Is face recognition 'special'? Evidence from neuropsychology

Martha J. Farah *

University of Pennsylvania, Philadelphia, PA, USA

Received 17 December 1994; accepted 14 March 1995

Abstract

Is face recognition 'special,' in the sense of relying on functionally and anatomically distinct mechanisms from those required for other kinds of pattern recognition? A number of different neuropsychological dissociations involving recognition and learning of faces and nonface objects are reviewed. In addition, studies of the nature of shape representation in normal face and object recognition are reviewed. The evidence from brain-damaged and normal subjects suggests that face recognition is, indeed, 'special,' and provides some clues to the functional differences between face and object recognition.

Keywords: Face recognition; Neuropsychological dissociation; Object recognition

1. Introduction

The recognition of faces and many other types of three-dimensional objects seem to pose essentially the same problem. Depending upon the angle from which we view them, or the positions of their moveable parts, they may cast radically different images on our retinas, as demonstrated in Fig. 1. The primary task our visual systems confront in recognizing these stimuli is to create a representation that discriminates among them yet is invariant over at least a range of viewing conditions (e.g., [12]). When thought about in this way, the recognition of common objects and of faces seem to pose largely the same problem. Indeed, Marr [12] discussed them together and suggested an explanation of the effects of orientation on terms of his theory of object recognition (p. 310). The assumption that face and object recognition require essentially the same mechanisms can be found in many different disciplines in which vision is studied. Within neuropsychology, for example, Damasio and colleagues have asserted that face recognition is an instance of the more general problem of recognizing a particular exemplar of an object category (e.g., [2]), be it one person's face among all faces, or one particular chair among all chairs. Within neurophysiology, researchers may address general issues in object recognition, such as the distinction between object and location representation or the use of viewer-centered versus object-centered representations, yet confine their set of objects to faces (e.g., [11,15]).

In contrast, other lines of research have focused on the ways in which face recognition may be different from common object recognition. In addition to the neuropsychological support to be discussed in this chapter, evidence from normal subjects suggests that face recognition is, at least in some ways, different from other types of object recognition. For example, infants are born with a preference to gaze at faces rather than other objects. At just 30 min of age, they will track a moving face farther than other moving patterns of comparable contrast, complexity, and so on (see [14] for a review of this and other studies of infant face perception). The effects of orientation on face recognition, already mentioned, provides another indication that face recognition is special. Whereas most objects are a bit harder to recognize upside down that rightside up, inversion makes faces dramatically harder for normal adult subjects to recognize (see [18] for a review).

These findings from normal subjects indicate differences between face recognition and the recognition of other objects: Face recognition has earlier developmental precursors and is more orientation-sensitive than the recognition of other types of object. However, they do not necessarily imply that face recognition is accomplished using a different system from object recognition.
Fig. 1. Examples of how images of faces and other objects vary according to viewpoint.

Faces could be the first type of shape that a general-purpose system represents. Similarly, faces could require a special orientation-sensitive type of shape representation that is derived within a physically unitary and functionally indivisible system.

The hypothesis of interest in the present article is whether or not face and nonface object recognition are accomplished by different systems. I will therefore begin by stating explicitly what I mean by 'different systems.'

For systems to be considered different, they should be functionally independent, such that either can operate without the other. They should also be physically distinct (not necessarily redundant with the first criterion, as two programs running on the same computer could be functionally independent). Finally, they should not be mere physical and functional duplicates of one another, but should process information in different ways. In the present article, I will review neuropsychological evidence
relevant to the hypothesis that face recognition depends upon a different system from the recognition of common objects.

2. Visual agnosia and prosopagnosia

The most relevant neuropsychological impairment for present purposes is visual agnosia. The term agnosia refers to an impairment of object recognition that is not attributable to a loss of general intellectual ability or to an impairment in elementary perceptual processes, such as brightness, acuity, depth, or color vision (see [4] for a detailed overview). Thus by definition, agnoses retain full knowledge of the nonvisual aspects of the object, enabling them to recognize it by touching it, hearing its characteristic sounds, or hearing a verbal definition of it. They can also perceive at least some of its visual properties.

In associative agnosia perception can be remarkably preserved, to the extent that the person may be able to draw a good copy of a drawing or object that he cannot recognize. Indeed, the term was coined, in the 19th century, because it seemed that perception could not be at fault in these cases, and that the problem must therefore lie in associating perception with knowledge of the objects. Our understanding of vision has now progressed to the point where we can identify different levels of visual representation, from those early and intermediate representations that make explicit the edges and surfaces in an image, to higher level representations that make explicit the more stable shape properties of the distal object and thus solve the 'problem' of visual object recognition mentioned earlier. Associative visual agnosia is probably best viewed as an impairment in the highest levels of visual representation, rather than as an inability to associate normal visual representations with other types of knowledge [4]. According to this view, the ability of associative agnosics to draw the object results from their use of lower level visual representations, whereas recognition requires higher level representations. The observation that associative visual agnosics tend to copy drawings slowly and slavishly, and mistake objects for visually similar objects, is suggestive of an impairment in visual perception (see [4]).

Associative visual agnosia does not always seem to affect the recognition of all types of stimuli equally. The selectivity observed in some cases of agnosia lends support to the view that there are specialized systems for recognizing particular types of stimuli. The best known example of this is prosopagnosia, the inability to recognize faces after brain damage. Prosopagnosics cannot recognize familiar people by their faces alone, and must rely on other cues for recognition such as a person's voice, or distinctive clothing or hairstyles. The disorder can be so severe that even close friends and family members will not be recognized. Although many prosopagnosics have some degree of difficulty recognizing objects other than faces, in some cases the deficit appears strikingly selective for faces.

2.1. Is prosopagnosia really selective for faces

The most straightforward interpretation of prosopagnosia is that the highest levels of visual representation are subdivided into specialized systems, and prosopagnosics have lost the system that is necessary for recognizing faces and not necessary, or less necessary, for recognizing other types of object. However, it is possible that faces and other types of object are recognized using a single recognition system, and that faces are simply the most difficult type of object to recognize. Prosopagnosia could then be explained as a mild form of agnosia, in which the impairment is detectable only on the most taxing form of recognition task. This account has the appeal of parsimony, in that it requires only one, single type of visual recognition system, and perhaps for this reason has gained considerable popularity.

To determine whether prosopagnosia is truly selective for faces, and hence whether the human brain has specialized mechanisms for recognizing faces, we must therefore assess the prosopagnostic performance on faces and nonface objects relative to the difficulty of these stimuli for normal subjects. One technical difficulty encountered in such a project is that normal subjects will invariably perform nearly perfectly on both face and nonface recognition tasks. The resultant ceiling effect will mask any differences in difficulty that might exist between tasks, making it pointless to test normal subjects in the kinds of recognition tasks that have traditionally been administered to patients. With this problem in mind, researchers have devised visual recognition tasks that test learning of face and nonface objects. By having subjects learn to recognize specific new exemplars of faces and other types of object, it is possible to titrate normal subjects' level of recognition performance so that it falls between ceiling and floor.

The first researchers to address this issue directly were McNeil and Warrington [13]. They studied case W.J., a middle-aged professional man who became prosopagnostic following a series of strokes. After becoming prosopagnostic, W.J. made a career change and went into sheep farming. He eventually came to recognize many of his sheep, although he remained unable to recognize most humans. The authors noted the potential implications of such a dissociation for the question of whether human face recognition is 'special,' and designed an ingenious experiment exploiting W.J.'s new-found career. They assembled three groups of photographs, human faces, sheep faces of the same breed kept by W.J., and sheep faces of a different breed, and attempted to teach subjects names for each face. Normal subjects performed at intermediate levels between ceiling and floor in all
conditions. They performed better with the human faces than with sheep faces, even those who, like W.J., worked with sheep. In contrast, W.J. performed poorly with the human faces, and performed normally with the sheep faces. These data suggest that W.J.’s recognition impairment does not affect the recognition of all groups of visually similar patterns, but is selective for human faces.

Karen Levinson, Karen Klein and I took a similar approach, but used common objects rather than faces of another species to compare with human face recognition [6]. Our subject was L.H., a well-educated professional man who had been prosopagnosic since an automobile accident in college. L.H. is profoundly prosopagnosic, unable to recognize reliably his wife, children, or even himself in a group photograph. Yet he is highly intelligent, and seems to have little or no difficulty recognizing other types of visual patterns such as printed words or objects. Although he has a degree of recognition impairment with drawings of objects, this is still less severe than his impairment with faces.

We employed a recognition memory paradigm, in which L.H. and control subjects first studied a set of photographs of faces and nonface objects, such as forks, chairs, and eyeglasses. The photographs in Fig. 1 were drawn from this experiment. Subjects were then given a larger set of photographs, and asked to make ‘old’/‘new’ judgements on them. This larger set was designed so that for each face and nonface object in the ‘old’ set there was a highly similar item in the ‘new’ set. For example, there were two upholstered swiveling desk chairs with arms, one in the ‘old’ set and one in the ‘new.’ Verbal descriptions would therefore be of minimal help in performing this memory task. Whereas normal subjects performed equally well with the faces and nonface objects, achieving on average 85% correct, L.H. showed a significant performance disparity, achieving only 62% correct for faces and 92% correct for objects.

In a second experiment, we compared recognition of exemplars of the category ‘face’ and an equivalent number of highly similar exemplars all drawn from a single nonface category, namely eyeglass frames. Examples of faces and eyeglass frames are shown in Fig. 2. The photographs were divided evenly into sets of ‘old’ items, which would appear in the study and test phases of the experiment, and sets of ‘new’ items, which would appear only at test. Similar-looking eye glass frames were separated into ‘old’ and ‘new’ sets to make the task more challenging (e.g., there were both ‘old’ and ‘new’ horn rims, and ‘old’ and ‘new’ aviator style frames). In this experiment normal subjects found face recognition considerably easier than eye glass frame recognition. Normal undergraduates achieved on average 87% faces correct and 67% eye glass frames correct. A second group of normal subjects, matched in age and education level with L.H., showed the same disparity, achieving on average 85% faces correct and 69% eye glass frames correct. As before, L.H. was disproportionately impaired at face recognition relative to nonface recognition, when his performance is considered relative to normal subjects. Specifically, he showed significantly less face superiority in this task than normal subjects, achieving 64% faces correct and 63% eye glass frames correct. Like the first experiment, this one also suggests that L.H.’s impairment in face recognition cannot be attributed to a general problem with object recognition. The present results also suggest that the problem does not lie with the recognition of specific exemplars from any visually homogeneous category, but is specific to faces.

A final experiment was undertaken to address the specificity of L.H.’s face recognition impairment. In essence, the design of the previous experiments amounts to a comparison between a prosopagnosic’s performance with faces and his performance with stimuli that are like faces (in their recognition difficulty, in their membership in a visually homogeneous category) without being faces, or rather without being processed by the hypothesized face-specific recognition mechanism. Stating the experimental design in this way suggests the ideal nonface comparison stimulus: upside-down faces. As mentioned earlier, inverting a face makes it much harder for normal subjects to recognize. On the basis of the face inversion effect, it is generally assumed that if a specialized face recognition mechanism exists, it is specialized for the processing of upright faces. Inverted faces therefore constitute ideal comparison stimuli: They are equivalent to upright faces in virtually all physical stimulus parameters, including complexity and inter-item similarity, but do not engage (or engage to a lesser extent) the hypothesized face-specific processing mechanisms.

Kevin Wilson, Maxwell Drain, James Tanaka and I [8] reasoned that if L.H.’s underlying impairment was not face-specific, then he would show a normal face inversion effect. In other words, he would perform normally with upright faces relative to his performance on inverted faces. In contrast, if he had suffered damage to neural tissue implementing a specialized face recognition system, he would show an absent or attenuated face inversion effect. That is, he would be disproportionately impaired with upright faces, relative to his performance on the comparison stimuli, inverted faces.

L.H. and normal subjects were tested in a sequential matching task, in which an unfamiliar face was presented, followed by a brief interstimulus interval, followed by a second face, to which the subject responded ‘same’ or ‘different’. The first and second faces of a trial were always in the same orientation, and upright and inverted trials were randomly intermixed. As expected, normal subjects performed better with the upright than with the inverted faces, replicating the usual face inversion effect: 94% and 82% correct, respectively.

In contrast, L.H. was significantly more accurate with inverted faces! He achieved 58% correct for upright and
72% correct for inverted faces. This outcome was not among the alternatives we had considered. We had assumed that if he had an impaired face processor, it would simply not be used in this task, leading to the prediction of an absent or attenuated face inversion effect. Instead, it appears he has an impaired face-specific processor, which is engaged by the upright but not the inverted faces, and used despite being disadvantageous. This result was confirmed in additional studies, which invariably showed either statistically significant or non-significant trends in the same direction.

The 'inverted inversion effect' found in this prosopagnosic subject has strong implications for the selectivity of prosopagnosia. Inverted faces are the perfect control stimulus for equating faces and nonface objects for such factors as complexity and interitem similarity. L.H.'s disproportionate impairment on upright relative to inverted faces therefore implies that an impairment of face-specific processing mechanisms underlies his prosopagnosia.

A second conclusion that follows from these results concerns the 'control structure' of visual processing. L.H.'s specialized face perception system was apparently contributing to his performance even though it was impaired and clearly maladaptive. This suggests that the specialized face system operates mandatorily, reminiscent of Fodor's [10] characterization of special-purpose perceptual 'modules' as engaged mandatorily by their inputs. The idea that the face system cannot be prevented from processing faces, even when damaged, may also explain why W.J. was able to learn to recognize individual sheep after his strokes but could not learn to recognize human faces.

The general conclusion of these studies of W.J. and L.H. is that prosopagnosia represents the selective loss of visual mechanisms necessary for face recognition, and not necessary (or less necessary) for other types of object recognition. In terms of the organization of the normal visual system, these studies suggest that faces are recognized differently than other objects.
3. Selective impairment of new face learning

Individuals such as W.J. and L.H. are impaired at both new face learning and recognition of previously familiar faces, as would be expected if the substrates of face representation were damaged. Lynette Tippett, Laurie Miller and I recently encountered someone with an even more selective impairment: Case C.T. is impaired at learning new faces, but is relatively preserved in his recognition of previously familiar faces and in his learning of other nonface visual objects [17]. This pattern of performance is consistent with a disconnection between intact face representations and an intact medial temporal memory system. As such, it provides additional evidence that the neural substrates of face representation are distinct from the representation of other objects, as they can be selectively disconnected from the substrates of new learning.

C.T.'s face perception was normal as measured by the Benton and Van Allen face matching task [1]. He also performed normally on the face inversion task used with L.H., in terms of overall level of performance and the presence of an inversion effect. His learning of verbal material and even visual material other than faces is also normal. However, when given the face and eyeglass learning task, he performed similarly to L.H., achieving 58% correct for faces and 63% correct for eyeglasses. Additional evidence of his inability to learn faces comes from his identification of famous faces. For people who were famous prior to C.T.'s head injury, he performed within the range of 8 age-matched control subjects on a forced choice ‘famous/not famous’ task, whereas for more recently famous individuals he performed at chance. One celebrity allows for an especially interesting comparison between premorbid and current face recognition: In the case of Michael Jackson, the singer’s extensive plastic surgery following C.T.’s injury provides us with a ‘within-celebrity’ comparison of face recognition. Despite the greater popularity and media exposure of Michael Jackson in recent years, C.T. recognized a picture taken in the 1970s and failed to recognize an up-to-date photograph.

4. Preserved face recognition in visual agnosia

Some associative visual agnosics appear to have more difficulty with object recognition than with face recognition, presenting us with the mirror image of the prosopagnosic’s impaired and spared abilities [5,9]. This pattern of impairment is interesting for two reasons. First, it offers further disconfirmation of the hypothesis that prosopagnosia is just a mild disorder of a general-purpose object recognition system, with faces simply being harder to recognize than other objects. If this were true, how could it be possible for a person to do better with faces than with other objects? Second, it distinguishes two possible relationships that might hold between the specialized face system and the nonface object system, shown in Fig. 3. The two systems could be arranged in series, such that all stimuli are first processed by one general system, and then faces receive further processing by the system that is specialized for faces. Alternatively, the two systems could be arranged in parallel, such that a stimulus is recognized either by one or the other, independently of one another.

Given the intuition that face recognition requires processing that is somehow more elaborate or demanding than object recognition, which presumably motivated the alternative accounts described in the last section, one might expect the first type of arrangement to hold. According to this view there is specialized face system, but it is not functionally independent from the object system; it requires input from the object system and performs further processing on that input. This arrangement contrasts with the first one, according to which earlier visual processes deliver their products to two parallel, independent systems, one of which is required to recognize faces and the other of which is required to recognize other objects.

If there are, indeed, associative visual agnosics with relatively spared face recognition, then the systems subsuming object and face recognition must be arranged in parallel. In on-going work, Marlene Behrmann and I have confirmed experimentally the clinical observation that faces can be disproportionately spared. We used the same faces and eyeglasses experiment that had earlier been administered to L.H.

The subject in this experiment was C.K., a young man who suffered a closed head injury resulting in severe visual agnosia, with seemingly spared face recognition. C.K.’s pattern of performance differs from normality in the opposite direction from L.H.’s: He obtains 98% correct for the faces and only 48% correct for the eyeglasses. This result, taken together with opposite
dissociation, implies that the systems specialized for face and object recognition are functionally independent. Put more precisely, there is one system that is more important for face recognition than for nonface object recognition, and another system (or systems) that is more important for nonface object recognition than for face recognition, and they are arranged in parallel as shown on the right side of Fig. 3.

5. Functional differences between face and object recognition

Given the evidence just reviewed of at least two specialized subsystems underlying visual recognition, let us now turn to the question of what these specialized systems might be specialized for, in terms of the way in which they represent shape information. In order to address this question, it would be helpful to first introduce the concept of a structural description. Many current theories of object recognition posit structural descriptions, which are representations of object shape in terms of parts, which are themselves explicitly represented as shapes in their own right. The more extensive the part decomposition, the more parts there will be in an object's representation, but the simpler those parts will be. The less the part decomposition, the fewer the parts there will be in an object's representation, but the more complex those parts will be.

The conjecture being put forth here is that face recognition is holistic, in the sense that it involves virtually no part decomposition, and hence requires the ability to represent complex parts. At first glance this hypothesis may seem already disconfirmed by the experiments of Davidoff and Donnelly [3]. They compared the availability of part representations during face and chair perception in a sequential matching task, and found no difference between faces and chairs. However, this result speaks against only the stronger hypothesis that subjects cannot voluntarily attend to facial parts as well as they can to chair parts, when stimuli are physically present. In collaboration with James Tanaka and others, I have carried out several tests of the hypothesis that normal or 'default' face recognition differs from the recognition of other kinds of objects. Taken together, these findings suggest that the face recognition system damaged in prosopagnosia is a system of relatively holistic representation. The purpose of the final experiment is to test this hypothesis directly.

Maxwell Drain, Jim Tanaka and I compared the relative advantage of whole faces over face parts for normal subjects and for the prosopagnostic L.H. Our initial plan was to administer the same task that Tanaka and I used with the normal subjects to L.H., but despite intensive effort, L.H. could not learn to recognize a set of faces. We therefore switched to a short-term memory paradigm, in which a face was presented for study, followed by a blank interval, followed by a second presentation of a face. The subject's task was to say whether the first and second faces were the same or different. There were two different conditions for the presentation of the first face: Either 'exploded' into four separate frames containing the head, eyes, nose and mouth (in their proper relative spatial position within each frame), or intact. The second face was always presented in the normal format, so that the two conditions can be called 'parts-to-whole' and 'whole-to-whole'. Normal subjects performed better in the 'whole-to-whole' than in the 'part-to-whole' condition, on average
93% and 74% correct, respectively, thus providing further evidence that their perception of a whole face is not equivalent to the perception of its parts. L.H. showed abnormally little difference between the two conditions, 74% and 73% correct, respectively. This finding is consistent with the hypothesis that he has lost the ability to see faces as wholes.

5.1. Conclusions

Is face recognition 'special'? The evidence reviewed in this article suggests that it is. The selective impairment of face recognition in two prosopagnosic subjects, McNeill and Warrington's case W.J., and the subject of our studies, L.H., implies that we are endowed with a specialized system for recognizing faces. This system is not necessary for (or is less important for) recognizing objects other than human faces, even when such objects form a large and visually homogeneous category such as sheep faces or eyeglass frames. The selective deficit for new face learning observed in C.T. also supports the existence of functional specialization for face representation. Furthermore, the system is anatomically distinct, in that it can be selectively damaged and selectively
disconnected from medial temporal areas by stroke or head injury. The opposite pattern of impairment observed in the agnostic subject C.K. suggests that the specialized face recognition system does not merely elaborate the processing of the object system, but rather processes stimuli in parallel with it, and is at least partially functionally independent of the other system.

In addition to being physically distinct and functionally independent, the two systems also appear to differ in the way they represent shape. Research with normal subjects suggests that faces are recognized as single complex wholes, without decomposition into separately represented parts. A study with L.H. showed that his short-term memory for faces does not benefit from the opportunity to perceive the face as a whole, consistent with the idea that he has lost holistic perception of faces.

Referring back to the issue raised at the outset, we can now offer a tentative answer: Face recognition and object recognition appear to depend on different systems, which are anatomically separate, functionally independent, and differ according to the degree of part decomposition used in representing shape.

Acknowledgment

This research was supported by ONR grant N00014-93-10621, NIMH grant R01 MH48274, NINDS grant R01 NS34030, Alzheimer’s Association/Hearst Corporation Research Grant PRG-93-153, an NSF STC grant to the Institute for Research in Cognitive Science at the University of Pennsylvania, and the University of Pennsylvania Research Foundation. I thank Vincent Walsh for helpful comments and advice.

References


