Dissociated Overt and Covert Recognition as an Emergent Property of a Lesioned Neural Network

Martha J. Farah, Randall C. O'Reilly, and Shaun P. Vecera

Covert recognition of faces in prosopagnosia, in which patients cannot overtly recognize faces but nevertheless manifest recognition when tested in certain indirect ways, has been interpreted as the functioning of an intact visual face recognition system deprived of access to other brain systems necessary for consciousness. The authors propose an alternative hypothesis: that the visual face recognition system is damaged but not obliterated in these patients and that damaged neural networks will manifest their residual knowledge in just the kinds of tasks used to measure covert recognition. To test this, a simple model of face recognition is lesioned in the parts of the model corresponding to visual processing. The model demonstrates covert recognition in 3 qualitatively different tasks. Implications for the nature of prosopagnosia, and for other types of dissociations between conscious and unconscious perception, are discussed.

In recent years, neuropsychology has seen what Weiskrantz (1990) has called an "epidemic" of dissociations involving the loss of conscious awareness in particular perceptual or cognitive domains. Many of these dissociations involve vision. In such cases, patients may deny being able to see or recognize visual stimuli, and indeed perform poorly on certain direct tests of visual perception, but may nevertheless manifest considerable knowledge of the stimulus on certain other, generally indirect, tests of perception. We defer discussion of most of these syndromes until the General Discussion section and begin with the primary focus of the present article, namely covert face recognition in prosopagnosia.

Prosopagnosia is an impairment of face recognition following brain damage, which can occur relatively independently of impairments in object recognition and which is not caused by impairments in lower level vision or memory. In at least some cases of prosopagnosia, there is a dramatic dissociation between the loss of face recognition ability as measured by standard tests of face recognition, as well as patients' own introspections, and the apparent preservation of face recognition when tested by certain indirect tests. Our goal is to elucidate the underlying causes of this dissociation and its implications for both the na-

ture of prosopagnosia and for the neural correlates of conscious and unconscious perception.

Evidence for Covert Recognition of Faces in Prosopagnosia

Demonstrations of covert recognition in prosopagnosia have made use of extremely varied methodologies so that it is unlikely that any simple methodological artifact underlies the phenomenon. The relevant research includes psychophysiological measures such as skin conductance responses (SCRs) and event-related potentials (ERPs) as well as behavioral measures such as reaction time (RT) and learning trials to criterion.

In the absence of theories relating psychophysiological indexes to mechanistic accounts of cognition or neural information processing, it is difficult to use the psychophysiological findings to constrain a mechanistic model of covert recognition. Therefore, we focus primarily on the behavioral data implicating covert recognition and provide just a brief review of some representative psychophysiological data here.

Psychophysiological Evidence

Bauer (1984) presented a prosopagnosic patient with a series of photographs of familiar faces. While viewing each face, the patient heard a list of names read aloud, one of which was the name of the person in the photograph. This test has been called the *Guilty Knowledge Test* because it is based on a technique used to assess suspects' familiarity with the details of crimes of which they deny knowledge. For normal subjects, the SCR is greatest to the name belonging to the pictured person, regardless of whether the subject admits to knowing that person. Bauer found that, although the prosopagnosic patient's SCRs to names were not as strongly correlated with the names as a normal subject's would be, they were nevertheless significantly correlated. In contrast, the patient performed at chance levels when asked to select the correct name for each face.

In a different use of the SCR measure, Tranel and Damasio

Martha J. Farah, Randall C. O'Reilly, and Shaun P. Vecera, Department of Psychology, Carnegie Mellon University.

Correspondence concerning this article should be addressed to Martha J. Farah, who is now at Department of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia, Pennsylvania, 19104-6196.

(1985, 1988) showed that prosopagnosic patients had larger SCRs to familiar faces than to unfamiliar faces, even though their overt ratings of familiarity versus unfamiliarity did not reliably discriminate between the two.

Renault, Signoret, Debruille, Breton, and Bolger (1989) recorded ERPs to familiar and unfamiliar faces that had been intermixed in different proportions within different blocks of trials. In general, the P300 component of the ERP is larger to stimuli from a relatively infrequent category. They found that a prosopagnosic patient showed larger P300s to whichever type of face, familiar or unfamiliar, was less frequent in a block of trials, even though the patient was poor at overtly discriminating familiar from unfamiliar faces.

Behavioral Evidence

The first evidence of covert recognition in prosopagnosia was gathered by Bruyer, Laterre, Seron, Feyereisen, Strypstein, Pierrard, and Rectem (1983) in the context of a paired-associate, face–name relearning task, and this task has become the most widely applied measure of covert recognition in prosopagnosia. Bruyer et al.'s patient was asked to learn to associate the facial photographs of famous people with the names of famous people. When the pairing of names and faces was correct, the patient required fewer learning trials than when it was incorrect, suggesting that the patient did possess at least some knowledge of the people's facial appearance. Unfortunately, this demonstration of covert recognition is not as meaningful as it could be, because Bruyer et al.'s subject was not fully prosopagnosic; he could manifest an appreciable degree of overt recognition on conventional tests of face recognition such as forced-choice, face-naming tests.

Recently, several more severe prosopagnosic patients have been tested in the face–name relearning task, and some have shown the same pattern of faster learning of correct than incorrect face–name associations, despite little or no success at the overt recognition of the same faces. For example, de Haan, Young, and Newcombe (1987b) documented consistently faster learning of face–name and face–occupation pairings in their prosopagnosic subject, even when the stimulus faces were selected from among those that the patient had been unable to identify in a preexperiment stimulus screening test.

Greve and Bauer (1990) used a different form of learning as evidence of covert recognition in prosopagnosia. They showed a prosopagnosic patient a set of unfamiliar faces and then showed him the same faces, each paired with another face, at which time he was asked the following two questions about each pair: Which of these faces have you seen before? Which of these faces do you like better? Normal subjects tend to prefer stimuli that they have seen previously, whether or not they explicitly remember having seen these stimuli, and this has been attributed to a "perceptual fluency" advantage for previously seen stimuli (Jacoby, 1984). Perceptual fluency refers to the facilitation in processing a stimulus that has already been perceived, which leads to a subjective sense of the stimulus seeming more salient, which may in turn be attributed by the subject to the attractiveness of the stimulus. Although the prosopagnosic patient was unable to discriminate previously seen from novel faces, he did show a normal preference for the previously seen faces.

Evidence of covert recognition has also come from RT tasks in which the familiarity or identity of faces are found to influence processing time. In a visual identity match task (see Posner, 1978) with simultaneously presented pairs of faces, de Haan et al. (1987b) found that a prosopagnosic patient was faster at matching pairs of previously familiar faces than unfamiliar faces, as is true of normal subjects. In contrast, he was unable to name any of the previously familiar faces. De Haan et al. then went on to show another similiarity between the performance of the patient in this task and that of normal subjects. If the task is administered to normal subjects with either the external features (e.g., hair and jaw line) or the internal features (e.g., eyes, nose, and mouth) blocked off and with instructions to match on the visible parts of the face, normal subjects show an effect of familiarity only for the matching of internal features. The same result was obtained with the prosopagnosic patient.

In another RT study, de Haan et al. (1987a, 1987b) found evidence that photographs of faces could evoke covert semantic knowledge of the depicted person, despite the inability of the prosopagnosic patient to report such information about the person when tested overtly. Their task was to categorize a printed name as belonging to an actor or a politician as quickly as possible. On some trials, an irrelevant (i.e., to be ignored) photograph of an actor's or polician's face was simultaneously presented. Normal subjects are slower to categorize the names when the faces come from a different-occupation category relative to a no-photograph baseline. Even though their prosopagnosic patient was severely impaired at categorizing the faces overtly as belonging to actors or politicians, he showed the same pattern of interference from different-category faces.

A related finding was reported by Young, Hellawell, and de Haan (1988) in a task involving the categorization of names as famous or nonfamous. Both normal subjects and a prosopagnosic patient showed faster RTs to the famous names when the name was preceded by a picture of a semantically related face (e.g., the name "Diana Spencer" preceded by a picture of Prince Charles) than by an unfamiliar or an unrelated face. Furthermore, the same experiment was carried out with printed names as the priming stimulus so that the size of the priming effect with faces and names could be compared. The prosopagnosic patient's priming effect from faces was not significantly different from the priming effect from names. However, the patient was able to name only 2 of the 20 face prime stimuli used.

In summary, a wide variety of methods has been used to document covert recognition of faces in prosopagnosia. Although in some cases fairly coarse-grained discriminations are taken as evidence of recognition (e.g., between familiar and unfamiliar faces, de Haan et al., 1987a; Tranel & Damasio, 1985), in other cases the patients successfully discriminate among unique individuals (e.g., Bauer, 1984; Bruyer et al., 1983; de Haan et al., 1987a, 1987b) or classify individuals in a way that presumably requires individual recognition (actors vs. politicians, de Haan et al., 1987b). Although we argue that not all viable interpretations of these phenomena have been considered and we urge consideration of a new interpretation, it would seem that the correct interpretation is very unlikely to be any kind of

methodological artifact. The investigators in this area have been vigorous in attempting to eliminate possible artifacts in each of the experimental paradigms they have used. Furthermore, the sheer diversity of such paradigms makes an artifactual explanation unlikely. Finally, the absence of covert recognition in some cases (e.g., Etcoff, Freeman, & Cave, 1991; Newcombe, Young, & de Haan, 1989; Sergent & Villemure, 1989) suggests that it is not a result of the experimental paradigms themselves.

Interpretations of Covert Recognition in Prosopagnosia and Their Implications

The foregoing results would appear to indicate that, at least in those cases of prosopagnosia that show covert recognition, the underlying impairment is not one of visual recognition per se but of conscious access to visual recognition. Indeed, all of the interpretations so far offered of covert recognition in prosopagnosia include this assumption.

For example, Tranel and Damasio (1988) said, of their patients' SCRs, that they are "not the result of some primitive form of perceptual process, but rather an index of the rich retro-co-activation produced when representations of stimuli successfully activate previously acquired, non-damaged, and obviously accessible facial records" (p. 248). Similarly, de Haan et al. (1987a) described their subject's prosopagnosia as involving a "breakdown (or disconnection) in the mechanisms that allow people to be aware of what has been recognized [rather] than a breakdown in recognition mechanisms per se" (p. 315). In a recent computer simulation of semantic priming effects, described in greater detail later, this group modeled covert recognition as a partial disconnection separating intact visual recognition units from the rest of the system, again preserving the assumption of intact visual recognition (Burton, Young, Bruce, Johnston, & Ellis, 1991). Bruyer (1991) offered a similar interpretation, in terms of personal (i.e., conscious agent) and subpersonal (i.e., comprising at least the visual recognition system) levels of description: "The conscious subject does not recognize or identify familiar faces, while her/his 'information processing system' does?" (p. 230).

Figure 1 shows a depiction of de Haan et al.'s (1992) conception of the relation between face recognition systems and other systems needed for conscious awareness and the functional lesion responsible for prosopagnosia with covert recognition. According to their model, the face-specific visual and mnemonic processing of a face (carried out within the face-processing module) proceeds normally in covert recognition, but the results of this process cannot access the conscious awareness system because of a lesion at location number 1.

A different type of explanation was put forth earlier by Bauer (1984), who suggested that there may be two neural systems capable of face recognition, only one of which is associated with conscious awareness. According to Bauer, the ventral cortical visual areas, which are damaged in prosopagnosic patients, are the location of normal conscious face recognition. The dorsal visual areas are hypothesized to be capable of face recognition as well, although they do not mediate conscious recognition but, instead, mediate affective responses to faces. Covert recognition is explained as the isolated functioning of the dorsal face system. This interpretation is similar to the others in that it hy-
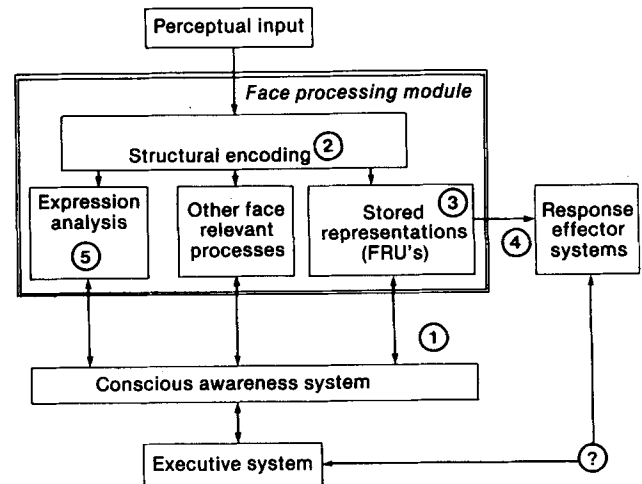

Figure 1. De Haan, Bauer, and Greve's (1992) framework for analyzing disorders of face recognition. (Covert recognition in prosopagnosia is thought to result from a functional disconnection at [1] between normally activated stored face representations and the conscious awareness system. FRU = face recognition units. From "Behavioral and Physiological Evidence for Covert Recognition in a Prosopagnosic Patient" by E. H. F. de Haan, R. M. Bauer, and K. W. Greve, 1992, Cortex, 28, p. 89. Copyright 1992 by Erminio Capitani. Reprinted by permission.)

pothesizes some form of intact visual recognition. It is distinctive in that the dissociation between recognition and conscious awareness is not a form of disconnection (functional or anatomical) between the visual recognition system and other brain systems that mediate conscious awareness brought about by brain damage but is the normal state of affairs for the dorsal face recognition system.

These interpretations of covert recognition have implications both for the nature of prosopagnosia and, more generally, for the neural bases of conscious awareness. With regard to prosopagnosia, current interpretations of covert recognition imply that there are at least two kinds of prosopagnosia, with different underlying causes: one in which visual recognition is intact but unavailable to consciousness (in the case of patients with covert recognition) and one in which visual recognition is impaired (in the case of patients without covert recognition).

At a more general level, these interpretations have implications for the broad issue of the neural bases of consciousness, in that they hypothesize distinct stages of processing, and corresponding distinct neural substrates, for face recognition on the one hand and awareness of face recognition on the other. The assignment of separate brain mechanisms to information processing and awareness of information processing has roots as far back as Descartes's writings on the mind–body problem (with the pineal gland subserving awareness in that case), and in the context of modern neuroscience this has been dubbed "Cartesian materialism" by Dennett and Kinsbourne (1992). Perhaps the most general and lucid expression of this idea, applied to a variety of neuropsychological syndromes including covert recognition by prosopagnosic patients, was put forth by Schacter, McAndrews, and Moscovitch (1988). They tentatively

proposed that "(a) conscious or explicit experiences of perceiving, knowing and remembering all depend in some way on the functioning of a common mechanism, (b) this mechanism normally accepts input from, and interacts with, a variety of processors or modules that handle specific types of information, and (c) in various cases of neuropsychological impairment, specific modules are disconnected from the conscious mechanism" (p. 269).

An Alternative Hypothesis: Residual Functioning of an Impaired Visual Recognition System

We argue that the available evidence on covert face recognition in prosopagnosics is consistent with an impairment in visual recognition per se. This interpretation has implications for our understanding of prosopagnosia in that it dispenses with the necessity of postulating different forms of prosopagnosia that are due to different underlying causes. Instead, cases with covert recognition are hypothesized to have more residual functioning of the visual face recognition system than cases without. It also has implications for our understanding of the neural bases of conscious awareness in that conscious awareness of recognition is not attributed to a distinct neural system from the one subserving recognition per se. Instead, the same neural system subserves both overt and covert recognition.

The primary challenge for such an account is to explain the dissociation between overt and covert recognition given that these two sets of phenomena are hypothesized to rely on the same neural substrates. We argue that the difference between them lies in the robustness to brain damage of performance of the two kinds of tasks, in other words, the degree of preserved neural information processing that is required in each case. Specifically, we argue that lower quality visual information processing is needed to support performance in tests of covert recognition (e.g., to show savings in relearning and the various RT facilitation and interference effects) relative to the quality of information processing needed to support normal overt recognition performance (e.g., naming a face or sorting faces into those of actors and politicians).

One very general way of stating this hypothesis is to say that the covert tests of recognition are more sensitive to the residual knowledge encoded in a damaged recognition system than are the overt tests. Thus, very impaired performance on overt tests might be associated with only moderately or slightly impaired performance on the covert tests. Stating the hypothesis in this way calls attention to two questions important for evaluating the hypothesis: First, what are the precise levels of patient performance on tests of overt and covert recognition, and are they consistent with the hypothesis of a single damaged system being tapped by tests of differing sensitivity? Normal-size covert recognition effects are unlikely to be due to the functioning of a damaged system (although it would not, strictly speaking, be impossible if the "ceiling" on covert performance was very low relative to the ceiling on overt performance). Better than chance performance by prosopagnosic patients on overt tests would also be consistent with residual functioning of the visual recognition system (although, by the same token, there is no logical reason why overt performance could not have its "floor" of chance performance above the floor of the covert tests). Second,

is there any independent reason to believe that the covert tests would be more sensitive measures of residual recognition ability in a damaged recognition system?

Empirical Evidence Relevant to Testing the Alternative Hypothesis

To answer the first question, it is impossible to compare directly the covert recognition performance of prosopagnosic patients and normal subjects on the basis of the evidence currently available, so we cannot know whether their covert recognition is normal or merely present to some degree. In some cases data from normal subjects have either not been reported, as in the P300 study, or would be impossible to obtain, as when familiar faces and names are retaught with either the correct or incorrect pairings. In other cases the problem of comparing effect sizes on different absolute measures arises. In both the SCR and RT paradigms, covert recognition is measured by differences between the dependent measures in two conditions (e.g., familiar and unfamiliar faces). Unfortunately, patients' SCRs are invariably weaker than those of normal subjects (Bauer, 1986), and their RTs are longer (de Haan et al., 1987b). It is difficult to know how to assess the relative sizes of differences when the base measures are different. For example, is an effect corresponding to a 200-ms difference between RTs on the order of 2 s bigger than, comparable to, or smaller than an effect corresponding to a 100-ms difference between RTs of less than a second? The true scaling of RT in any given task is an empirical issue; using proportions may be a better approximation to the scale than linearity, but one cannot a priori know the true scale (see Snodgrass, Corwin, & Feenan, 1990, for a discussion of these issues).

The study that comes closest to allowing a direct comparison of covert recognition in patients and normal subjects is the priming experiment of Young et al. (1988). Recall that they found equivalent effects of priming name classification for their prosopagnosic patient with either photographs or names of semantically related people. Of course, this fact alone does not imply that the face-mediated priming was normal, as face-mediated priming in this task might normally be larger than name-mediated priming. To address this problem, Young et al. cited their earlier experiment, reported in the same article, in which normal subjects were also found to show equivalent effects of face-mediated and name-mediated priming. Unfortunately, the earlier experiment differed in several ways from the latter, which could conceivably shift the relative sizes of the face-mediated and name-mediated priming effects: Normal subjects in the earlier experiment performed only 30 trials each, whereas the prosopagnosic patient performed 240 trials; items were never repeated in the earlier experiment, whereas they were in the later one; the type of prime was varied between subjects in the earlier experiment, whereas the prosopagnosic patient received both types; different faces and names were used in the two experiments; and the primes were presented for about half as long in the earlier experiment as in the later one. Ideally, to answer the question of whether this prosopagnosic patient shows normal priming from faces, a group of normal control subjects should be run through the same experiment as the patient.

Turning now to the question of whether the prosopagnosic

patients who show covert recognition also show some degree of overt recognition, consistent with a damaged but not obliterated visual recognition system, the evidence is similarly difficult to evaluate. For example, some patients' chance performance on overt tasks is consistent with the use of extreme response biases, which would mask any degree of remaining sensitivity. Among the three prosopagnosic patients studied by Tranel and Damasio (1988), two rated almost all faces as unfamiliar, and the one who used a larger portion of the rating scale narrowly missed the .05 significance level in discriminating familiar from unfamiliar faces.

Statistical naivete concerning the concept of chance performance has also led to confusion. In some cases, the term "chance performance" has been used synonymously with poor performance. For example, de Haan et al. (1987a) presented the results of an overt politician–nonpolitician face judgment task with their patient and described the score of 30/48 in a two-alternative, forced-choice task as being at chance. In fact, there is only a .06 probability of achieving such a high score by guessing alone. In other cases, performance is truly not statistically different from chance (e.g., in Young & de Haan, 1988, 12/30 in a three-alternative, forced-choice familiarity task), but the small number of trials makes this a relatively weak test for purposes of obtaining confidence in the null hypothesis. In addition, the ability of this patient and others to occasionally identify a face by name, a task whose "chance level" is difficult to estimate but is certainly close to 0% correct, also indicates that visual recognition has not been entirely obliterated. For example, this same patient was able to identify 2 out of 20 of the faces used in the semantic-priming study of Young et al. (1988).

One way in which investigators have attempted to control overt recognition performance and measure covert recognition in the absence of overt recognition is by testing patients only on faces that were not successfully identified in a screening test. For example, de Haan et al. (1987b) used only the faces that their prosopagnosic patient had failed to recognize in their face–name relearning task. This procedure is certainly a more conservative way of testing for a dissociation between overt and covert recognition than simply pooling the data from all faces. However, there could still be some measurement error in the overt task, which could in principle lead some overtly recognizable faces to remain in the covert test. Perhaps more likely, overt identification could be a less sensitive test of recognition than savings in relearning. That these considerations are not purely academic was demonstrated by Wallace and Farah (1992), who followed the same screening procedure of eliminating successfully identified faces with normal subjects on faces that had been learned 6 months before the experiment and nevertheless found savings in relearning the original face–name associations, relative to new pairings.

Computational Rationale for the Alternative Hypothesis

The empirical data reviewed so far fail to distinguish between the original hypothesis of intact face recognition deprived of access to consciousness and the alternative hypothesis that face recognition is impaired and that covert tasks are more sensitive than overt tasks to detecting residual functioning. Our reason for prefering the alternative hypothesis is based on a consider-

ation of the relative computational demands of the overt and covert tests. To explain how these differ, we first provide a very brief overview of computation in recurrent neural networks. More extensive background can be found in Rumelhart and McClelland's (1986) book on parallel distributed processing models of cognition.

In parallel distributed processing models, representations consist of a pattern of activation over a set of highly interconnected neuronlike units. The extent to which the activation of one unit causes an increase or decrease in the activation of a neighboring unit depends on the "weight" of the connection between them; positive weights cause units to excite each other, and negative weights cause units to inhibit each other. For the network to learn that a certain face representation goes with a certain name representation, the weights among units in the network are adjusted so that presentation of either the face pattern in the face units or the name pattern in the name units causes the corresponding other pattern to become activated. On presentation of the input pattern to the input units, all of the units connected with those input units will begin to change their activation in accordance with the activation value of the units to which they are connected and the weights on the connections. These units might in turn connect to others and influence their activation levels in the same way. In recurrent, or attractor, networks, the units downstream will also begin to influence the activation levels of the earlier units. Eventually, these shifting activation levels across the units of the network settle into a stable pattern, or attractor state. The attractor state into which a network settles is determined jointly by the input pattern (stimulus) and the weights of the network (stored knowledge).

Accordingly, much of the behavior of the network depends on the pattern of weights. For example, the weights determine not only which pattern becomes activated in association to an input pattern but they also determine how quickly this pattern becomes stable and how quickly a given unit or set of units reaches some predetermined threshold of activation. Not surprisingly, the current pattern of weights will also determine how many training cycles are needed to teach the network a new association. In ways that we elaborate on shortly, these aspects of network behavior seem closely related to the behavioral measures of covert recognition reviewed earlier: speed of perception (corresponding to settling time), speed of classifying actors and politicians (corresponding to how quickly actor or politician representations reach threshold), and, of course, paired-associate learning (a direct correspondence).

When a network is damaged by eliminating units, it will be less effective at associating the patterns that it knew previously. This can be understood in terms of the idea that knowledge is stored in the weights by viewing unit damage as the permanent zeroing of all weights going into and out of the eliminated units. As more units are eliminated, the ability of the network to correctly associate previously known patterns will steadily decline until it reaches chance levels.

The impetus for our project comes from the following key idea, first set out by Hinton and Sejnowski (1986; see also Hinton & Plaut, 1987): The set of the weights in a network that cannot correctly associate patterns because it has never been trained (or has been trained on a different set of patterns) is different in an important way from the set of weights in a net-

work that cannot correctly associate patterns because it has been trained on those patterns and then damaged. The first set of weights is random with respect to the associations in question, whereas the second is a subset of the necessary weights. Even if it is an inadequate subset for performing the association, it is not random; it has, "embedded" in it, some degree of knowledge of the associations.

Consideration of the kinds of tests used to measure covert recognition suggest that the covert measures might be sensitive to this embedded knowledge. The most obvious example is that a damaged network would be expected to relearn associations that it originally knew faster than novel associations because of the nonrandom starting weights. Less obvious, but nevertheless plausible for reasons to be elaborated on later, the network might settle faster when given previously learned inputs than novel inputs, even though the pattern into which it settles is not correct, because the residual weights come from a set designed to create a stable pattern from that input. Finally, to the extent that the weights continue to activate partial and subthreshold patterns over the nondamaged units in association with the input, these resultant patterns could prime (i.e., contribute activation toward) the activation of patterns by intact routes. We later discuss these mechanisms in greater detail in the context of the individual simulations. For present purposes, the general implication of these ideas is that as a neural network is increasingly damaged, there might be a window of damage in which overt associations between patterns (e.g., faces and names) would be extremely poor, whereas the kinds of performance measures tapped by the covert tasks might remain at high levels. Note that if this is true, it does more than just undermine the prevailing hypothesis of intact face recognition systems in those prosopagnosic patients who manifest covert recognition. It offers a specific, mechanistic hypothesis explaining the overt–covert dissociations in terms of general principles of computation in neural networks.

To test this hypothesis, we developed a very simple model of face recognition, and explored the effects of damage to visual input units on network performance of three different types of tasks, corresponding to the savings in relearning paradigm, the physical matching paradigm, and the priming paradigm. Before presenting the model and simulations themselves, we explain the concepts of activation space and weight space, which are helpful for understanding the behavior of the model.

Spatial Analogies for Understanding the Behavior of Attractor Networks

Spatial analogies are useful for visualizing certain aspects of network dynamics, including the way in which the network's patterns of activation change under the influence of an input and the way in which the ensemble of weights changes during learning. These analogies are also useful in understanding the behavior of the present network under damage.

The activation state of the network at any point in time can be represented as a point in a high-dimensional space called *activation space*. The dimensions of this space represent the level of activation of each unit in the network, assuming a fixed set of weights. In addition to the dimensions representing the activation levels of the units, there is one additional dimension,

representing the overall fit between the current activation pattern and the weights.

When units that are both active have a large positive weight between them, so that they reinforce each other's activation, this is an example of a good fit. If one unit is positively activated, another is negatively activated, and the weight connecting them is positive or if both units are positive and there is a negative (i.e., inhibitory) weight between them, the fit would be poor. This measure of fit is called *energy*, with low energy representing a better fit. The energy value associated with each pattern of activation defines a surface in activation space.

When an input pattern is presented to the network, the corresponding initial position in activation space is defined by the activation levels on the input units, along with resting level values for the dimensions representing the other units in the network. The weights in the rest of the network will not fit well with uniform resting level activation values over their portion of the network (assuming that they have been trained to associate a pattern with the input). Thus, the initial point in activation space will be in a region of high energy. As activation propagates through the network, the pattern of activation changes, and the point representing this pattern moves along the energy surface in activation space. The movement will generally be downward as the network lowers its energy, much as a ball rolls down a hill to lower its potential energy. To see why this would happen in terms of network dynamics, rather than by analogy with rolling balls, consider the examples given earlier of high- and low-energy activation states. For example, active units connected by negative weights (a poor fit, high-energy pattern) will tend to change their activations until one is active and the other is not (a good fit, low-energy pattern).

The energy minima toward which the network tends are the attractors mentioned earlier in this article. Attractors are useful in network computation not only for associating patterns and completing partial patterns but also for their ability to "clean up" a noisy input, by transforming a pattern similar to a known pattern into that known pattern (i.e., a pattern just uphill from an attractor will roll down into the attractor).

How quickly the network settles when presented with an input pattern depends on how quickly it can traverse the distance between its starting point in activation space and the attractor into which it "rolls." This in turn depends on the shape of the energy "landscape" because the network's activation pattern will travel more directly (and therefore quickly) down a steep, smooth incline than along more bumpy, winding terrain. The shape of the energy landscape is determined by the network's weights. In an untrained network, the landscape will be generally flat with random dips. When the network has learned a certain association, its weights will create an energy landscape in activation space in which the point corresponding to the input pattern and the attractor point corresponding to the complete associated pattern are connected by a smoothly and steeply sloping path that causes the one state to roll directly down into the other. Because some patterns will have the same value on some dimensions (i.e., they will have units activated in common), the network will need barriers to prevent confusion among trajectories for different patterns. Paths bounded by these barriers can be thought of as ravines.

The weights that underlie the attractor structure of activation

space can themselves be used to define a space, and this space is useful for visualizing the process of learning. In *weight space*, each of the weights in a network corresponds to one dimension of a space so that we can represent the sum total of the network's knowledge as a point in this high-dimensional space. If one additional dimension is now added to the space, representing the performance of the network at associating names and faces, then there will be a surface defined by each combination of weights and their associated performance. The energy of the point in activation space to which the network settles with a given set of weights is a measure of performance, with low energy (that is, good fit between the weights and the resultant activation pattern) being better performance. If, when we present the input, we also fix the activation values for the units for the associated pattern ("clamping"), then the desired weights will be those that create an energy minimum (with respect to activities) at this pattern. Learning consists of moving along this energy surface in weight space, changing weight values, until a sufficiently low point has been reached.

The Model

The present model is intended to illustrate some very general, qualitative aspects of the behavior of damaged neural networks in the kinds of tasks used with prosopagnosic patients. It is accordingly very simple. Figure 2 shows the architecture of the model. There are five pools of units. The face input units subserve the initial visual representation of faces, the semantics units subserve representation of the semantic knowledge of people that can be evoked by either the person's face or name, and the name units subserve the representation of names. In a model of this kind, hidden units are helpful to learn the associations among patterns of activity in each of these three layers. These are located between the face and semantic units, (called the *face hidden units*) and between the name and the semantic units (called the *name hidden units*). Thus, there are two pools of units that make up the visual face recognition system in our

model in that they represent visual information about faces: the face input units and the face hidden units.

The connectivity among the different pools of units was based on the assumption that to name a face, or to visualize a named person, one must access semantic knowledge of that person (Young, Hay, & Ellis, 1985). Thus, face and name units are not directly connected, but rather they send activation to one another through hidden and semantic units. The arrows in Figure 2 show the bidirectional connectivity between layers and the within-layer connectivity. Furthermore, each unit had a bias weight that learned the average activation level of that unit (a technique for improving the ability of the network to learn).

Units in this model have a threshold of zero. Thus, when the activation value of a unit is positive, it will activate those units to which it is connected by positive weights and inhibit those units to which it is connected by negative weights, and when its activation value is negative, it will have the opposite effects.

Faces and names are represented by random patterns of 5 active units out of the total of 16 in each pool. Semantic knowledge is represented by 6 active units out of the total of 18 in the semantic pool. The model makes no committment to any particular form of representation, beyond supposing that the representations are distributed; that is, each face, semantic representation, or name is represented by multiple units, and each unit represents multiple faces, semantic representations, or names. The information encoded by a given unit will be some "microfeature" (Hinton, McClelland, & Rumelhart, 1986) that may or may not correspond to an easily labeled feature (such as eye color in the case of faces). The only units for which we have assigned an interpretation are the "occupation units" within the semantic pool. One of them represents the semantic microfeature "actor" and the other represents the semantic microfeature "politician."

We created 40 distinct individuals, each consisting of a random name, face, and semantic pattern (over the 16 unlabeled semantics units). Ten individuals were actors (i.e., their semantic pattern had the actor unit active in addition to the other five active semantics units), 10 were politicians, and the remaining 20 were not assigned either of these two occupations. These 20 individuals were not tested in the simulations to be reported but were included in training to simulate the fact that subjects know many more people than are ever tested in a given experiment. Of the 10 actors and 10 politicians, 5 of each were not used in training so that we could compare the effects of familiarity on network performance in Simulation 2, resulting in a training set of 30 patterns.

The network was trained to be able to associate an individual's face, semantics, and name whenever one of these was presented, using the contrastive Hebbian learning (CHL) algorithm (Hinton & Sejnowski, 1986; Movellan, 1990; Peterson & Anderson, 1987). The CHL function can be used to train stochastic (e.g., Boltzman machine, Hinton & Sejnowski, 1986) and deterministic (e.g., mean field networks, Peterson & Anderson, 1987) recurrent networks. It stipulates that the weight change is proportional to the difference between the product of the activations of the two units on either side of the weight in the positive and negative phases, where the positive phase has both input and desired output patterns presented and the negative phase has just the input presented. In this way, the difference
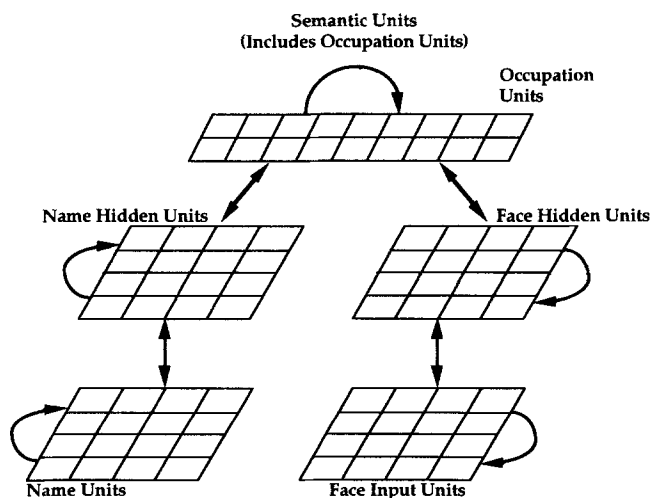

Semantic Units
(Includes Occupation Units)

Occupation Units

Name Hidden Units   Face Hidden Units

Name Units   Face Input Units

Figure 2. Model of face recognition.

between the desired activation state and the one that results from the input is minimized. The specific formulation used here follows Movellan (1990) in the use of the interactive activation and competition (IAC) activation function (McClelland & Rumelhart, 1981), with a step size of .01, maximum of 1, minimum of −1, rest of 0, and decay of 2.

For each training epoch we presented one of the three representations for each individual (face, semantics, or name) and trained the network to reproduce the other two. The learning rate was .01. The network was trained for 320 epochs on the complete set of 30 individuals and for an additional 5 epochs on the set of the 10 individuals to be tested later to ensure 100% accuracy for these individuals in the undamaged network.

Simulation 1: Savings in Relearning Face–Name Associations

The primary goal of this simulation was to examine the effects of different degrees of damage to the visual units (face input and face hidden units) on both overt identification of face patterns and on the difference in the number of cycles needed to relearn previously known name–face associations, relative to the number needed to learn to associate the same names and faces paired differently. Hinton and Sejnowski (1986) demonstrated savings in relearning after a variety of types of damage to a recurrent network, including unit ablation. Hinton and Plaut (1987) presented similar findings obtained with a feedforward network. If there is some degree of damage to the face units that can result in poor overt performance while preserving significant savings in relearning, then the savings in relearning observed in prosopagnosic patients need not imply that visual recognition per se has been spared.

Method. The network was lesioned in two different ways: by eliminating randomly chosen units from the face input pool and from the face hidden unit pool. These are the two pools of units in the model that correspond to visual face recognition. Seven different levels of damage were used, corresponding to removal of 2, 4, 8, 10, 12, and 14 units from the pools of 16 units, corresponding to 12.5%, 25%, 37.5%, 50%, 62.5%, 75%, and 87.5% damage.

The basic measure of overt recognition, used for comparison with covert performance in all of the simulations to be reported, was the percentage of correct name identifications of faces in a 10-alternative, forced-choice task among the 10 test patterns. A face was considered correctly identified if the resultant name pattern matched the correct name pattern more closely than any of the other 9 test patterns. Degree of match was quantified by the number of units having the same sign (positive or negative). This is a more lenient method of scoring overt recognition than requiring a perfect match, or even a match to within one bit.

In the first simulation, the names and faces for the 10 familiar actors and politicians were paired correctly. In the second simulation, they were paired incorrectly, although never across occupation categories, because this would confound the correct–incorrect distinction with the compatability of the occupation unit pattern. To expedite learning, each network was required to learn only five name–face pairs at a time. These were presented to the network after damage for retraining in separate simulations. To simulate the training procedure used with patients in which they are asked to name the face on each trial rather than select from a multiple-choice set of names, we used the pattern that resulted in the name units of the network following presentation of the face as

the simulation's response. This was scored as correct if it matched the target pattern to within 2 units.

To measure savings in relearning for correctly paired names and faces, the damaged network was retrained for 10 epochs, and its performance on overt identification was assessed. This procedure was repeated 10 times with different sets of random lesions to assess the reliability of the results.

Results and discussion. Table 1 and Figure 3 show the overt identification performance of the network in the 10-alternative, forced-choice task after different amounts of damage to the two pools of visual units. By 50% damage to either pool of units, the network is correct for only about 1 in 4 faces. With higher levels of damage, performance drops further. At 62.5% and 75% damage to face input units, only about 1 in 6 faces are correctly identified. At these same levels of damage to face hidden units, performance is not significantly different from 1 in 10, or chance performance.

Despite the network's poor performance in the overt tasks under damage, it manifests covert knowledge of the faces by relearning correct name–face pairings more quickly than incorrect ones. Table 2 shows the average percentage correct naming, to within a 2-unit matching criterion of the correct name, for each degree of damage to the face input and hidden units after 0 and 10 epochs of learning for correctly and incorrectly paired faces and names. Figure 4 shows the learning curves for the network after 50%, 62.5%, and 75% damage to the face input and face hidden units for the same pairings. Although not all levels of damage lead to equivalent performance for correct and incorrect pairings at the outset of training, the learning curve is steeper, that is, learning is faster for the correct pairings in all cases. The chance probability of six out of six cases showing the predicted form of interaction is $(.5)^6$, or .016. Furthermore, the faster relearning is found even with the four cases in which the

Table 1
Overt Identification in 10-Alternative, Forced-Choice Tasks

| Amount of damage (%) | Percentage correct | |
	M	SE
Hidden unit damage		
12.5	62	4.9
25.0	43	5.0
37.5	43	4.8
50.0	24	4.3
62.5	14	3.5
75.0	13	3.4
87.5	8	2.7
Input unit damage		
12.5	64	4.8
25.0	56	5.0
37.5	41	4.9
50.0	26	4.4
62.5	17	3.8
75.0	17	3.8
87.5	19	3.9

(A) Hidden Unit Lesions
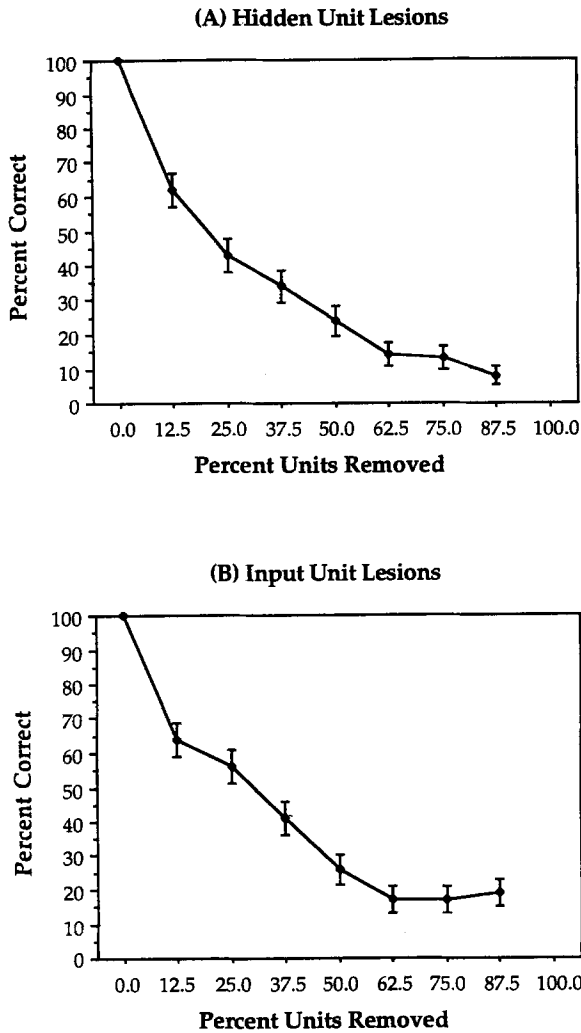

(B) Input Unit Lesions

Figure 3. Performance of model at forced-choice naming of faces after varying degrees of damage to (A) face hidden units and (B) face input units.

pretraining performance of the damaged network is comparable for correct and incorrect pairings: 62.5% and 75% damage to face input units and 50% and 75% damage to face hidden units. Tested individually, over the 10 different random lesions, these differences in learning rate have chance probabilities of .0004, .0763, .0671, and .0148 by t test, respectively.

Why does the damaged network relearn correct (old) name–face associations faster than incorrect (novel) ones? The answer lies in the ability of the network to incorporate the residual knowledge in the remaining weights into a new pattern of weights that enable it to associate names and faces. This can be verified by comparing the pattern of weights after damage but before relearning with the pattern of weights after relearning correct and incorrect face–name associations. Accordingly, we computed the total sum of the squares (TSS) distance between the weight matrices for the initial damaged network (which is a

subset of the original set of weights, specifically all weights except those going into and out of the eliminated units) and both the correct and incorrect face–name associations after relearning. This procedure was repeated for 10 different networks at each lesion size. As shown in Table 3, the pattern of weights in the damaged network before relearning is more similar to (lower TSS) the pattern of weights for the relearned correct face–name pairings than for the incorrect face–name pairings.

Simulation 2: Speed of Visual Perception

The goal of this simulation was to examine the effect of different degrees of damage to visual units on the speed of visual analysis of face patterns and specifically whether speed of analysis will depend on face familiarity at levels of damage where faces are not reliably identified. This question is of interest primarily because of de Haan et al.'s (1987b) demonstration that their prosopagnosic subject could perform physical same–different matching on faces more quickly when the faces were previously known to him. Presumably, the effect of familiarity on speed in this paradigm is not dependent on same–different matching per se but reflects a difference in the speed of deriving a visual representation that can be used to compare the appearance of the two faces. Therefore, we have not tried to implement a same–different matching paradigm here. The relevant issue is whether visual analysis of a face pattern proceeds more quickly when the face is familiar than when it is unfamiliar.

In the present model, the speed of visual perception is most directly measured by the number of cycles needed for the visual units of the network to settle into a stable pattern after presenta-

Table 2
Savings in Relearning Correct Relative to Incorrect Face–Name Pairings

| | Percentage correct | | | | | | | |
| | Correct pairings | | | | Incorrect pairings | | | |
| | 0 epochs | | 10 epochs | | 0 epochs | | 10 epochs | |
Amount of damage (%)	M	SE	M	SE	M	SE	M	SE
	Hidden unit lesion							
12.5	58.0	5.0	98.0	1.4	6.0	2.4	10.0	3.0
25.0	26.0	4.4	82.0	3.9	8.0	2.7	14.0	3.5
37.5	34.0	4.8	62.0	4.9	8.0	2.7	18.0	3.9
50.0	18.0	3.9	50.0	5.1	10.0	3.0	12.0	3.3
62.5	20.0	4.0	36.0	4.8	4.0	2.0	6.0	2.4
75.0	12.0	3.3	36.0	4.8	18.0	3.9	14.0	3.5
87.5	6.0	2.4	24.0	4.3	16.0	3.7	12.0	3.3
	Input unit lesion							
12.5	68.0	4.7	98.0	1.4	0.0	0.0	4.0	2.0
25.0	58.0	5.0	96.0	2.0	8.0	2.7	4.0	2.0
37.5	32.0	4.7	72.0	4.5	8.0	2.7	4.0	2.0
50.0	20.0	4.0	74.0	4.4	6.0	2.4	18.0	3.9
62.5	16.0	3.7	18.0	3.9	10.0	3.0	10.0	3.0
75.0	12.0	3.3	46.0	5.0	6.0	2.4	24.0	4.3
87.5	10.0	3.0	18.0	3.9	2.0	1.4	20.0	4.0
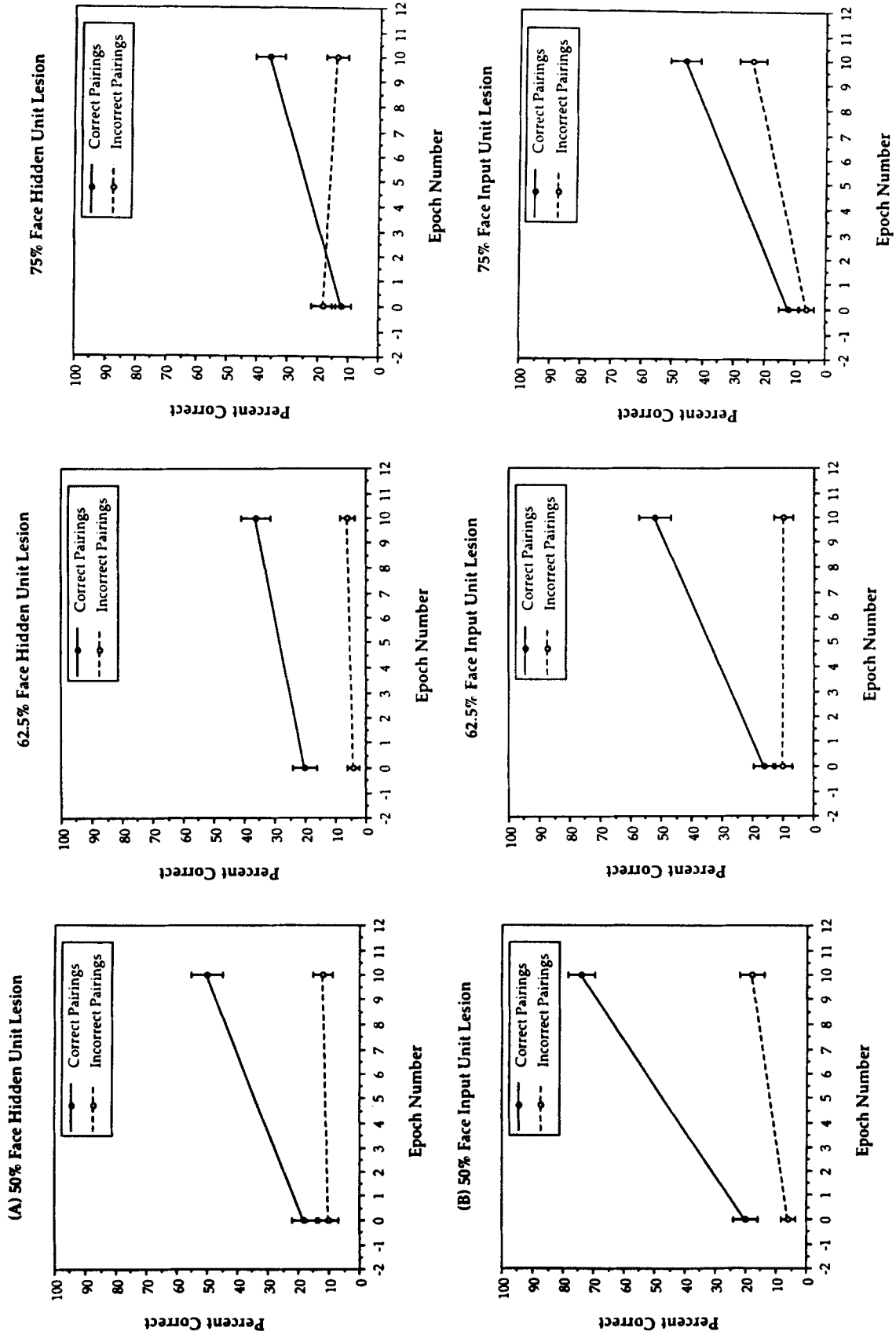
Figure 4. Savings in relearning face–name associations. (Performance of model after varying degrees of damage to [A] face hidden units and [B] face input units, initially and after 10 epochs of training, on correct and incorrect face–name associations.)

Table 3

Distances Between Weight Matrix Before and After 10 Epochs of Learning for Correct and Incorrect Face–Name Pairings

	Distance[a]			
	Correct pairing		Incorrect pairing	
Amount of damage (%)	M	SE	M	SE
Hidden unit lesion				
12.5	4.84	0.54	19.90	0.60
25.0	7.58	0.76	18.71	0.47
37.5	10.74	0.94	19.57	0.58
50.0	12.67	0.69	19.56	0.68
62.5	15.33	0.77	19.84	0.75
75.0	15.99	0.94	18.13	0.76
87.5	16.14	0.49	15.45	0.71
Input unit lesion				
12.5	2.88	0.65	19.20	0.40
25.0	6.81	0.77	19.38	0.62
37.5	11.19	0.87	19.51	0.78
50.0	14.90	1.17	20.68	0.66
62.5	15.87	1.07	19.87	0.66
75.0	16.71	0.87	21.00	0.74
87.5	20.67	1.13	19.10	0.58

[a] Distance is the total sum of the squares.

tion of a face pattern. Note that we need assume only a monotonic relationship between model settling time and human RT to interpret the results of the present simulation. Nevertheless, no transformations of the data are necessary to capture the qualitative pattern of human RTs.

Method. The model was lesioned as in the previous simulation. The face portion of the 10 actor and 10 politician patterns were then presented to the network. As explained earlier, in the description of the model, the network had been trained on half of these patterns, divided equally into 5 actors and 5 politicians. The number of cycles needed for the visual units (input and hidden) of the network to settle was recorded for each face pattern. The visual units were considered to have settled when the average change in activation of the units in a cycle was less than .001. The face input unit activations were allowed to settle by presenting the input pattern as a component of the net input to each unit, instead of simply clamping the activations (i.e., "soft" clamping). As for the previous simulation, 10 replications were performed with different random patterns of damage.

Results and discussion. The settling times for familiar and unfamiliar face patterns are shown in Table 4 and presented graphically in Figure 5. At levels of damage causing poor or chance overt performance (see Table 1), the settling time for familiar face patterns is nevertheless faster than for unfamiliar patterns. This pattern is maintained throughout all degrees of damage to the face hidden units and is present with as much as 50% damage to the face input units.

Why should the familiarity of the pattern affect how quickly it settles? In an intact network, a familiar input pattern will roll into an attractor representing the correct pattern of activation to which it should be associated, because the energy landscape has been tailored for this purpose. Given that much of the activation space has been shaped by learning, the trajectory of the network when presented with the input portion of an unfamil-

iar pattern will also tend toward attractors for the familiar patterns. However, because the activation space has not been specifically shaped to bring the unfamiliar patterns into attractors, their trajectories will typically be less direct and more circuitous. For this reason, familiar input patterns will settle faster than unfamiliar input patterns in an intact network.

When the network is lesioned, the loss of units reduces the dimensionality of the space, and the loss of weights distorts the shape of the new, lower dimensional energy landscape. Whereas the new surface will lack some of the topographic features that draw familiar input patterns into the correct attractors, it will retain others, so that at least part of the trajectory of at least some of the familiar patterns will tend to be maintained. In contrast, this will be no more true of unfamiliar patterns after damage than before.

The faster settling of familiar patterns is also relevant to Greve and Bauer's (1990) finding of greater attractiveness ratings for faces seen previously by their prosopagnosic subject, which they interpreted as greater perceptual fluency, or speed of perceptual analysis. With exposure to new patterns, the damaged network will alter its weights to begin to form attractors for those patterns, although it will arrive at the best set of weights more slowly than a network that has a larger number of weights (cf. the slower learning of the increasingly damaged networks in Simulation 1). This leads to faster settling times for faces seen postdamage than for completely novel faces, even before the network has learned to accurately recognize the patterns.

Indeed, examination of the settling times for the novel patterns of Simulation 1 (i.e., the novel combinations of faces and names) shows that, at early stages of learning associated with chance overt performance in the damaged network, settling time is nevertheless reduced relative to no learning. As shown in Table 5, at all levels of hidden unit damage and at 4

Table 4

Settling Time for Familiar and Unfamiliar Face Patterns

	Number of cycles			
	Familiar face		Unfamiliar face	
Amount of damage (%)	M	SE	M	SE
Hidden unit lesion				
12.5	154.3	6.9	278.1	14.1
25.0	176.5	11.5	256.3	12.8
37.5	170.6	10.2	267.4	13.8
50.0	162.5	10.8	223.5	14.3
62.5	145.0	8.5	191.6	10.1
75.0	124.2	6.9	162.5	8.3
87.5	119.9	12.0	138.3	7.4
Input unit lesion				
12.5	187.4	10.2	284.2	18.4
25.0	222.3	10.8	276.4	15.2
37.5	255.9	14.2	255.7	14.3
50.0	258.3	11.2	306.6	18.1
62.5	255.3	14.4	273.3	15.0
75.0	293.7	14.0	296.6	14.7
87.5	368.9	20.8	359.2	18.8

(A) Face Hidden Unit Lesions
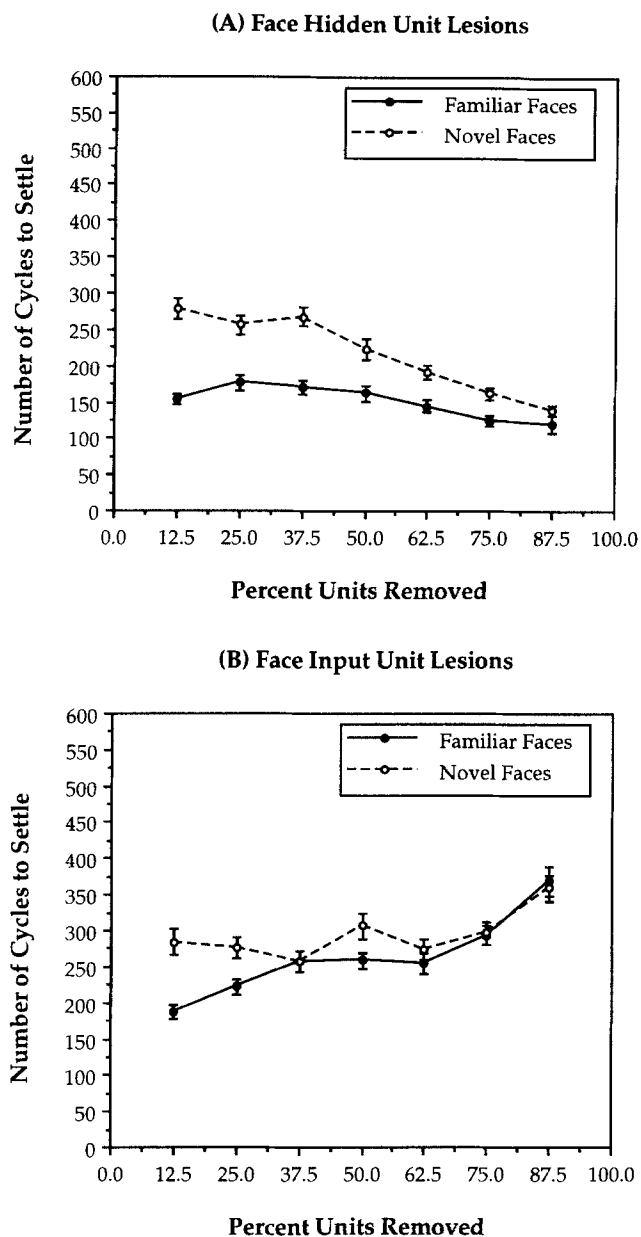

(B) Face Input Unit Lesions

Figure 5. Speed of perception of familiar and unfamiliar faces. (Settling time for [A] face hidden units and [B] face input units after varying degrees of damage, for familiar and unfamiliar faces.)

out of 7 levels of input unit damage, settling time is faster after just 5 epochs of training than before.

Simulation 3: Semantic Priming of Occupation Decisions

The goal of this simulation was to examine the effects of different amounts of damage to the visual units on facilitation and interference caused by a face prime when judging the occupation of a named person. As a related measure of overt perfor-

mance, the network was presented with the face input patterns alone to classify according to occupation.

Method. The model was lesioned as in the previous simulations. The name portions of the 5 familiar actor and 5 familiar politician patterns were presented to the network, paired with face patterns from the same group of individuals. Each of the 10 names was presented in three conditions: alone, paired with the nonidentical same-occupation faces, and paired with the different-occupation faces. The number of cycles needed for one of the occupation units, actor and politician, to attain a positive activation value was recorded. (The bias weights, learned during training, were largely inhibitory, leaving the units in a negative state in the absence of input activation.) As usual, 10 replications of the simulation with different random lesions in each of the two pools of units were carried out.

The overt ability of the network to derive occupation information from the face patterns was measured by recording which occupation unit reached threshold (i.e., became positive) after presentation of the face. For trials on which neither unit reached threshold, the network was assumed to guess with probability .5 of being correct. The rationale for scoring performance in this way, rather than taking the larger activation of the two regardless of whether either are positively activated, is that units, like neurons, have a categorical quality to their state. In the present model, there is a categorical difference between the way in which positive- and negative-valued activation in a unit affects the other units to which it is connected. For example, a negatively activated unit will inhibit units to which it is connected by positive weights but will excite them when its activation goes positive. Note that the method of scoring overt categorization was lenient in that we only require the sign of the activation to be correct.

Results and discussion. The performance of the network on the overt occupation decision for faces is shown in Table 6. With lesions to hidden units or input units of 50% and 62.5%, the network's performance falls in the range of 59%–65% correct. This is roughly similar to the performance of the prosopagnosic

Table 5
Settling Time for Novel Patterns Before and After a Small Amount of Learning

| | Number of cycles | | | |
| | Before learning | | After training 5 epochs | |
Amount of damage (%)	M	SE	M	SE
	Hidden unit lesion			
12.5	376.0	18.8	365.0	21.1
25.0	476.2	26.9	430.9	26.3
37.5	475.3	23.9	419.1	30.0
50.0	513.9	29.2	465.8	25.7
62.5	506.6	24.9	438.3	23.6
75.0	521.6	25.7	467.2	25.3
87.5	669.4	35.0	431.6	18.2
	Input unit lesion			
12.5	369.5	22.0	365.8	22.5
25.0	473.3	27.3	446.7	29.3
37.5	464.3	21.4	486.1	33.4
50.0	529.4	31.4	474.3	23.4
62.5	506.3	26.9	462.8	26.6
75.0	538.5	25.3	468.2	21.7
87.5	469.6	22.6	544.2	36.5

Table 6
Overt Occupation Categorization

Amount of damage (%)	Percentage correct	
	M	SE
Hidden unit damage		
12.5	85.5	3.1
25.0	77.0	3.3
37.5	74.0	4.4
50.0	62.5	4.5
62.5	59.5	5.1
75.0	53.0	5.0
87.5	51.5	4.5
Input unit damage		
12.5	88.0	1.2
25.0	86.5	2.3
37.5	73.0	3.4
50.0	64.5	4.4
62.5	59.5	4.9
75.0	57.5	4.8
87.5	57.0	5.0

patient reported by de Haan et al. (1987a, 1987b), who obtained 55.5% and 62.5% correct in comparable tasks.

Table 7 and Figure 6 show the number of cycles required for the correct occupation unit to become positive after presentation of a name, without an accompanying face, and with faces from the same- or different-occupation category. Fewer cycles are required for the occupation units to attain positive values when the face and name are from the same-occupation category than when they come from different-occupaton categories. The effect of the face is evident at all but the most extreme levels of damage. In particular, it is evident at the levels of damage to input and hidden units whose corresponding overt performance was discussed earlier. The data from the no-face condition suggest that, as in de Haan et al.'s (1987a, 1987b) studies, the effect is primarily one of interference rather than facilitation.

The mechanism by which faces affect performance in the present model is as follows: To the extent that the presentation of a face pattern causes any activation to propagate into the rest of the network, this activation will influence the activation of the occupation units, even if it is not in itself sufficient to bring them all the way to threshold. At first glance this would seem to predict both facilitation and interference. Facilitation would arise because the face would contribute activation toward its occupation unit, and if the name has the same occupation, less additional activation from the name pattern would be needed for that occupation unit to attain a positive value. Interference would be predicted because the negative correlation between the two occupation units' activations, over the set of known patterns, would have resulted in an inhibitory connection between them having been learned by the network, so that the activation of either occupation unit would tend to inhibit the activation of the other. In effect, the network learns which subpatterns are consistent and inconsistent with which others, and inconsistent subpatterns (e.g., the single-unit actor or the single-unit politician) will tend to inhibit each other. The lack of an observed

facilitation effect is attributable to mutual inhibition of the patterns for different individuals in the same-occupation category counteracting the facilitation mechanism just described. That is, some of the units activated by the name, which would normally contribute activation to the occupation unit, are themselves being inhibited by the influence of the face pattern.

A similar account has already been presented by Burton et al. (1991) to explain semantic priming by faces in prosopagnosia. They adapted a model of normal face recognition, developed by Burton, Bruce, and Johnston (1990), which was based on the McClelland and Rumelhart (1981) IAC architecture. In the Burton et al. model, there are three pools of units, face recognition units (FRUs), semantic information units, and name units, all connected by way of personal identity units (PINs), which serve to index the appropriate parts of each individual's representation to one another. They model covert recognition by leaving intact all units, including the FRUs, and attenuating the connections between the FRUs and the PINs. Semantic priming by faces is therefore the result of attenuated activation of semantic and name units coming from intact FRUs, which is insufficient to push semantic or name units over threshold for an overt response but lowers the amount of activation from name inputs needed to reach threshold.

The structural similarities between their model and ours include the three types of units (excluding PINs) and the use of a connectionist architecture. Perhaps more important, both models agree that the basic mechanism for semantic priming after damage involves subthreshold activation of semantic representations. The models differ structurally in that their representations are local, whereas ours are distributed, and their con-

Table 7
Time to Categorize Names According to Their Occupation Alone and in the Presence of Same- and Different-Category Faces

Amount of damage (%)	Cycles for correct occupation unit to attain positive activation					
	No face		Same category		Different category	
	M	SE	M	SE	M	SE
Hidden unit damage						
12.5	50.5	5.5	49.0	4.4	142.8	13.6
25.0	49.8	5.9	75.3	9.1	150.5	12.2
37.5	55.3	5.8	76.1	9.3	110.1	8.9
50.0	70.8	13.3	81.6	8.9	114.0	8.8
62.5	66.6	13.9	59.6	6.5	82.1	6.9
75.0	73.0	15.1	84.0	7.6	98.0	7.8
87.5	62.8	9.1	72.9	9.2	66.9	5.0
Input unit damage						
25.0	69.6	9.2	101.7	9.2	77.0	16.0
37.5	55.8	5.6	88.9	8.3	96.3	20.2
12.5	52.0	4.4	115.4	9.3	55.3	9.3
50.0	101.3	10.4	146.5	13.1	126.9	21.6
87.5	76.3	7.6	78.8	6.7	144.0	27.2
62.5	106.0	11.2	131.8	11.3	123.0	24.2
75.0	107.5	10.9	120.6	10.7	130.9	21.9

(A) Face Hidden Unit Lesions

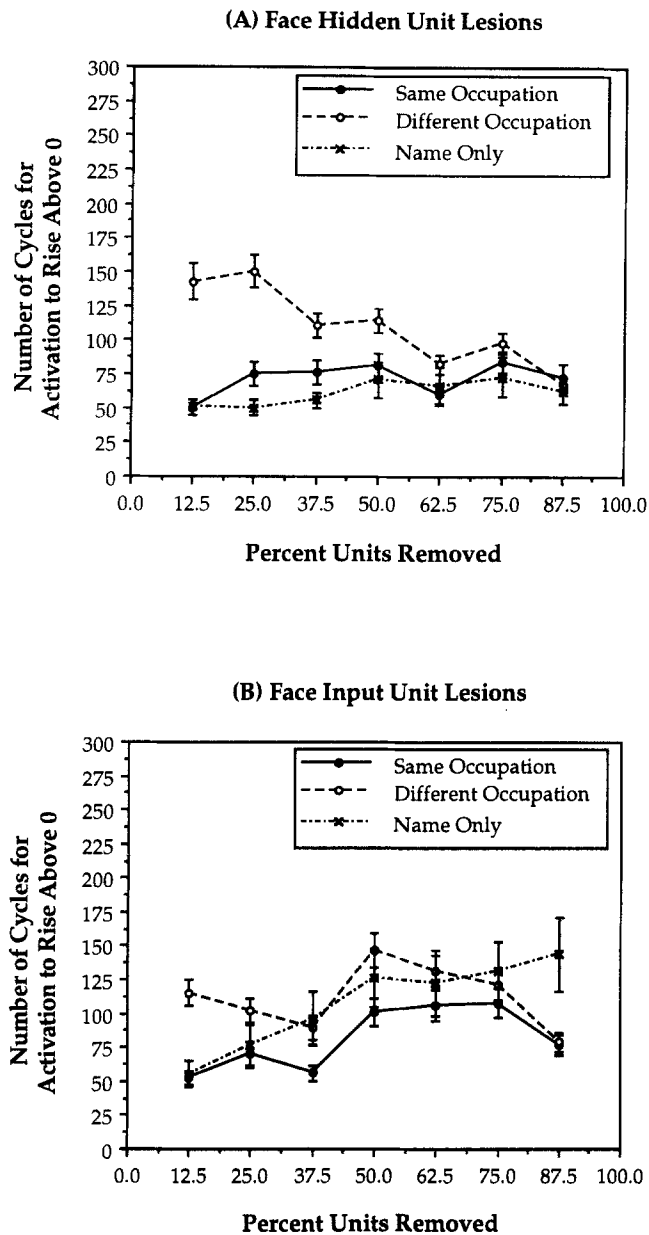## (B) Face Input Unit Lesions

Figure 6. Semantic priming. (Time needed for correct occupation unit to reach threshold when name input is accompanied by face of same occupation, different occupation, or no face, for varying degrees of damage to [A] face hidden units and [B] face input units.)

nections are hand wired, whereas ours are set by learning. One of the most obvious consequences of these differences is that our model can attempt to account for savings in face–name re-learning (Simulation 1), whereas the hand-wired model cannot. A further, indirect result of the learning process in our model was the development of inhibitory relationships among patterns for different individuals, which causes the model to evince more interference than facilitation in the occupation semantic-priming task (Simulation 3), just as real human subjects do. Our use of distributed, as opposed to local, representations allowed us

to simulate the effects of familiarity within that system on set-tling time (Simulation 2). Perhaps most important, in a model with distributed face representations, which undergo "graceful degradation," it seems natural to explore the effects of degrading the face recognition system proper, and this provided the basis for all of our simulations.

The same mechanism proposed here, in principle, explains Young et al.'s (1988) finding of semantic priming of names by associated faces in a name familiarity task. In this case, the relevant locus of priming would not be limited to occupation units but would involve any part of semantics shared by the priming and target individuals. Indeed, Burton et al. (1991) were also able to simulate the two kinds of priming task in the same way.

Simulation 4: Covert Recognition of Overtly Unrecognized Faces

In this final section, we demonstrate that the preserved covert recognition ability in the damaged network is not the result of the network's preserved overt recognition ability for a subset of the familiar patterns. The demonstration takes the form of a series of existence proofs. For each of the three tasks that were simulated, we tested the covert recognition performance of the network just on the subset of faces that it failed to recognize in the overt recognition tests.

Method. A randomly selected 50% of the face hidden units were damaged, and the resulting network was tested on the overt 10-alternative, forced-choice recognition test. The 2 faces out of 10 that were correcty identified were eliminated from the set of test faces. For the semantic-priming experiment, only the 5 faces that were not correctly categorized as actors or politicians were retained in the test set. The damaged network was then tested for covert recognition in the three previously described tasks.

Results and discussion. The network relearned the correct associations among the eight faces and names faster than the incorrect: After damage and before learning, it obtained a score of 0% correct for both the correct and incorrect name–face pairs. After 10 epochs of learning, more learning had taken place for the correct pairs: The network obtained a score of 50% correct for the correct pairs and 0% correct for the incorrect pairs.

As before, presentation of a face from the wrong occupation category delayed the relevant occupation unit from reaching threshold when a name was presented. The mean number of cycles to reach threshold was 70.0 when no face was presented, 33.6 when a face from the same category was presented, and 94.9 when a face from the other category was presented. Settling time in the face units was faster for the 8 previously learned faces than for the 10 novel faces, on average 200.8 and 232.2 cycles, respectively.

In summary, the covert recognition abilities displayed by damaged attractor networks does not depend on the presence in the test set of any overtly identified face patterns.

General Discussion

We have shown that some very general properties of neural networks lead to preserved performance, after network damage, for the types of tasks used to measure covert recognition in prosopagnosia. Specifically, we have simulated in varying de-

grees of detail three types of behavioral tasks used to document covert recognition. At levels of damage associated with low overt identification and categorization performance of face patterns, the network continues to manifest knowledge of the faces when tested by the covert tasks. Of additional interest is the fact that visual portions of the network were damaged in these simulations, demonstrating that one need not conclude that visual recognition is intact in cases of prosopagnosia with covert recognition. In the remainder of this article, we discuss the implications of these results for our understanding of covert face recognition, other covert visual abilities, prosopagnosia, and consciousness.

Covert Face Recognition

Previous attempts to explain covert recognition of faces in prosopagnosia have assumed that covert and overt recognition are dependent on at least partly distinct components of the cognitive architecture, somehow disconnected by brain damage, and that the visual recognition component is intact. In contrast, we have shown that the same system may subserve both overt and covert visual recognition and that damage to this system may spare covert recognition relative to overt recognition. This result is another example of how connectionist architectures can provide unified accounts of dissociations that initially appear to imply the existence of separate mechanisms (Rumelhart & McClelland, 1986; Seidenberg & McClelland, 1989).

Of course, the results of our simulations do not prove that our account is correct, merely that it is possible. Nevertheless, we find it plausible for three reasons: First, it follows from a set of independently motivated computational principles. These include the robustness and efficiency of distributing knowledge among a large number of connections (relevant to all simulations, but especially Simulation 1), the notion that there is a degree of fit between a network and any input pattern (relevant to Simulation 2), and the utility of having a threshold for activation flow between units (relevant to Simulation 3). Second, it is consistent with the available data on overt and covert performance in prosopagnosic patients, specifically the occasional success in overt tasks by these patients. Third, it is a parsimonious account. It is not necessary to invoke separate brain centers for recognition and overt awareness of recognition, and only one face recognition system is hypothesized (cf. Bauer, 1984). Furthermore, consideration of the lesion sites and associated perceptual deficits in cases of prosopagnosia suggest that the visual system is likely to have been damaged.

In our view, the phenomenon of covert recognition in prosopagnosia is no less interesting or important if it is explained in terms of incomplete damage to the face recognition system. The fact that recognition can be manifest in different ways, some of which are accompanied by conscious awareness and others are not, and that this distinction appears to be coextensive with their vulnerability to brain damage, is of obvious high importance to the understanding of perception and the brain. We are merely pointing out that the most straightforward explanation of this dissociation, that the face recognition system is spared and the impairment in overt recognition tasks arises elsewhere, is not the only possibility. In addition to questioning the prevail-

ing hypothesis, we are offering a new one that has the advantage of being more explicit about mechanism.

Covert Recognition in Other Syndromes

In addition to covert recognition in prosopagnosia, there are several other syndromes in which indirect tests of visual perception seem to reveal greater capacities than are apparent on standard overt tests (see Farah, in press, for a review). Can these dissociations also be explained in terms of a single damaged network capable of supporting performance of some tasks and not others? In principle, they could, although there is no reason to assume that all of the syndromes will have the same explanation. In some cases, there is clear evidence favoring the involvement of at least partially distinct systems subserving overt and covert perception. We briefly review the other syndromes and assess the applicability of the present hypothesis, that covert perceptual abilities reflect the functioning of a damaged, but not obliterated, visual system.

Blindsight. The phenomenon of blindsight, in which cortically blind patients who deny having any visual experience can localize and discriminate visual stimuli, was the first neuropsychological dissociation involving conscious awareness to be studied in detail. Although it was initially subject to much skepticism, two decades of careful research have demonstrated to most people's satisfaction that the dissociation is real, and current efforts center on elucidating the specific neural systems responsible for the nonconscious components of visual perception in blindsight (see Weiskrantz, 1990, for a review). Although it has been suggested that the visual abilities in blindsight may be mediated by residual functioning of spared primary visual cortex (Campion, Latto, & Smith, 1983), there is evidence of disproportionate involvement of the subcortical visual system in at least some of these abilities. For example, asymmetries in the processing of visual stimuli between nasal and temporal hemifields suggests that the subcortical visual system (which receives disproportionate input from the temporal hemifield) plays a primary role in some of the covert visual abilities in this syndrome (e.g., Rafal, Smith, Krantz, Cohen, & Brennan, 1990). A wealth of other evidence (summarized by Weiskrantz, 1990) suggests that more than one visual pathway may be involved in the preserved abilities of blindsight patients but that the residual functioning of the primary visual cortex is not a necessary factor.

Implicit reading. Another form of visual recognition in the absence of conscious awareness of recognition can be found in certain patients with pure alexia. Pure alexic patients are, by definition, impaired in reading but have roughly normal auditory word comprehension and writing, and their underlying deficit is therefore inferred to be one of visual word recognition. To the extent that they are able to read, they do so by a slow and laborious letter-by-letter strategy, and their reading can therefore be obliterated entirely by presenting words briefly. However, with brief presentations of words, some pure alexic patients are able to derive considerable information from the words, even though they report being unable to recognize the words and even though they cannot name the words (e.g., Shallice & Saffran, 1986; Coslett & Saffran, 1989). For example, with presentations too brief for any explicit reading, these pa-

tients are able to discriminate words from orthographically legal nonwords and to classify words as belonging to a certain category (e.g., animals or foods) at levels far above chance. Implicit reading may also be carried out by different systems from those subserving normal explicit reading. The hypothesis of right-hemisphere mediation of implicit reading (in contrast with the predominant role of the left hemisphere in normal reading) is supported by the relative absence of implicit reading for abstract words, function words, and grammatical inflections, and the lack of access to phonology, all characteristics of the right-hemisphere lexicon (Coslett & Saffran, 1989). Nevertheless, it is conceivable that this profile of reading abilities would also emerge from damage to the left-hemisphere reading system, as discussed by Shallice and Saffran (1986). For example, differences between word classes such as word frequency and availability of collateral support from semantic representations may confer different degrees of robustness to damage on them. However, differences in the regularity of mapping among print, meaning, and phonology could also affect the robustness of these mappings in the network after damage and would seem to predict lesser rather than greater vulnerability of phonology relative to semantics.

Unconscious perception of extinguished visual stimuli. Extinction refers to the impairment in perception of a contralesional stimulus when presented simultaneously with an ipsilesional stimulus. Volpe, Ledoux, and Gazzaniga (1979) tested the ability of right parietal-damaged patients to perceive contralesional visual stimuli in two ways. First, the patients were shown a tachistoscopic presentation of a pair of stimuli (line drawings or words), one on each side of fixation, and were asked to name what they saw. In this task, the patients manifested visual "extinction" of the left stimulus by the right, which is typical of right parietal-damaged patients: the right stimulus was generally named correctly, but the left stimulus was not, and patients sometimes even denied that the left stimulus had been presented. In contrast, the patients performed well in a second kind of task with the same stimuli. When asked whether the two stimuli presented on a given trial were the same or different, the patients were highly accurate, even though this task requires perception of the left stimulus. Volpe et al. interpreted their findings as revealing "a breakdown in the flow of information between conscious and nonconscious mental systems" (p. 724). However, it is possible that these findings can be explained in terms of the residual functioning of a damaged visual system, rather than a dissociation between two types of system, one conscious and one unconscious.

Farah, Monheit, and Wallace (1991) showed that the dissociation observed by Volpe et al. could be obtained in normal subjects simply by placing a translucent sheet of drafting stock over the left half of the display to degrade subjects' perception of stimuli on the left. We also showed that the dissociation could be eliminated in parietal-damaged patients with extinction when the overt and covert tasks were matched for the precision of visual perception required by each. This implies that the dissociation between overt and covert perception after parietal damage is also due to differences in the quality of information needed to support performance in the two types of task, with performance in the covert task again more robust to low-quality information. The nature of the information degradation ap-

pears to be different in the two cases, however. In prosopagnosia, what is degraded is the pattern of previously learned associations within the visual recognition system, so that the effects of prior learning on perception are disrupted. In extinction, there is no structural impairment of representations, as evidenced by the ability of patients with extinction to perceive normally in the absence of a simultaneously occuring ipsilesional stimulus. Rather, the locus of degradation appears to be before visual recognition, affecting the input to visual recognition memory. This is consistent with our ability to simulate covert recognition in extinction by degrading the stimulus input to normal subjects.

The most general implication of the present model for the study of covert perception is that it demonstrates another mechanism by which overt and covert processing can be dissociated, beyond those previously considered by neuropsychologists. Schacter et al. (1988) listed three general types of account for overt–covert dissociations: conservative response bias in the overt tasks, disconnection from language (on the assumption that language is more involved in the overt tasks) and truly distinct and thus dissociable processing systems for overt and covert performance. To these we would add a fourth: differential susceptibility to damage of overt and covert performance. We have shown how knowledge can reside in a damaged network but be inaccessible for most purposes, for reasons quite distinct from either the signal-detection theory concept of bias or a disconnection from other systems. As Hinton and Plaut (1987) showed, savings in relearning reveals the knowledge retained in damaged networks that is not apparent in the overt performance of the network. We have extended these findings to two other measures of covert knowledge in damaged networks: speed of settling and priming.

Prosopagnosia

The finding that some prosopagnosic patients manifest covert recognition and others do not has been taken as an indication that there are two different types of prosopagnosia: one caused by a visual perceptual impairment per se and the other caused by a disconnection of visual recognition and other, conscious, mental systems (e.g., Bauer, 1986; Bruyer, 1991; Burton et al., 1991; de Haan et al., 1992; Newcombe et al., 1989). However, our analysis suggests that these two groups of prosopagnosic patients are more likely to differ in severity than in kind. In particular, the similarity of the effects obtained when we lesioned face input units and face hidden units suggests that the presence of covert recognition may not be a precise way of discriminating different functional loci of damage. In fact, lesions further downstream in our model also showed similar effects to the ones reported here. This is a consequence of the highly interactive nature of the model. The nonlocalizability of errors resulting from damage in interactive models has been discussed in detail by Hinton and Shallice (1991) for their model of reading. This is not to deny that prosopagnosia may be accompanied by, or may even result from, impairments of different levels of visual face processing.

Consciousness

The dissociations between covert and overt perception in prosopagnosia and in other syndromes are of interest indepen-

dent of the association between overt perception and consciousness. The fact that knowledge may be accessible in certain tasks and not in others is somewhat counterintuitive and promises insights into how information is represented in the brain. Indeed, this has been the focus of the present article. However, it cannot be denied that part of the fascination of these dissociations comes from the involvement of consciousness, specifically the patients' seemingly earnest denials of conscious awareness of stimulus properties of which they show knowledge in certain tasks. On the basis of our research, can we offer any insights into consciousness?

We believe that neuropsychological evidence can indeed answer certian questions about consciousness, specifically those concerning its physical correlates. In the present case, it appears that conscious awareness of recognition is correlated with a certian minimal quality of information representation within the visual recognition system. Why should the likelihood of conscious awareness depend on the quality of the representation? Kinsbourne (1988) has suggested that stimuli reach conscious awareness only when they are integrated into the global information-processing state of the brain as a whole. He has also suggested that faulty perceptual processing of a stimulus will decrease the likelihood that it will become integrated. This idea is consonant with the dynamics of neural networks, particularly attractor networks such as the one presented here: When a new input pattern is presented, it pulls the network into a new attractor state, with each part of the network taking on new activation values that are integrated in the sence of being mutually compatible. When the input patterns are degraded by damage to perceptual units, they loose their ablility to pull the remaining parts of the network into an integrated state consistent with themselves (i.e., semantics and name units no longer take on the patterns corresponding to the presented face), even for levels of degradation at which residual information can be detected by other means (that do not involve propagating information throughout the system). In summary, our model suggests that an important neural correlate of conscious awareness in perception is the quality of perceptual representation, perhaps because perceptual quality limits the ability of perceptual representations to draw other parts of the system into an integrated state with themselves.

References

Bauer, R. M. (1984). Autonomic recognition of names and faces in prosopagnosia: A neuropsychological application of the guilty knowledge test. *Neuropsychologia, 22,* 457–469.

Bauer, R. M. (1986). Aspects of face processing. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *The cognitive psychophysiology of prosopagnosia* (pp. 253–267). Dordrecht, The Netherlands: Martinus Nijhoff.

Bruyer, R. (1991). Covert face recognition in prosopagnosia: A review. *Brain and Cognition, 15,* 223–235.

Bruyer, R., Laterre, C., Seron, X., Feyereisen, P., Strypstein, E., Pierrard, E., & Rectem, D. (1983). A case of prosopagnosia with some preserved covert remembrance of familiar faces. *Brain and Cognition, 2,* 257–284.

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology, 81,* 361–380.

Burton, A. M., Young, A. W., Bruce, V., Johnston, R. A., & Ellis, A. W. (1991). Understanding covert recognition. *Cognition, 39,* 129–166.

Campion, J., Latto, R., & Smith, Y. M. (1983). Is blindsight an effect of scattered light, spared cortex, and near-threshold vision? *The Behavioral and Brain Sciences, 3,* 423–447.

Coslett, H. B., & Saffran, E. M. (1989). Evidence for preserved reading in 'pure alexia.' *Brain, 112,* 327–359.

de Haan, E. H. F., Bauer, R. M., & Greve, K. W. (1992). Behavioral and physiological evidence for covert recognition in a prosopagnosic patient. *Cortex, 28,* 77–95.

de Haan, E. H. F., Young, A. W., & Newcombe, F. (1987a). Faces interfere with name classification in a prosopagnosic patient. *Cortex, 23,* 309–316.

de Haan, E. H. F., Young, A. W., & Newcombe, F. (1987b). Face recognition without awareness. *Cognitive Neuropsychology, 4,* 385–415.

Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences, 15,* 183–247.

Etcoff, N. L., Freeman, R., & Cave, K. R. (1991). Can we lose memories of faces? Content specificity and awareness in a prosopagnosic. *Journal of Cognitive Neuroscience, 3,* 25–41.

Farah, M. J. (in press). Visual perception and visual awareness after brain damage: A tutorial overview. In M. Moscovitch & C. Umilta (Eds.), *Conscious and unconscious information processing: Attention and performance XIV.* Cambridge, MA: MIT Press.

Farah, M. J., Monheit, M. A., & Wallace, M. A. (1991). Unconscious perception in the extinguished hemifield: Re-evaluating the evidence. *Neuropsychologia, 29,* 949–958.

Greve, K. W., & Bauer, R. M. (1990). Implicit learning of new faces in prosopagnosia: An application of the mere exposure paradigm. *Neuropsychologia, 28,* 135–141.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations* (pp. 77–109). Cambridge, MA: MIT Press.

Hinton, G. E., & Plaut, D. C. (1987). Using fast weights to deblur old memories. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Hinton, G. E., & Sejnowski, T. J. (1986). Learning and relearning in Boltzmann machines. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations* (pp. 282–317). Cambridge, MA: MIT Press.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98,* 74–95.

Jacoby, L. L. (1984). Incidental vs. intentional retrieval: Remembering and awareness as separate issues. In L. R. Squire & N. Butters (Eds.), *The neuropsychology of human memory.* New York: Oxford University Press.

Kinsbourne, M. (1988). An intergrated field theory of consciousness. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 159–182). Oxford, England: Claredon Press.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88,* 375–407.

Movellan, J. R. (1990). Contrastive Hebbian learning in the continuous Hopfield model. In D. S. Touretzky, G. E. Hinton, & T. J. Sejnowski (Eds.), *Proceedings of the 1989 Connectionist Models Summer School* (pp. 10–17). San Mateo, CA: Morgan Kaufman.

Newcombe, F., Young, A., & de Haan, E. H. F. (1989). Prosopagnosia and object agnosia without covert recognition. *Neuropsychologia, 27,* 179–191.

Peterson, C., & Anderson, J. A. (1987). A mean field theory learning algorithm for neural networks. *Complex Systems, 1,* 995–1019.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Erlbaum.

Rafal, R., Smith, J., Krantz, J., Cohen, A., & Brennan, C. (1990). Extrageniculate vision in hemianopic humans: Saccade inhibition by signals in the blind field. *Science, 250,* 118–121.

Renault, B., Signoret, J. L., Debruille, B., Breton, F., & Bolgert, F. (1989). Brain potentials reveal covert facial recognition in prosopagnosia. *Neuropsychologica, 27,* 905–912.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations.* Cambridge, MA: MIT Press.

Schacter, D. L., McAndrews, M. P., & Moscovitch, M. (1988). Access to consciousness: Dissociations between implicit and explicit knowledge in neuropsychological syndromes. In L. Weiskrantz (Ed.), *Thought without language* (pp. 242–278). New York: Oxford University Press.

Seidenberg, M., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Sergent, J., & Villemure, J. G. (1989). Prosopagnosia in a right hemispherectomized patient. *Brain, 112,* 975–995.

Shallice, T., & Saffran, E. (1986). Lexical processing in the absence of explicit word identification: Evidence form a letter-by-letter reader. *Cognitive Neuropsychology, 3,* 429–458.

Snodgrass, J. G., Corwin, J., & Feenan, K. (1990). The pragmatics of measuring priming: Applications to normal and abnormal memory. *Journal of Clinical and Experimental Neuropsychology, 12,* 61.

Tranel, D., & Damasio, A. R. (1985). Knowledge without awareness: An automatic index of facial recognition by prosopagnosics. *Science, 228,* 1453–1454.

Tranel, D., & Damasio, A. R. (1988). Nonconscious face recognition in patients with face agnosia. *Behavioral Brain Research, 30,* 235–249.

Volpe, B. T., Ledoux, J. E., & Gazzaniga, M. S. (1979). Information processing of visual stimuli in an "extinguished" field. *Nature, 282,* 722–724.

Wallace, M. A., & Farah, M. J. (1992). Savings in relearning as evidence for covert recognition in prosopagnosia. *Journal of Cognitive Neuroscience, 4,* 150–154.

Weiskrantz, L. (1990). Outlooks for blindsight: Explicit methodologies for implicit processes. *Proceedings of the Royal Society of London, B239,* 247–278.

Young, A. W., & de Haan, E. H. F. (1988). Boundaries of covert recognition in prosopagnosia. *Cognitive Neuropsychology, 5,* 317–336.

Young, A. W., Hay, D. C., & Ellis, A. W. (1985). The faces that launched a thousand slips: Everyday difficulties and errors in recognising people. *British Journal of Psychology, 76,* 495–523.

Young, A. W., Hellawell, D., & de Haan, E. H. F. (1988). Cross-domain semantic priming in normal subjects and a prosopagnosic patient. *Quarterly Journal of Experimental Psychology, 38A,* 297–318.

1994 APA Convention "Call for Programs"

The "Call for Programs" for the 1994 APA annual convention appears in the September issue of the *APA Monitor.* The 1994 convention will be held in Los Angeles, California, from August 12 through August 16. The deadline for submission of program and presentation proposals is December 3, 1993. Additional copies of the "Call" are available from the APA Convention Office, effective in September. As a reminder, agreement to participate in the APA convention is now presumed to convey permission for the presentation to be audiotaped if selected for taping. Any speaker or participant who does not wish his or her presentation to be audiotaped must notify the person submitting the program either at the time the invitation is extended or before the December 3 deadline for proposal submission.