comparing different regions of the brain (even when these regions do not differ in size, as they may differ in many other ways), although it allows for some individual

Figure 4
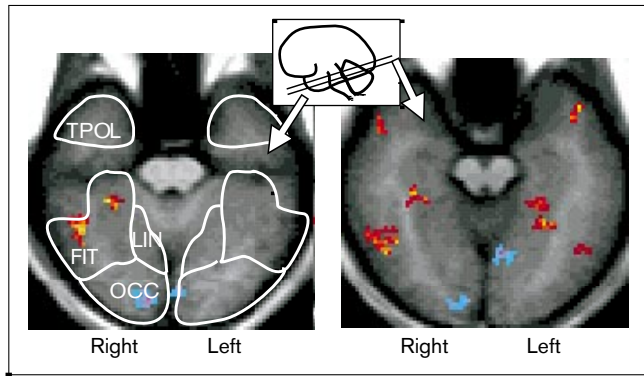
Distribution, averaged across eight subjects, of positive and negative neural activation in the [visual (subordinate − basic) − semantic (subordinate − basic)] double subtraction ($t = \pm\, 0.3$; cluster size = 11). As well as the three regions of interest depicted in Figure 2, the figure includes a schematic representation of a fourth ROI (temporal poles, TPOL) included in the analysis for the double subtraction.

differences in the localization of a 'subordinate' area. On the other hand, averaged activation maps are more sensitive to spatial inconsistencies between subjects. Here, the results using either of the two methods supported our hypothesis that the additional visual processing required to access the subordinate level of a non-face object, over and above its basic level, engages the region of the brain previously defined as the face-selective area. This result appears to be consistent with those from other studies that have implicated the fusiform gyri in non-face pattern processing [29,30]. As mentioned previously, the temporal poles area, which was activated in seven out of eight subjects (five bilaterally, one on only the left and one on only the right) has also been previously associated with the processing of faces [11,26]. It has been suggested that the right temporal pole may be involved in the recognition of unique entities, and that part of the left temporal pole may be engaged by lexical processing in general [26]. Our results do seem to indicate a general role for the temporal poles in subordinate-level processing, but our design was not expected to isolate lexical processes. The left activation obtained in this area could have been due in part to the left-handed subjects (all three had bilateral activation in the temporal poles for the double subtraction).

Discussion

The term modularity — describing the idea that the brain contains areas, or modules, that are specialized for particular tasks — has been used in different ways, only some of which are addressed by the present study. The more extreme (systems-level) view of modularity (sometimes referred to as 'Fodorian') has typically been promoted using neuropsychological and neuroimaging evidence. According to this view, the face recognition system is functionally independent and physically distinct from the non-face object recognition system (leading to face-specific recognition deficits after fairly large lesions of the temporal lobe) and processes visual information in a qualitatively different fashion [7–9,13–15,28,31]. The face-specific module is believed to operate rapidly and mandatorily upon the presentation of a face, and its development is thought to be largely endogenously determined. This is the view, quite popular in the literature [8,13,15,31], targeted by our study, as well as by other studies that have concluded that faces may not be 'special' in that sense [3–5,12,21–24].

A second, less extreme, use of the term modularity that appears widely in the physiology literature is more restricted and describes the spatial organization of neurons with similar response properties. Thus, the inferior temporal lobe appears to consist of small columns, about 0.4 mm in width, each containing cells that respond to similar complex visual features, such as faces [20,32–35]. Technical and methodological limitations, however, prevent any strong claim about the specificity of these cell assemblies. Indeed, recent findings suggest that the neurons in visual areas such as V2, V4 and MT that were once thought to be specific feature detectors may instead be general-purpose analyzers [36]. Moreover, it is unlikely that any one of these columns is sufficient for the recognition of all exemplars of an object class, such as faces, or that a single column can be selectively damaged by a stroke. Whether these columns can be interpreted as 'modules' in the Fodorian sense will depend on a number of unresolved issues, including the type and extent of interactions between neighboring columns.

In contrast, this second view of modularity is entirely compatible with the existence of a large network of detectors of complex visual features that can be tuned to the properties of any class of object depending on experience and task requirements. Indeed, several physiologists have explicitly suggested that the cells that are specifically engaged by faces are not otherwise unique and that their responses may be explained by our experience of faces [20,34,37]. An important empirical demonstration comes from Logothetis and colleagues [38,39], who trained monkeys to identify novel 'paper-clip' objects at the subordinate level and found cells in the inferior temporal lobe that showed the same degree of object-selectivity and view-selectivity typically found with face-specific cells (few people would argue for a Fodorian 'paper-clip module'). The existence of clusters of cells that respond preferentially to one class of object (such as faces or paper-clip objects) is, therefore, somewhat less controversial than is the putative modularity, in the strict sense, of their processing.

The theory that there are cells specialized for the features present in faces [20,32–35] is independent from the hypothesis of a face-specific recognition module. (Indeed

one can imagine a face-specific recognition module composed of a distributed network within which none of the cells behaves like typically described face-specific cells.) Moreover, regardless of a given techniques' resolving power, a more extreme modular hypothesis that incorporates face-specific processing should be tested using task manipulations with non-face objects — the only means by which such a hypothesis can possibly be refuted. Our results indicate that FIT as well as the temporal poles can be selectively engaged by subordinate-level judgments of common non-face objects. Since the same pictures were used in the basic-level and subordinate-level conditions for the visual task, our results are unlikely to be due to physical or experiential differences between stimuli. They also cannot be due to the task instructions, as these were held constant for both levels of categorization. Such factors, as well as the processing requirements tied to different categorization levels, are often confounded with stimulus category in studies in which the putative face-specific area is engaged.

Our interpretation of previous group-averaged data that have been taken as evidence for a face-specific neural module is that the typical activation obtained in this area of the brain may arise in part from a difference in categorization level. We are not arguing, however, that this factor alone is responsible for all dissociations between faces and non-face objects: on the contrary, we hope that future neuroimaging studies will adopt designs that manipulate other important differences between faces and non-face objects, such as the level of expertise [3–5], the size of the class formed by exemplars with similar shapes, and the social or personal value of the exemplars to perceivers [23]. Passive or even anesthetized viewing of faces and non-face objects, however, may not be appropriate to test the modularity hypothesis because subordinate-level categorization may be especially automatized for faces. Finally, because the face-sensitive portion of the ventral cortex in any subject represents only a small portion of a group-averaged face-specific area, further studies should compare the face-specific and subordinate-level areas within individual subjects. Our findings have direct implications for other studies that have used the group-average method and/or compared faces with non-face control stimuli and suggest that part of the ventral temporal cortex may be organized along general dimensions of visual recognition (such as the level of categorization) rather than along class boundaries.

Materials and methods

Subjects
Eight neurologically normal subjects (four males and three females, five right handed and three left handed) took part in the study (approved by the Yale University Human Investigation Committee). All subjects were healthy and were not on any medication.

Behavioral studies
In order to select the stimuli, behavioral studies were conducted on a Macintosh computer using the same ABAB design as the fMRI study.

In the first pilot study, 20 subjects performed the visual task at both the basic and subordinate levels for 120 shaded grey-scale pictures of common objects. We selected 72 pictures, each showing a basic-level advantage in response time (mean response times for selected pictures were 995 ms for the subordinate level and 849 ms for the basic level, $t = 10.6$, $p < 0.001$) and 80% or higher accuracy at both levels. The mean accuracy for the two levels also differed reliably for selected items (97% for the basic level, 88% for the subordinate level, $t = 8.6$, $p < 0.01$). In the second pilot study, 20 subjects performed the semantic task at both the basic and subordinate levels for 240 labels (including all the labels matching the pictures selected in pilot study 1). Results were used to produce two different lists of words for the visual and semantic tasks – these lists were matched in mean word length, word frequency, and response times in the semantic task. The semantic task also produced a reliable basic-level advantage in response time (744 ms for the basic level, 800 ms for the subordinate level, $t = 2.2$, $p < 0.03$) and accuracy (90% for the basic level, 86% for the subordinate level, $t = 2.2$, $p < 0.03$). Note that this difference in difficulty between levels is an integral part of their definition.

fMRI scan acquisition and analysis
Imaging was performed on a 1.5 Tesla Signa scanner (GE Medical Systems), with Instascan echo planar imaging capabilities (Advanced NMR Systems). A single shot, gradient echo, echo planar pulse sequence was used to acquire images (64×128 voxels) over a field of view of 20×40 cm. The imaging parameters were echo time 60 ms, repetition time 1,547 ms, flip angle 60 degrees, slice thickness 7 mm and slice spacing 0 mm. During each repetition time interval, six axial-oblique slices were imaged . Each run produced 128 images per slice. The first two images (per slice) were discarded to decrease the effect of non-steady state longitudinal magnetization. Changes in image intensity were analyzed on a voxel-by-voxel basis: voxels were considered activated if the t-values in both halves of the bisected dataset were equal or higher than one and if they were contiguous with a set of at least two voxels also above threshold. In a separate study, simulated activation was added to resting state data in order to estimate receiver operating characteristic (ROC) curves for our method of analysis [40]. Conservative estimates indicate that 2% of voxels are incorrectly identified as activated whereas 76% of activated voxels are correctly identified. Image data were corrected for motion using the SPM 96 software (Wellcome Department of Cognitive Neurology, London). The t-maps were superimposed on T1 weighted anatomical images of the corresponding slice.

fMRI averaging
For averaging purposes, anatomical landmarks were defined in T1 weighted axial-oblique and midline sagittal images of each subject (they consisted, in the oblique plane, of the outer edges of the brain, the optic chiasma and the anterior edge of the cerebral aqueduct, and in the sagittal plane, of the most dorsal point of the thalamus, the optic chiasma, inferior colliculi and most posterior point of the fourth ventricle). Functional and anatomical images were transformed by piece-wise linear warping in 12 brain subvolumes to register the results for each subject in a common coordinate system. The transformed functional maps were then averaged across subjects and thresholded to obtain subject-average functional maps. The double subtraction was first performed on each subject's data, and the resultant subject maps were then averaged together.

Region of interest analysis
ROIs were selected on the basis of previous neuropsychological and neuroimaging results: they included the area described previously as face-selective in humans [6–14], which includes part of FIT, LIN – which has a spatial, but not functional, relationship with the face-specific area [6,20] – and OCC, a portion of the occipital cortex including part of the striate cortex. The two slices that included the most ventral temporal cortex were selected and the regions of interest were defined for each subject according to anatomical landmarks within each hemisphere. OCC was defined as the occipital cortex from the back of the

brain to a line drawn between the temporo-occipital incisure laterally and the lingual sulcus medially. LIN was defined as the region anterior to OCC, medial to the collateral sulcus and posterior to the parahippocampal gyrus. FIT was defined as the region anterior to OCC, lateral to the collateral sulcus, including the posterior aspect of the inferior temporal gyrus and the fusiform gyrus, going as anterior as the head of the hippocampus but not including any hippocampus or parahippocampal gyrus. Temporal poles included all the superior and middle temporal gyrus cortex visible in the two slices selected. A $2 \times 2 \times 2$ ANOVA conducted on the simple subtractions included the following factors: ROI (OCC and FIT), hemisphere (left and right) and task (visual and semantic). LIN was dropped from the analysis as it did not reliably activate in any task and showed significantly less activation than the other regions of interest. A 2×4 ANOVA was conducted for the double subtraction and included the following factors: hemisphere (right and left) and ROI (OCC, temporal poles, FIT, LIN). Temporal poles were included on the basis of the averaged results for the double subtraction which indicated activation in that region. The dependent measure for both ANOVAs was the number of voxels activated above $t = 1$. The selection of a t-value for the ROI analysis is essentially arbitrary and $t = 1$ was selected in order to be conservative and to ensure that the distribution of numbers entering the ANOVA was not overly skewed.

Acknowledgements

References

1. Hay DC, Young AW: **The human face.** In *Normality and Pathology in Cognitive Function.* Edited by Ellis AW. New York: Academic Press; 1982:173-202.
2. Tovée MJ: **What are faces for?** *Curr Biol* 1995, **5**:480-482.
3. Diamond R, Carey S: **Why faces are and are not special: an effect of expertise.** *J Exp Psychol [Gen]* 1986, **115**:107-117.
4. Gauthier I, Tarr MJ: **Becoming a "Greeble" expert: exploring mechanisms for face recognition.** *Vision Res* 1997, **37**:1673-1681.
5. Rhodes G, McLean IG: **Distinctiveness and expertise effects with homogeneous stimuli: towards a model of configural coding.** *Perception* 1990, **19**:773-794.
6. Haxby JV, Horwitz B, Ungerleider LB, Maisog JM, Pietrini P, Grady CL: **The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations.** *J Neurosci* 1994,**14**:6336-6353.
7. Kanwisher N, Chun MM, McDermott J, Ledden P: **Functional imaging of human visual recognition.** *Cogn Brain Res* 1996, **5**:55-67.
8. Kanwisher N, McDermott J, Chun MM: **The fusiform face area: a module in human extrastriate cortex specialized for face perception.** *J Neurosci* 1997, **17**:4302-4311.
9. Puce A, Allison T, Gore JC, McCarthy G: **Face-sensitive regions in human extrastriate cortex studied by functional MRI.** *J Neurophysiol* 1995, **74**:1192–1199.
10. Sergent J, Ohta S, MacDonald B: **Functional neuroanatomy of face and object processing.** *Brain* 1992, **115**:15-36.
11. Sergent J, Signoret JL: **Functional and anatomical decomposition of face processing: evidence from prosopagnosia and PET study of normal subjects.** *Philos Trans R Soc Lond [Biol]* 1992,**335**:55-62.
12. Damasio AR, Damasio H, Van Hoesen GW: **Prosopagnosia: anatomic basis and behavioral mechanisms.** *Neurology* 1982, **32**:331-341.
13. Farah MJ, Wilson KD, Drain HM, Tanaka JW: **The inverted face inversion effect in prosopagnosia: evidence for mandatory, face-specific perceptual mechanisms.** *Vision Res* 1995, **35**:2089-2093.
14. McNeil JE, Warrington EK: **Prosopagnosia, a face-specific disorder.** *Q J Exp Psychol* 1993, **46A**:1-10.
15. Fodor J: *Modularity of Mind.* Cambridge: MIT Press; 1983.
16. Tanaka JW, Gauthier I: **Expertise in object and face recognition.** *Mech Percept Learning* 1997, **36**:83-125.
17. Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P: **Basic objects in natural categories.** *Cogn Psychol* 1976, **8**:382-439.
18. Jolicoeur P, Gluck MA, Kosslyn SM: **Pictures and names: making the connection.** *Cogn Psychol* 1984, **16**:243-275.
19. Tanaka JW, Taylor M: **Object categories and expertise: is the basic level in the eye of the beholder?** *Q J Exp Psychol* 1991, **46A**:225-245.
20. Tanaka K, Saito H, Fukada Y, Moriya M: **Coding visual images of objects in the inferotemporal cortex of the macaque monkey.** *J Neurophysiol* 1991, **66**:170-189.
21. Damasio AR: **Prosopagnosia.** *Trends Neurosci* 1985, **8**:132-135.
22. Damasio AR: **Category-related recognition defects as a clue to the neural substrates of knowledge.** *Trends Neurosci* 1990, **13**:95-98.
23. Damasio AR, Tranel D, Damasio H: **Face agnosia and the neural substrates of memory.** *Annu Rev Neurosci* 1990, **13**:89-109.
24. Gauthier I, Behrmann M, Tarr MJ, Anderson AW, Gore JC, McClelland J: **Subordinate-level categorization in human inferior temporal cortex: converging evidence from neuropsychology and brain imaging.** *Neurosci Abstr* 1996, **1**:10-11.
25. Kosslyn SM, Alpert NM, Thompson WL: **Identifiying objects at different levels of hierarchy: a positron emission tomography study.** *Hum Brain Mapping* 1995, **3**:107-132.
26. Damasio H, Grabowski TJ, Tranel D, Hichwa RD, Damasio AR: **A neural basis for lexical retrieval.** *Nature* 1996, **380**:499-505.
27. Kosslyn SM, Ochsner KN: **In search of occipital activation during visual mental imagery.** *Trends Neurosci* 1994, **17**:290-292.
28. Allison T, McCarthy G, Nobre A, Puce A, Belger A: **Human extrastriate visual cortex and the perception of faces, words, numbers, and colors.** *Cereb Cortex* 1994, **5**:544-554.
29. Martin A, Wiggs CL, Ungerleider LG, Haxby JV: **Neural correlates of category-specific knowledge.** *Nature* 1996, **379**:649-652.
30. Schacter DL, Reiman E, Uecker A, Polster MR, Yun LS, Cooper LA: **Brain regions associated with retrieval of structurally coherent visual information.** *Nature* 1995, **376**:587-590.
31. Farah MJ: **Is face recognition 'special'? Evidence from neuropsychology.** *Behav Brain Res* 1996, **76**:181-189.
32. Fujita I, Tanaka K, Ito M, Cheng K: **Columns for visual features of objects in monkey inferotemporal cortex.** *Nature* 1992, **36**:343-346.
33. Wang G, Tanaka K, Tanifuji M: **Optical imaging of functional organization in the monkey inferotemporal cortex.** *Science,* **272**:1665-1668.
34. Perrett DI, Oram MW: **Neurophysiology of shape processing.** *Image Vision Comput* 1993, **11**:317-333.
35. Hasselmo ME, Rolls ET, Baylis GC: **The role of expression and identity in the face selective response of neurons in the temporal visual cortex of the monkey.** *Behav Brain Res* 1989, **32**:203-218.
36. Schiller P: **On the specificity of neurons and visual areas:** *Behav Brain Res* 1996, **76**:21-35.
37. Logothetis NK, Sheinberg DL: **Visual object recognition.** *Annu Rev Neurosci* 1996, **19**:577-621.
38. Logothetis NK, Pauls J, Bülthoff HH, Poggio T: **View-dependent object recognition in monkeys.** *Curr Biol* 1994, **4**:401-414.
39. Logothetis NK, Pauls J, Poggio T: **Shape representation in the inferior temporal cortex of monkeys.** *Curr Biol* 1995, **5**:552-563.
40. Constable RT, Skudlarski P, Gore JC: **An ROC approach for evaluating functional brain MR imaging and postprocessing protocols.** *Magn Reson Med* 1995, **34**:57-64.