Human visual object recognition:
What have we learned from neuroimaging?

G. K. AGUIRRE and M. J. FARAH
Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania

We review the neuroimaging literature for studies whose designs permit the identification of regions specialized for visual recognition. The results of these 17 studies are combined and analyzed with the goal of answering the following questions: (1) How well localized are areas involved in visual recognition across subjects? (2) Are there cortical areas that are specialized for the perception of different categories of stimuli (e.g., faces, words, and general objects)? The concept of specialization is defined and examined, as are the inferential limitations of neuroimaging methodology. Local maxima across studies were poorly colocalized within posterior inferior cortex, and there was no consistent segregation of activation sites dependent on the category of stimulus used. We discuss several possible reasons why the results of this review do not agree with the predictions of lesion and neurophysiology studies.

What parts of the human brain are involved in visual object recognition? Is visual recognition a single, general-purpose system, or are there specialized subsystems for recognizing different types of visual stimuli? These long-standing questions of neuropsychology and cognitive psychology have traditionally been addressed using evidence from human brain-damaged patients and from laboratory studies of animals. With the advent of functional neuroimaging, a new source of evidence has become available. The primary goal of this article is to review the evidence from two neuroimaging techniques—positron emission tomography (PET) and functional magnetic resonance imaging (fMRI)—concerning specialization for visual object recognition in the human brain and to attempt a synthesis both among the results of neuroimaging studies and between the results of neuroimaging studies and the methods of clinical neuropsychology and animal neurophysiology. We begin by discussing what the term *specialized* means in the context of these classic neuropsychological and cognitive psychological questions and consider some issues of inference in interpreting neuroimaging data relative to these concepts. Although the meaning of the term might seem perfectly straightforward, there are nuances to the concept that bear on our discussion.

Specialization

What does it mean for a brain region to be specialized? The definition we consider here is:

A brain region is specialized for a particular cognitive process if damage to the region greatly impairs that cognitive process in at least one circumstance, but lesions of the region never cause more than mild impairments in other cognitive processes.

Several features of the definition merit discussion. First, the qualifier "in at least one circumstance" may strike the reader as odd. It is included in our definition because lesions to a specialized region might not result in any impairment if there are other, redundant regions that can support the process. Thus, a specialized region need not be strictly necessary for a cognitive process but instead causally involved. An involved region is one that, under at least one circumstance, is necessary (Zarahn, 1998). Such a circumstance might be the case in which the other, parallel regions are themselves damaged. In this setting, the region is now truly necessary for the cognitive process, since no other region exists to support the process in its absence. A second feature of the definition is that under no circumstances does damage to a specialized region greatly impair other cognitive processes. Third, as is clear from the use of such words as *greatly* and *mild*, we consider specialization not as an all-or-none phenomenon but instead as a property that can vary across a continuum. If the area is involved in one and only one function then it is specialized in the extreme. But an area could also participate primarily in one function and to a lesser but not insignificant extent in others. In this case, it displays more specialization than an area that participates equally in a number of functions. Finally, there are two general types of evidence that might be marshaled for or against the existence of a specialized region: lesion and neuroimaging studies. We briefly discuss next how these types of evidence might be used to support claims of functional specialization.

Strong positive evidence for the existence of specialization is provided by the demonstration that lesions of a

cortical area produce isolated cognitive impairments—this follows rather simply from the definition provided above. Isolated lesion studies are weaker, however, at providing evidence *against* specialization. Consider the case in which a lesion impairs several cognitive processes, including the putatively specialized cognitive process of interest. One might argue that damage was sustained by adjacent cortical areas in addition to the hypothesized specialized area. This type of criticism is particularly relevant in human studies in which the lesions are typically large and not under experimental control. Also, as was considered above, the case in which no impairment follows the lesion of a putatively specialized neural substrate is not conclusive. It is possible that other regions operating in parallel continue to support the process. For example, the frontal eye fields, parietal cortex, and superior colliculus are thought to form such a redundant system for the neural control of saccadic eye movements (Lynch & McLaren, 1989). Within the field of object recognition, lesions in the macaque of areas that contain "face cells" do not lead to overt impairment on face recognition tasks (Heywood & Cowey, 1992), perhaps because other, redundant areas can support these processes.

The second type of evidence for specialization is provided by neuroimaging techniques, such as PET, fMRI, event-related potentials (ERP), and electrophysiological recording. Tests of this kind rank neural activity along some axis (e.g., rate of neuronal firing) and assume that a region that is specialized for a given cognitive process will respond maximally when that cognitive process is evoked (Barlow, 1995). This probabilistic induction is supported by demonstrations that the manipulation (i.e., lesion or stimulation) of populations of neurons with a particular maximal response impacts the very behavioral state to which the neurons are "tuned" (Funahashi, Bruce, & Goldman-Rakic, 1993; Merigan, Katz, & Maunsell, 1991; Merigan & Maunsell, 1990; Salzman, Britten, & Newsome, 1990). Thus, the finding that a cortical region fires maximally in response to face stimuli, for example, supports the hypothesis that the cortical region is specialized for face perception.

Unlike lesion studies, however, the results of neuroimaging experiments cannot be taken as conclusive evidence for the involvement in (or necessity of) a region for a given cognitive process. The primary cause of this state of affairs is the observational, correlative nature of neuroimaging (Sarter, Berntson, & Cacioppo, 1996). Although one wishes to make inferences regarding cognitive processes, these processes are not themselves directly subject to experimental manipulations. Instead, the investigator controls the presentation of stimuli and instructions, with the hope that these circumstances will provoke the subject to enter a certain cognitive state and no other. However, although cooperative, the subject may unwittingly engage in confounding cognitive processes in addition to that intended by the experimenter or, alternatively, may fail to differentially engage the pro-

cess (i.e., it may already be "on"). It is therefore not possible to know if observed changes in neural activity in a brain region are the result of the evocation of the cognitive process of interest or an unintended, confounding process. As a result, positive results from a neuroimaging study cannot be interpreted as demonstrating the necessity of a region for a cognitive process. Additionally, negative findings might result not from the noninvolvement of a region but from the insensitivity of the neuroimaging technique to the critical change in metabolic activity (e.g., timing of neuronal firing as opposed to rate; McCarthy, Puce, Gore, & Allison, 1997; Puce, Allison, Spencer, Spencer, & McCarthy, 1997; Riehle, Grun, Diesmann, & Aertsen, 1997). In the case of early visual processing, previous work allows us to be relatively confident of our assumptions and less leery of these potential inferential failures, but they deserve to be held in mind nonetheless.

Having defined our terms, let us now turn to the empirical literature on the neural bases of object recognition and attempt to reach some conclusions regarding the specializations of the human visual system for object recognition. We will begin with a brief survey of lesion studies in humans and monkeys and single-unit recording in monkeys, which have been the subject of more detailed reviews elsewhere (Desimone, 1991; Farah, 1990, 1991; Miyashita, 1993; Plaut & Farah, 1990; Tanaka, 1996). We will then turn to a recent body of work that has yet to be fully reviewed and synthesized—namely, studies of visual object recognition using PET and fMRI.

Studies of Visual Object Recognition in Human Agnosics and Nonhuman Primates

Research with nonhuman primates and with human agnosic patients has provided some evidence for the localization of visual object recognition. Both literatures (monkey and human) suggest that inferior temporal cortex is necessary for visual object recognition. Human agnosic patients, who are impaired at the recognition of visual stimuli despite grossly intact elementary visual capabilities, generally have lesions within inferior temporal cortex (Farah, 1990). Studies with monkeys have demonstrated impairments in the ability to learn discriminations based on the intrinsic shape of objects (as opposed to position or orientation) following lesions to roughly analogous areas (Weiskrantz & Saunders, 1984). Additionally, recordings from this brain region in monkeys have revealed cells whose responses are determined largely by the shape and identity of the stimulus, with relatively little dependence on properties of the appearance determined by the viewer's vantage point (e.g., size, position, orientation; Desimone, Schein, Moran, & Underleider, 1985), consistent with a specialization of this region for object representation.

Several features of the studies conducted to date have, however, prevented more precise statements concerning anatomical localization. With human patients, the typically large lesions and their haphazard placement make precise localization difficult. There are also several rea-

sons why the monkey may not provide a perfect model for the study of human visual recognition. One class of problems is simply historical: The ways in which visual object recognition has typically been studied in monkeys make the correspondence to human visual object recognition less than direct. In lesion studies, for example, monkeys are not tested on their recognition of premorbidly familiar objects. Instead, they are trained and then tested in visual discrimination tasks that have little surface similarity to normal human object recognition.

Even if the methodological difficulties of the monkey model were circumvented, there would remain the very real possibility of species differences in the localization of visual object recognition. One indication of how different the layout of monkey and human visual association cortex may be comes from the attempt to identify the human homologue of monkey area V4. Using converging evidence including cytoarchitecture, myeloarchitecture, and functional characteristics, some researchers suggest that the relatively high lateral location of V4 in the monkey brain corresponds to a location on the inferior surface of the human brain (Clarke & Miklossy, 1990). Others suggest that such phylogenetic leap-frogging of areas is unlikely, and they question whether the homology can be made at all on the basis of the currently available evidence (Heywood & Cowey, 1993). That V4 is the main source of projections into the inferotemporal areas important for object representation makes this example all the more discouraging for the prospect of localizing human visual object recognition by reference to the monkey.

The presence of language in the human brain, as well as the concomitant functional specialization of the hemispheres, introduces yet further obstacles to a straightforward mapping between the substrates of visual object recognition in monkey and man. Functionally, there is an intimate relation between the recognition of an object and the naming of it. Anatomically, the posterior language areas of the human left hemisphere border on visual association cortex, and visual object recognition disorders are sometimes seen in patients with posterior aphasias. Indeed, one authority on aphasia treats visual agnosia as part of the syndrome of transcortical sensory aphasia (Kertesz, 1979). And, of course, for one very important class of visual "objects" that humans recognize—namely, printed words—language clearly plays a central role.

Turning from questions of localization to questions of functional organization, many of the same difficulties are encountered. Is there a single general-purpose system for recognizing all types of visual objects, or are there specialized subsystems for different classes of object? For example, is face recognition accomplished by a separate system from common object recognition? Is the visual recognition of words during reading also functionally distinct from other forms of object recognition? As mentioned above, naturally occurring human lesions may affect multiple functionally distinct subsystems, resulting in an underestimation of the division of labor within human visual object recognition. Monkey lesion studies do not gener-

ally contrast different categories of stimuli, such as faces and nonface objects, and, of course, they cannot test hypotheses concerning visual word recognition. Similarly, single-unit recording studies in monkeys have only recently explicitly compared face and nonface object recognition (e.g., Mikami, Nakamura, & Kubota, 1994; Nakamura, Matsumoto, Mikami, & Kubota, 1994), and they are incapable of addressing the question of whether reading makes use of general-purpose visual pattern recognition or requires a distinct subsystem for visual word recognition. The limitations inherent in the traditional methods of neuropsychology and neuroscience provide the motivation for functional neuroimaging.

Neuroimaging Studies of Visual Object Recognition: New Strengths and Weaknesses

In recent decades, cognitive neuroscientists have begun to exploit a variety of techniques that allow measurement of regional cerebral blood flow in humans—most notably, PET and fMRI. These techniques allow the experimenter to measure regional brain activity, as indexed by blood flow, and correlate it with experimentally manipulated cognitive processes. These techniques have a number of advantages over the older methods, including two enormous ones: First, they allow us to study regional activity in human brains directly, rather than extrapolating from monkey brains. Second, their spatial resolution is far better than the typical size of naturally occurring lesions in humans.

The idea that the visual recognition system employs functional subcomponents that are specialized for the detection of specific categories of stimuli (i.e., faces, words, and general objects) is testable using neuroimaging methods. Ideally, one would examine the responsiveness of cortical areas to different classes of stimuli (e.g., faces, words, and general objects) in a single subject and test for the existence of regions with significantly greater responses to one class or another. The existence of any observed division might be further examined in additional subjects to determine whether these properties of cortical organization generalize to a larger population.

While there have been many neuroimaging studies of visual object processing (Haxby et al., 1991; Haxby et al., 1994; Haxby et al., 1996; Kanwisher, Chun, McDermott, & Ledden, 1996; Kanwisher, McDermott, & Chun, 1997; Kanwisher, Woods, Iacoboni, & Mazziotta, 1997; Kosslyn, Alpert, & Thompson, 1995; Kosslyn et al., 1994; Malach et al., 1995; Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995; Martin, Wiggs, Ungerleider, & Haxby, 1996; McCarthy et al., 1997; Menard, Kosslyn, Thompson, Alpert, & Rauch, 1996; Petersen, Fox, Posner, Mintun, & Raichle, 1988; Petersen, Fox, Snyder, & Raichle, 1990; Price et al., 1994; Puce, Allison, Asgari, Gore, & McCarthy, 1996; Puce, Allison, Gore, & McCarthy, 1995; Schacter et al., 1995; Sergent, Ohta, & MacDonald, 1992; Sergent, Zuck, Lévesque, & MacDonald, 1992; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996), this particular experiment has not been conducted. Instead,

several studies have attempted to localize the neural responses evoked by a single class of stimuli, often averaging across multiple subjects. In isolation, these studies are unable to provide definitive evidence of functional specialization. One possibility, for example, is that the regions identified in each study as responding more to a particular stimulus than to a control are in fact in the same location. It might be possible, however, to combine the findings from this body of literature to provide for stronger inference. If, for, example, the studies that have attempted to isolate word processing consistently activate neuroanatomical regions distinct from those identified by face processing tasks, then some evidence exists for functional specialization.

Thus, pooling together previous individual neuroimaging studies of visual object recognition might reveal specialization if, indeed, visual processing is functionally and anatomically componential. We present here the results of such a review, in which the reported foci of activation from 17 different neuroimaging studies were pooled in an attempt to identify consistent patterns of dissociation across studies. We note at the outset that a positive finding for a review of this kind is not only dependent on the true nature of visual processing but also dependent on a number of methodological aspects of the review and the studies themselves. These issues are discussed in some detail below.

METHOD

We identified PET and fMRI studies that have examined visual object recognition (including face and word studies). While many studies have used various "object" stimuli, only a subset of these were intended to identify the neural substrates of isolated object processing. Experimental designs that were not considered here include studies of implicit memory of object shape (Schacter et al., 1995) and studies of semantic knowledge of object category (Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Martin et al., 1995; Martin et al., 1996; Vandenberghe et al., 1996) that did not include contrasts with nonobject conditions. Instead, we focused our attention on studies that have contrasted visual object processing with "low-level" visual controls, such as fixation, scrambled lines, or textures. These contrasts were emphasized here so that any areas with common responses to different category types would also be identified, in addition to those found to be differentially sensitive. Only studies that reported local maxima of activation using the coordinate system of Talairach and Tournoux (1988; see below) were included in this review.

The studies were divided along two axes. First, we identified which studies used face, word, and general object stimuli. Comparison among these groups would constitute the main purpose of the review. Second, the studies were grouped by the nature of the task employed. Some of the studies made use of passive stimulus presentations and thus might be considered to be a more

"pure" examination of visual processing. Other studies used more complex tasks, raising the possibility of confounding and interacting cognitive processes introduced by failures in cognitive subtraction (Friston et al., 1996; Zarahn, Aguirre, & D'Esposito, 1997). We wished to examine whether any difference in localization would be found for passive versus "active" task conditions.

These studies have reported sites of significant signal change as coordinates of local maxima in standard Talairach space. With some assumptions regarding anatomical variability and methodological consistency (discussed in detail below), these coordinates can be used to combine and compare the results across studies. Local maxima, as opposed to the volume of activation reported for a given contrast, were used for several reasons. First, the coordinates are reported in a standardized fashion that facilitates comparison across studies, whereas this is not the case for volumetric data. Second, the statistical expectation for the location of a maximum should be relatively unaffected by choice of smoothing kernel or statistical power (Worsley, Marret, Neelin, & Evans, 1996), factors that vary from study to study. Volumetric results, on the other hand, are rather dependent on changes in exogenous smoothing or statistical power. Finally, the volumetric extent of significant clusters reported in most studies have been rather small (i.e., on the order of several cubic centimeters). As a result, the local maximum serves as a parsimonious representative of the cluster.

All reported local maxima posterior to the anterior commissure were tabulated and entered into a single data file. For two early studies (Petersen et al., 1988; Petersen et al., 1990), the anterior–posterior value was adjusted to reflect the authors' use of the midpoint between the anterior and posterior commissure as the origin, and left/right coordinates reflected across the midline to conform to the "left negative" convention. Visual inspection was used to determine the concordance of locations across study types. Quantitative assessments of the data (e.g., cluster analysis, multidimensional scaling) were not undertaken, although these might be applied in the future in an attempt to identify subtle relationships among the coordinates. The Appendix lists the coordinates from all studies.

RESULTS

A total of 82 coordinates were derived from twenty tasks from seventeen different studies. Table 1 presents the studies included in the review along with the task contrast evaluated. Figure 1 is a "glass brain" projection view of the local maxima showing their broad distribution, encompassing the posterior third of the cortex inferiorly and the posterior half superiorly. The failure to find more focal clustering is disappointing given the considerations that motivated this review—namely, the good spatial resolution of PET and fMRI relative to typical lesion sizes and their independence from the various factors that constrain the locations of lesions in the human brain.

Table 1
Studies Reviewed

Study	Task
	Words
Petersen et al., 1988	Passive viewing of words vs. fixation
Petersen et al., 1990	Passive viewing of words and pseudowords vs. passive viewing of false fonts
Howard et al., 1992	Read aloud visually presented words vs. view false fonts and say "crime"
Price et al., 1994, Experiment 1	Read aloud visually presented words vs. perform feature decision on false fonts
Price et al., 1994, Experiment 2	Passive viewing of words vs. passive viewing of false fonts
Menard et al., 1996	Passive viewing of words vs. fixation
Puce et al., 1996	Passive viewing of letter strings (nonwords) vs. passive viewing of textures
Polk e& Farah, 1998	Passive viewing of AltErNAtIng case words vs. passive viewing of consonant strings
	Objects
Sergent, Ohta, & MacDonald, 1992	Living/nonliving judgment regarding Snodgrass and Vanderwort (S&V) pictures vs. fixation
Sergent, Zuck, et al., 1992	Living/nonliving judgment regarding S&V pictures vs. judge gratings as vertical or horizontal
Kosslyn et al., 1994	Matching S&V pictures with their names vs. viewing random patterns of lines
Kosslyn et al., 1995	Picture verification performed on S&V style line drawings of objects and auditorally presented "entry level" words vs. scrambled lines and words
Malach et al., 1995	Passive viewing of objects vs. passive viewing of phase randomized pictures
Menard et al., 1996	Passive viewing of S&V pictures vs. fixation
Kanwisher et al., 1997	Passive viewing of S&V pictures (and novel S&V style objects) vs. passive viewing of scrambled lines
	Faces
Sergent, Zuck, et al., 1992	Gender categorization of faces vs. judge gratings as vertical or horizontal
Haxby et al., 1994	Matching faces across shifts of gaze vs. alternating buttonpresses to scrambled faces
Haxby et al., 1996	Encoding (viewing) faces vs. alternating buttonpresses to scrambled faces
Puce et al., 1996	Passive viewing of faces vs. passive viewing of textures
McCarthy et al., 1997	Passive viewing of faces among phase randomized objects vs. viewing of phase randomized objects

Perhaps the scatter apparent in Figure 1 masks a set of more focal localizations, corresponding to the three classes of stimuli: faces, general objects, and printed words. Figure 2 presents the maxima arranged by stimulus class. Again, the results are disappointing: The subsets of maxima associated with each stimulus class subtend almost as broad a swath of cortex as the combined set. A particularly surprising aspect of these results is the failure to reveal any pronounced hemispheric asymmetries by segregating the maxima according to stimulus class; faces have long been believed to require more right-hemisphere processing than left-hemisphere processing (DeRenzi, Perani, Calesimo, Silveri, & Fazio, 1994), and the opposite pattern would have been predicted for printed words (Coslett, 1997) and objects (Feinberg, Schindler, Ochoa, Kwan, & Farah, 1994).

Even when separated into active and passive studies (Figure 3), no clear segregation among stimulus types was apparent. Maxima for faces and objects, in particular, remained fully intermixed within the active and passive subsets. We noted a slight tendency for studies of word processing to report maxima superior to faces and objects in the passive tasks, but no statistical test was undertaken to rigorously assess this impression.

DISCUSSION

Across 17 neuroimaging studies, with designs capable of identifying brain regions associated with visual recognition, we found little convergence of results. Although the majority of activations were located within the posterior and inferior regions of cortex, there was considerable scatter within this broad zone and even some activations falling outside of it. This finding is in agreement with a smaller review undertaken by Bly and Kosslyn (1997), in which 10 studies were reviewed, and no reliable pattern of activation was discerned. Unlike that earlier review, which included 5 studies in common with this one and 5 with designs that did not isolate visual recognition per se, the present failure to find consensus cannot be attributed to a failure to distinguish between designs that permit the isolation of object recognition and those that do not (as discussed in the Method section).

A possible explanation for the scatter found here is that different regions of the posterior cortex are specialized for the recognition of different classes of visual stimuli. The observation of selective deficits of visual processing in brain-damaged patients has led some to propose that visual recognition is subserved by multiple, specialized neural subsystems (Farah, 1991). As was noted in the introduction, neuroimaging experiments conducted in neurologically intact subjects might have the ability to support such a model through the demonstration of separable neural regions with distinct maximal responses to different classes of visual stimuli. Several neuroimaging experiments have been reported in which the responses to single classes of visual stimuli (i.e., faces, general objects, or words) have been compared with low-level visual controls (e.g., scrambled pictures, textures, lines, fixation). We combined the results from these different studies here to test for the existence of separable neural regions devoted to face, word, and/or object processing, but we found scant evidence of segregation along these lines.

One possibility is that this review accurately reflects the nature of the functional substrates of visual processing. By this account, contrary to the neuropsychological
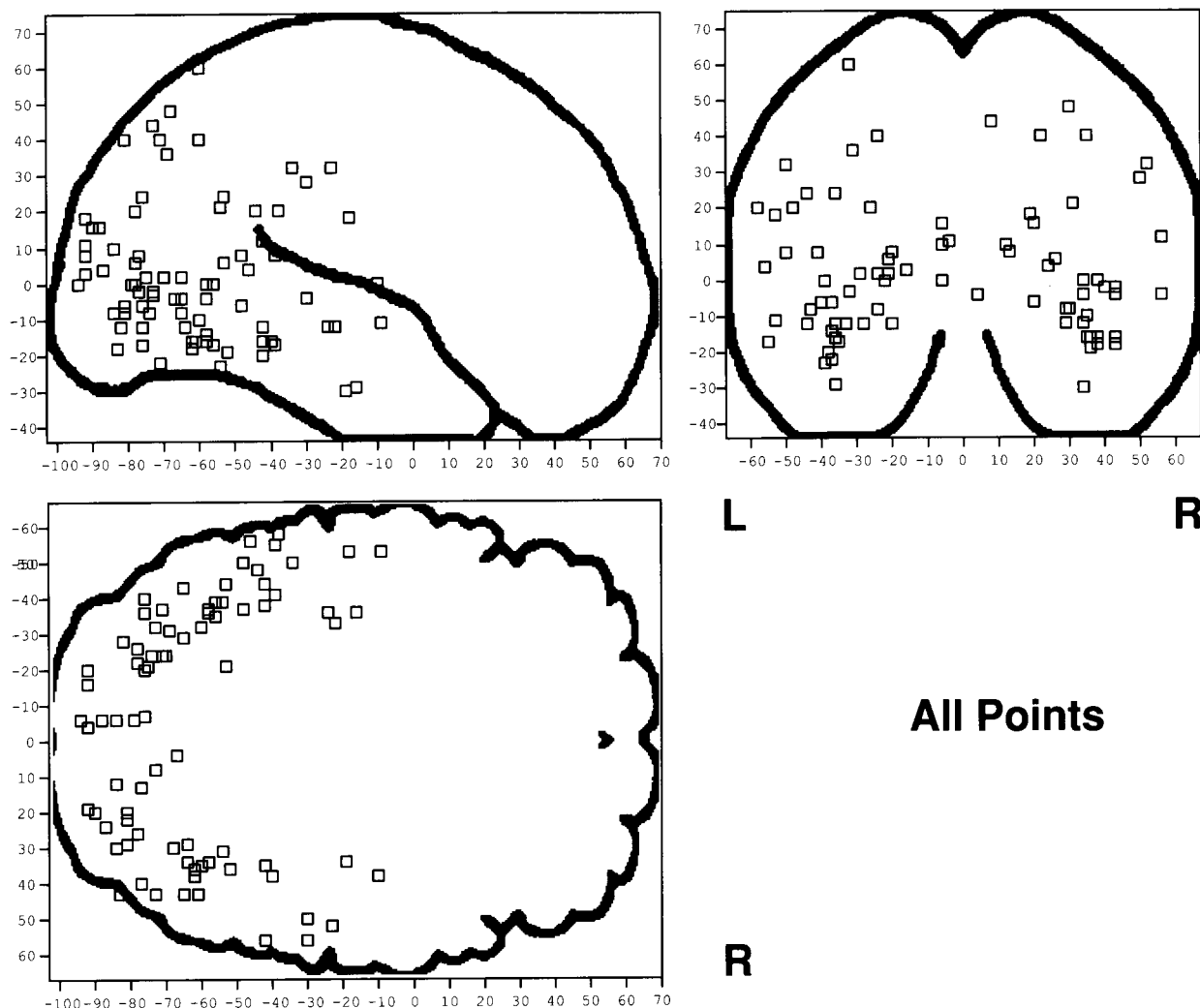
All Points

L   R

R

Figure 1. All 80 coordinate points from 17 different studies.

evidence, there are no functional subdivisions of visual recognition. Some might further interpret these findings as being consistent with "distributed code" models of visual recognition, in which populations of "prototypes" (Edelman & Duvdevani-Bar, 1997) share the representational burden. It should be noted, however, that the results presented here are consistent with distributed code models only in the weakest sense, in that the results do not strictly refute this class of model. This is because anatomical organization and degree of distribution (both in the spatial sense and representational sense) are parameters that might vary between different models. Consequently, a pattern of results opposite of that found here (i.e., clearly segregated regions that respond to either faces, objects, or words) could also be consistent with a distributed code model—albeit a model in which prototypes systematically vary in their tuning and spatial arrangement. Another way of stating these observations is that models that invoke functional specialization and dis-

tributed codes are not mutually exclusive. Finally, it should be noted that the results, if taken at face value, imply the same unpalatable features for any kind of model of visual processing, distributed or otherwise—namely, that there is tremendous variability from study to study in the location of cortical regions responsive to object recognition.

Another explanation for the failure of the results to accord with our hypotheses of localization and functional organization is that the review did not provide an appropriate test. In neuroimaging experiments such as those reviewed here, numerous layers of behavior, anatomy, physiology, and methodology intervene between the dependent data and the system that is the subject of theorizing. Even in the presence of true functional dissociations, a neuroimaging experiment (or review of such experiments) may fail to yield evidence of this aspect of neural organization due to limitations and/or properties of the intervening layers. Below, we briefly discuss some of these prop-
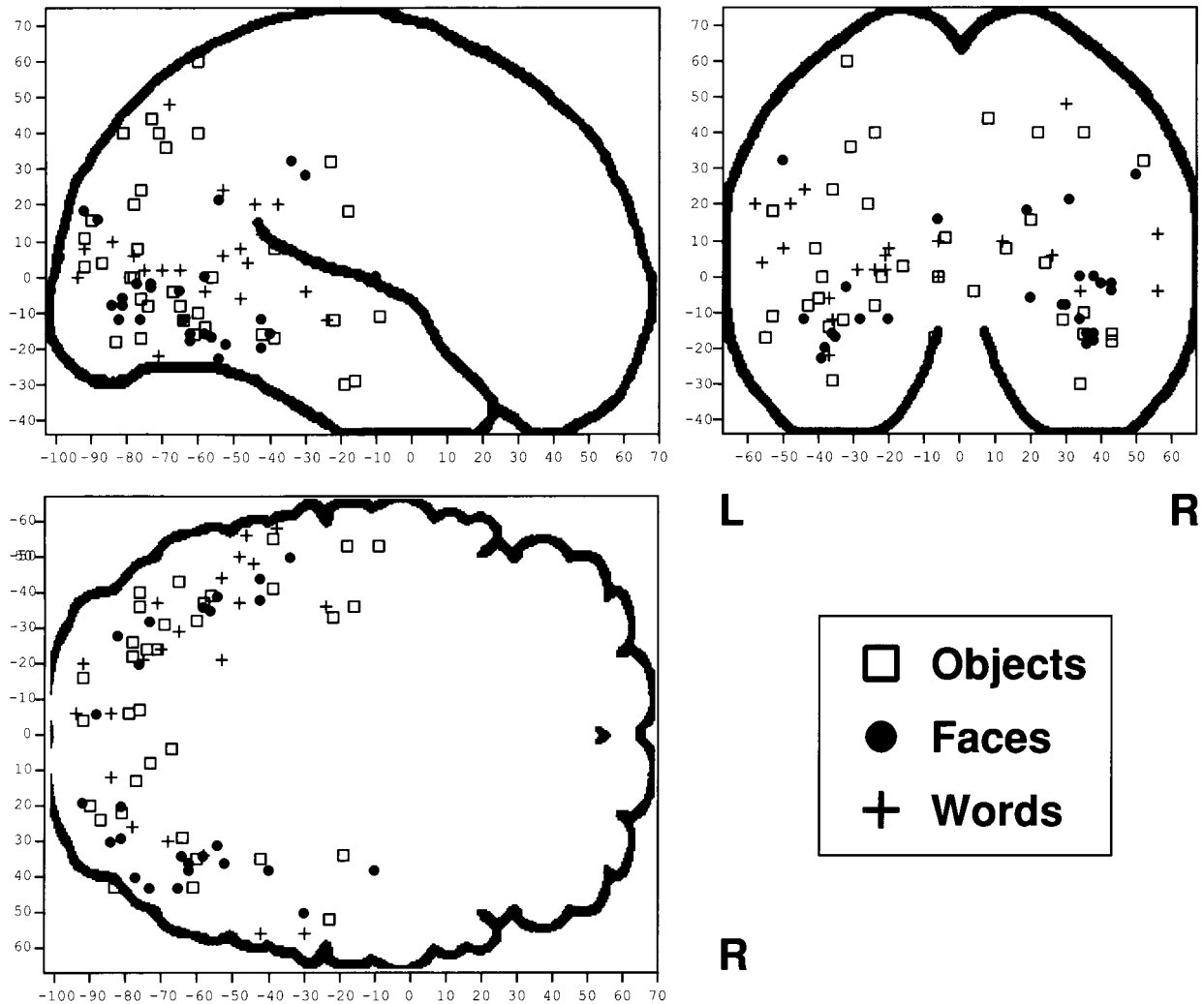
Figure 2. Coordinates classified by stimulus type.

erties. In general, these "methodological" limitations accrue when one attempts to make inferences across subjects and laboratories and act to increase the variability of observed locations of functional activity, impeding the detection of anatomical segregation. Potentially, properly conducted within-subject studies could ameliorate these effects and improve sensitivity for functional dissociations.

Differences Across Studies in Localizing Activations and Controlling False Positives

Because our review combined results obtained from different laboratories, variability across studies in the way in which functional imaging data are collected and analyzed might operate to disguise true anatomical dissociations. There are several possible sources of variability. First, different laboratories employ somewhat different methods of registering anatomical data into a standard space. The same "true" anatomical location might be reported at different locations due to differences in spa-

tial normalization techniques used in different laboratories (Woods, 1996). This will tend to disperse the locations of reported local maxima, reducing sensitivity for true spatial dissociations between tasks. Second, some laboratories may not control Type I error in a rigorous manner. False positive results might be the consequence of failures to employ a proper mixed effects model to allow for inference to be made regarding the population from which the subjects were drawn (Woods, 1996), failures to correct for the multiple voxels tested (Woods, 1996), or, in the case of fMRI, failures to account for the presence of temporally autocorrelated noise under the null hypothesis (Aguirre, Zarahn, & D'Esposito, 1998). The presence of local maxima that are the result of noise will tend to reduce the consistency of results across studies.

Differences Across Studies in Behavioral Design

Many neuroimaging studies, including all of those considered here, attempt to isolate the neural correlates of
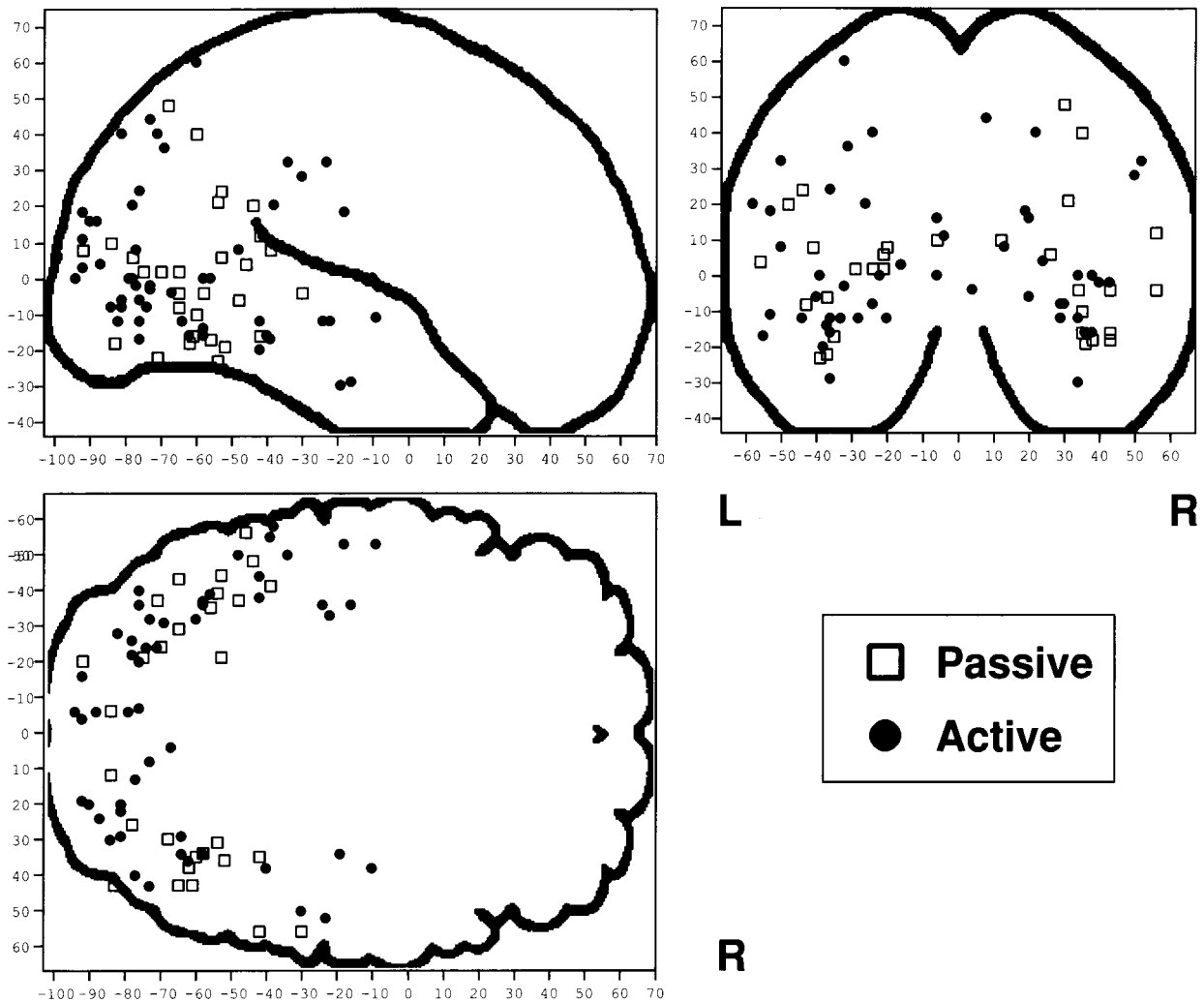
Figure 3. Coordinates classified by task design.

cognitive processes by cognitive subtraction (Posner, Petersen, Fox, & Raichle, 1988). In this approach, an experimental condition that engages a cognitive process of interest is compared with a control condition that is designed to account for some or all other uninteresting, confounding processes. Examination of Table 1 reveals that the studies reviewed here varied in the design of their behavioral subtractions. For example, different studies that attempted to isolate the process of word recognition used a variety of control conditions, including fixation (Menard et al., 1996), false fonts (Price et al., 1994), and textures (Puce et al., 1996). Differences in the pattern of neural activity evoked by these control conditions might be manifest as spread in the local maxima observed in the different studies. As a result, variability in behavioral design across studies can also contribute to reduced sensitivity for true anatomical dissociations in visual processing.

Differences Across Subjects in Anatomy, Functional Organization, and Behavior

The review conducted here considered data collected from many different subjects. Subject variability might obscure true neuroanatomical dissociations in several ways. First, there may be intersubject variability in anatomy that cannot be overcome by warping brains to a standard space. For example, the arrangement of the sulci in ventral occipitotemporal cortex is known to vary between subjects (Ono, Kubik, & Abernathey, 1990). Thus, while 2 subjects may have neural responses at the same "true" cytoarchitectonic location, the position of this site with respect to other landmarks in the brain may differ between subjects, leading to the spread of these locations when data are converted to a standard space (Woods, 1996). Second, even given rigid alignment of anatomy across subjects, there may be variability in the functional

organization of brain regions between subjects. For example, 2 subjects may truly have distinct face-selective neural regions, but these may be located in different sections of a cortical area as a consequence of differences in development or environment. Again, when normalized to a standard space, this variability in location will obscure functional dissociations. Third, there may be variability in behavior across subjects. Variability in functional activity might result from different subjects engaging in different cognitive processes in response to the same task. Finally, the degree of the hypothesized neural organization might vary across subjects, such that there might be true interactions between subject and, in this instance, modular organization of neural systems for visual processing. The existence of such interactions would tend to weaken the sensitivity of studies designed to test for consistent organization across subjects (and, indeed, should be properly accounted for with an appropriate mixed-effects model when inference regarding the entire population of subjects is desired).

Incidentally, variability in the degree of functional specialization across subjects is not necessarily incompatible with studies of patient populations that find evidence for highly specialized neural systems. This is because the subjects of neuropsychological studies are often selected based on the nature of their cognitive impairment. As a group, people in whom focal lesions produce isolated neuropsychological deficits would be expected to have had higher premorbid degrees of localized processing relative to the general population. One consequence of this observation is that models that posit functional specialization of neural systems based on consideration of patient deficits may not be found to hold for all members of the general population. Indeed, a study need only demonstrate that a proportion of neurologically intact subjects have the predicted functional specialization to be consistent with the lesion findings.

In summary, there are several features of the current analysis that reduce its ability to detect the hypothesized functional organization of extrastriate cortex for visual processing. These limitations are the result of pooling data across studies and subjects. It follows that studies performed within single subjects would avoid these difficulties and are therefore preferable. Of course, such experiments would also be capable of examining the consistency of results across subjects and thus provide for inference regarding functional organization within the population as a whole. Indeed, in contrast to the negative results obtained in this review, recent single-subject imaging experiments have demonstrated the existence of cortical areas with selective neural responses to classes of stimuli (Aguirre & D'Esposito, 1998; Kanwisher, McDermott, & Chun, 1997).

REFERENCES

AGUIRRE, G. K., & D'ESPOSITO, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: Evidence and implications. *Neuron*, **21**, 373-383.

AGUIRRE, G. K., ZARAHN, E., & D'ESPOSITO, M. (1998). A critique of the use of the Kolmogorov-Smirnov (KS) statistic for the analysis of BOLD fMRI data. *Magnetic Resonance in Medicine*, **39**, 500-505.

BARLOW, H. (1995). The neuron doctrine in perception. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 415-436). Cambridge, MA: MIT Press.

BLY, B. M., & KOSSLYN, S. M. (1997). Functional anatomy of object recognition in humans: Evidence from positron emission tomography and functional magnetic resonance imaging. *Current Opinion in Neurology*, **10**, 5-9.

CLARKE, S., & MIKLOSSY, J. (1990). Occipital cortex in man: Organization of callosal connections, related myelo- and cytoarchitecture, and putative boundaries of functional visual areas. *Journal of Computational Neurology*, **298**, 188-214.

COSLETT, H. B. (1997). Acquired dyslexia. In T. E. Feinberg & M. J. Farah (Eds.), *Behavioral neurology and neuropsychology* (pp. 197-208). New York: McGraw Hill.

DAMASIO, H., GRABOWSKI, T. J., TRANEL, D., HICHWA, R. D., & DAMASIO, A. R. (1996). A neural basis for lexical retrieval. *Nature*, **380**, 499-505.

DERENZI, E., PERANI, D., CALESIMO, G. A., SILVERI, M. C., & FAZIO, F. (1994). Prosopagnosia can be associated with damage confined to the right hemisphere—An MRI and PET study and review of the literature. *Neuropsychologia*, **32**, 893-902.

DESIMONE, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, **3**, 1-8.

DESIMONE, R., SCHEIN, S. J., MORAN, J., & UNDERLEIDER, L. G. (1985). Contour, color and shape analysis beyond the striate cortex. *Vision Research*, **25**, 441-452.

EDELMAN, S., & DUVDEVANI-BAR, S. (1997). A model of visual recognition and categorization. *Philosophical Transactions of the Royal Society of London: Series B*, **352**, 1191-1202.

FARAH, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision*. Cambridge, MA: MIT Press.

FARAH, M. J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology*, **8**, 1-19.

FEINBERG, T. E., SCHINDLER, R. J., OCHOA, E., KWAN, P. C., & FARAH, M. J. (1994). Associative visual agnosia and alexia without prosopagnosia. *Cortex*, **30**, 395-411.

FRISTON, K. J., PRICE, C. J., FLETCHER, P., MOORE, C., FRACKOWIAK, R. S. J., & DOLAN, R. J. (1996). The trouble with cognitive subtraction. *NeuroImage*, **4**, 97-104.

FUNAHASHI, S., BRUCE, C., & GOLDMAN-RAKIC, P. (1993). Dorsolateral prefrontal lesions and oculomotor delayed-response performance—Evidence for mnemonic scotomas. *Journal of Neuroscience*, **13**, 1479-1497.

HAXBY, J. V., GRADY, C. L., HORWITZ, B., UNGERLEIDER, L. G., MISHKIN, M., CARSON, R. E., HERSCOVITCH, P., SCHAPIRO, M. B., & RAPOPORT, S. I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Science*, **88**, 1621-1625.

HAXBY, J. V., HOROWITZ, B., UNGERLEIDER, L. G., MAISOG, J. M., PIETRINI, P., & GRADY, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *Journal of Neuroscience*, **14**, 6336-6353.

HAXBY, J. V., UNGERLEIDER, L. G., HOROWITZ, B., MAISOG, J. M., RAPOPORT, S. I., & GRADY, C. L. (1996). Face encoding and recognition in the human brain. *Proceedings of the National Academy of Science*, **93**, 922-927.

HEYWOOD, C. A., & COWEY, A. (1992). The role of the "face-cell" area in the discrimination and recognition of faces by monkeys. *Philosophical Transactions of the Royal Society of London: Series B*, **335**, 31-38.

HEYWOOD, C. [A.], & COWEY, A. (1993). Colour and face perception in man and monkey: The missing link. In B. Gulyas, D. Ottoson, & P. Roland (Eds.), *Functional organisation of the human visual cortex*. Oxford: Pergamon.

HOWARD, D., PATTERSON, K., WISE, R., BROWN, W. D., FRISTON, K., WEILLER, C., & FRACKOWIAK, R. (1992). The cortical location of the lexicons: Positron emission tomography evidence. *Brain*, **115**, 1769-1782.

JUST, M. A., & CARPENTER, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, **92**, 137-172.

KANWISHER, N., CHUN, M. M., MCDERMOTT, J., & LEDDEN, P. J. (1996). Functional imagining of human visual recognition. *Cognitive Brain Research*, **5**, 55-67.

KANWISHER, N., MCDERMOTT, J., & CHUN, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, **17**, 4302-4311.

KANWISHER, N., WOODS, R. P., IACOBONI, M., & MAZZIOTTA, J. C. (1997). A locus in human extrastriate cortex for visual shape analysis. *Journal of Cognitive Neuroscience*, **9**, 133-142.

KERTESZ, A. (1979). *Aphasia and associated disorders: Taxonomy, localization and recovery*. New York: Grune & Stratton.

KOSSLYN, S. M., ALPERT, N. M., & THOMPSON, W. L. (1995). Identifying objects at different levels of hierarchy: A positron emission tomography study. *Human Brain Mapping*, **3**, 107-132.

KOSSLYN, S. M., ALPERT, N. M., THOMPSON, W. L., CHABRIS, C. F., RAUCH, S. L., & ANDERSON, A. K. (1994). Identifying objects seen from different viewpoints: A PET investigation. *Brain*, **117**, 1055-1071.

LYNCH, J. C., & MCLAREN, J. W. (1989). Deficits of visual attention and saccadic eye movements after lesions of parietooccipital cortex in monkeys. *Journal of Neurophysiology*, **61**, 74-90.

MALACH, R., REPPAS, J. B., BENSON, R. R., KWONG, K. K., JIANG, H., KENNEDY, W. A., LEDDEN, P. J., BRADY, T. J., ROSEN, B. R., & TOOTELL, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences*, **92**, 8135-8139.

MARTIN, A., HAXBY, J. V., LALONDE, F. M., WIGGS, C. L., & UNGERLEIDER, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, **270**, 102-105.

MARTIN, A., WIGGS, C. L., UNGERLEIDER, L. G., & HAXBY, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, **379**, 649-652.

MCCARTHY, G., PUCE, A., GORE, J. C., & ALLISON, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, **9**, 605-610.

MENARD, M. T., KOSSLYN, S. M., THOMPSON, W. L., ALPERT, N. M., & RAUCH, S. L. (1996). Encoding words and pictures: A positron emission tomography study. *Neuropsychologia*, **34**, 185-194.

MERIGAN, W. H., KATZ, L. M., & MAUNSELL, J. H. (1991). The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. *Journal of Neuroscience*, **11**, 994-1001.

MERIGAN, W. H., & MAUNSELL, J. H. (1990). Macaque vision after magnocellular lateral geniculate lesions. *Visual Neuroscience*, **5**, 347-352.

MIKAMI, A., NAKAMURA, K., & KUBOTA, K. (1994). Neuronal responses to photographs in the superior temporal sulcus of the rhesus monkey. *Behavioural Brain Research*, **60**, 1-13.

MIYASHITA, Y. (1993). Inferior temporal cortex: Where visual perception meets memory. *Annual Review of Neuroscience*, **16**, 245-263.

NAKAMURA, K., MATSUMOTO, K., MIKAMI, A., & KUBOTA, K. (1994). Visual response properties of single neurons in the temporal pole of behaving monkeys. *Journal of Neurophysiology*, **71**, 1206-1221.

ONO, M., KUBIK, S., & ABERNATHEY, C. D. (1990). *Atlas of cerebral sulci*. New York: Thieme Medical Publishers.

PETERSEN, S. E., FOX, P. T., POSNER, M. I., MINTUN, M., & RAICHLE, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, **331**, 585-589.

PETERSEN, S. E., FOX, P. T., SNYDER, A. Z., & RAICHLE, M. E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science*, **249**, 1041-1044.

PLAUT, D. C., & FARAH, M. J. (1990). Visual object representation: Interpreting neurophysiological data within a computational framework. *Journal of Cognitive Neuroscience*, **2**, 320-343.

POLK, T. A., & FARAH, M. J. (1998). The neural development and organization of letter recognition: Evidence from functional neuroimaging, computational modeling, and behavioral studies. *Proceedings of the National Academy of Sciences*, **95**, 847-852.

POSNER, M. I., PETERSEN, S. E., FOX, P. T., & RAICHLE, M. E. (1988). Localization of cognitive operations in the human brain. *Science*, **24**, 1627-1631.

PRICE, C. J., WISE, R. J., WATSON, J. D., PATTERSON, K., HOWARD, D., & FRACKOWIAK, R. S. (1994). Brain activity during reading. The effects of exposure duration and task. *Brain*, **117**, 1255-1269.

PUCE, A., ALLISON, T., ASGARI, M., GORE, J., & MCCARTHY, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience*, **16**, 5205-5215.

PUCE, A., ALLISON, T., GORE, J. C., & MCCARTHY, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology*, **74**, 1192-1199.

PUCE, A., ALLISON, T., SPENCER, S. S., SPENCER, D. D., & MCCARTHY, G. (1997). Comparison of cortical activation evoked by faces measured by intracranial field potentials and functional MRI: Two case studies. *Human Brain Mapping*, **5**, 298-305.

RIEHLE, A., GRUN, S., DIESMANN, M., & AERTSEN, A. (1997). Spike synchronization and rate modulation differentially involved in motor cortical function. *Science*, **278**, 1950-1953.

SALZMAN, C. D., BRITTEN, K. H., & NEWSOME, W. T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, **346**, 174-177.

SARTER, M., BERNTSON, G., & CACIOPPO, J. (1996). Brain imaging and cognitive neuroscience: Toward strong inference in attributing function to structure. *American Psychologist*, **51**, 13-21.

SCHACTER, D. L., REIMAN, E., UECKER, A., POLSTER, M. R., YUN, L. S., & COOPER, L. A. (1995). Brain regions associated with retrieval of structurally coherent visual information. *Nature*, **376**, 587-590.

SERGENT, J., OHTA, S., & MACDONALD, B. (1992). Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain*, **115**, 15-36.

SERGENT, J., ZUCK, E., LÉVESQUE, M., & MACDONALD, B. (1992). Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. *Cerebral Cortex*, **2**, 68-80.

TALAIRACH, J., & TOURNOUX, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical Publishers.

TANAKA, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, **19**, 109-139.

VANDENBERGHE, R., PRICE, C., WISE, R., JOSEPHS, O., & FRACKOWIAK, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, **383**, 254-256.

WEISKRANTZ, L., & SAUNDERS, R. C. (1984). Impairments of visual object transformations in monkeys. *Brain*, **107**, 1033-1072.

WOODS, R. P. (1996). Modeling for intergroup comparisons of imaging data. *NeuroImage*, **4**, S84-S94.

WORSLEY, K., MARRET, S., NEELIN, P., & EVANS, A. (1996). Searching scale space for activation in PET images. *Human Brain Mapping*, **4**, 74-90.

ZARAHN, E. (1998). *The neural correlates of spatial mnemonic processing*. Unpublished doctoral dissertation, University of Pennsylvania.

ZARAHN, E., AGUIRRE, G. K., & D'ESPOSITO, M. (1997). A trial-based experimental design for fMRI. *NeuroImage*, **6**, 122-138.

APPENDIX
Study Coordinates

The Talairach coordinates used in this review are listed below. Locations are in millimeters relative to the anterior commissure, with negative values in *x* located on the left side. The experimental contrasts for each study are provided in Table 1.

Study	*x*	*y*	*z*	Study	*x*	*y*	*z*
Words				Kosslyn et al., 1995	−22	−78	0
Petersen et al., 1988*	−6	−84	10		−36	−76	24
	12	−84	10		−31	−69	36
	−24	−70	2		−39	−56	0
	26	−78	6		20	−90	16
	34	−58	−4		13	−77	8
Petersen et al., 1990*	−29	−65	2		4	−67	−4
	−21	−75	2	Malach et al., 1995	43	−83	−18
	−21	−53	6	Menard et al., 1996	−43	−65	−8
Howard et al., 1992	−50	−48	8		−41	−39	8
Price et al., 1994,	−6	−94	0		35	−60	40
Experiment 1†	−58	−38	20	Kanwisher et al., 1997	43	−61	−16
	−36	−24	−12		35	−60	−10
Price et al., 1994,	−20	−92	8		35	−42	−16
Experiment 2†	−56	−46	4	Faces			
	−48	−44	20	Sergent, Zuck, et al., 1992	19	−92	18
	56	−30	−4		40	−77	−2
	56	−42	12		43	−73	−2
Menard et al., 1996	−44	−53	24		29	−81	−8
	30	−68	48		−32	−73	−3
Puce et al., 1996	−37	−71	−22		20	−81	−6
Polk & Farah, 1998	−37	−48	−6	Haxby et al., 1994	30	−84	−8
Objects					36	−62	−16
Sergent, Zuck, et al., 1992	−55	−39	−17		34	−58	0
	−37	−58	−14		38	−40	−16
	−53	−9	−11		−28	−82	−12
	−40	−76	−6		−36	−58	−16
	−32	−60	60		−38	−42	−20
	52	−23	32	Haxby et al., 1996	34	−64	−12
	−53	−18	18		−20	−76	−12
Sergent, Ohta,	−4	−92	11		−44	−42	−12
& MacDonald, 1992	−16	−92	3		−50	−34	32
	29	−64	−12		38	−10	0
	−24	−74	−8		50	−30	28
	−7	−76	−17		−6	−88	16
	34	−19	−30	Puce et al., 1996	31	−54	21
	−36	−16	−29		38	−62	−18
Kosslyn et al., 1994	−6	−79	0		43	−65	−4
	−26	−78	20		−39	−54	−23
	−24	−71	40	McCarthy et al., 1997	36	−52	−19
	−33	−22	−12		−35	−56	−17
	24	−87	4				
	22	−81	40				
	8	−73	44				

*These coordinates have been altered from their reported values to conform to the conventions used here.   †These coordinates are for the summed effects of both durations (150 + 981 msec).