PERGAMON

Neuropsychologia 40 (2002) 187–204

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Dissociating 'what' and 'how' in visual form agnosia: a computational investigation

Shaun P. Vecera *

Department of Psychology, 11 Seashore Hall E, University of Iowa, Iowa City, IA 52242-1407, USA

Received 24 January 2000; received in revised form 25 March 2001; accepted 3 April 2001

Abstract

Patients with visual form agnosia exhibit a profound impairment in shape perception (what an object is) coupled with intact visuomotor functions (how to act on an object), demonstrating a dissociation between visual perception and action. How can these patients act on objects that they cannot perceive? Although two explanations of this 'what–how' dissociation have been offered, each explanation has shortcomings. A 'pathway information' account of the 'what–how' dissociation is presented in this paper. This account hypothesizes that 'where' and 'how' tasks require less information than 'what' tasks, thereby allowing 'where/how' to remain relatively spared in the face of neurological damage. Simulations with a neural network model test the predictions of the pathway information account. Following damage to an input layer common to the 'what' and 'where/how' pathways, the model performs object identification more poorly than spatial localization. Thus, the model offers a parsimonious explanation of differential 'what–how' performance in visual form agnosia. The simulation results are discussed in terms of their implications for visual form agnosia and other neuropsychological syndromes.

Keywords: Vision; Apperceptive agnosia; Visual form agnosia; Dorsal visual pathway; Ventral visual pathway

1. Introduction

A major finding that has influenced vision science since the early 1980s is the discovery of relatively separate visual pathways for object identification (what an object is) and spatial localization (where an object is located; see Refs. [31,47]). Ungerleider and Mishkin [47] proposed that object identification is computed by the ventral pathway from the occipital lobe into the temporal lobe, and that spatial localization is computed by the dorsal pathway from the occipital lobe into the parietal lobe. The evidence for separate 'what' and 'where' pathways was based originally on lesion studies in monkeys, and a number of neuroimaging (e.g. Ref. [46]), neuropsychological (e.g. Ref. [14,17,23]), and computational investigations (e.g. Ref. [10,24,40]) have since supported the separation of identity and location in the primate visual system.

Goodale and coworkers [19–21,29,30], based on neuropsychological studies, have suggested that the appropriate division of processing is between visual perception and visuomotor processing, rather than a division in identification and localization. Goodale and Milner thus have proposed a 'what' and 'how' division for the primate posterior cerebral cortex as an alternative to 'what' and 'where.' This 'what–how' division has been supported by a double dissociation between visual perception and visuomotor action in neuropsychological patients. The two patient groups that form this double dissociation are patients with apperceptive agnosia, or visual form agnosia, and patients with optic ataxia (see Ref. [21]).

One could argue that the separation of 'what' and 'where' or 'how' information is a basic principle in vision research. But why are visual perception and visuomotor control (action) dissociated in neuropsychological patients? What are the mechanisms that give rise to the inability of some patients to recognize objects (visual form agnosia) and the inability of other patients to act on objects (optic ataxia)? I explore one possible

* Tel.: +1-319-335-0839; fax: +1-319-3350191.
E-mail address: shaun-vecera@uiowa.edu (S.P. Vecera).

0028-3932/01/$ - see front matter
PII: S0028-3932(01)00072-0

mechanism of this 'what–where/how' dissociation in this paper, and I focus primarily on patients with visual form agnosia.

Patients with visual form agnosia[1] are characterized by an inability to identify objects and to copy visually presented objects following damage to the posterior visual cortex (see Ref. [14] for a review). Lissauer [27] noted that these patients could not form an adequate visual percept, which prevented object identification. In most cases, damage results from an anoxic episode, a lack of oxygen to the brain, often from carbon monoxide poisoning (e.g. Refs. [1,2,4,8,9,30]) or from heart failure (e.g. Ref. [48]). The damage tends to be diffuse, covering the posterior cortices, particularly the occipital lobe. In addition to the relative homogeneity in neuropathology, these patients share similar visual object identification problems, problems that are due to low-level impairments. One early-level ability that visual form agnosics have lost is simple shape matching. Mr S, Efron's [13] patient, was unable to discriminate two rectangular shapes that were matched for area, and he had an inability to distinguish between an X and an O. Adler's [1] patient was unable to discriminate between a circle and a triangle. Campion's [8] patient, RC, described a circle as 'a lot of dots' and was unable to distinguish an 'M' from a 'W.' The ability to copy visually presented shapes, which may require a relatively intact shape representation, also appears disrupted in visual form agnosia. (See Refs. [1,2,4] for published examples of patients' disrupted copying abilities.) Finally, recent analyses of the visual disturbances in these patients have implicated damaged preattentive visual processes, such as perceptual organization and figure–ground segregation (see Refs. [14,48–50]). Patients with visual form agnosia do not appear to be able to use gestalt grouping cues or figure–ground cues to organize the visual field into meaningful shapes [30,48]. For example, patients DF [30] and JW [48] were unable to segregate two overlapping line segments based on gestalt good continuation.

Although patients with visual form agnosia exhibit impaired visual perception and identification, they seem to posses intact visuomotor control — for example, they can reach out and grasp shapes appropriately. Patient DF [30] could and align her hand appropriately to place a card in an oriented slot (see Refs. [19,20]), although she could not perceive form or shape. DF also could grasp objects at appropriate points with a 'precision grip' [21], and could by accurately grasp objects by calibrating the aperture between her thumb and forefinger [19]. Similarly, patient JW, who was unable to perceive form or to use gestalt grouping cues [48], could scale the opening of his hand to correspond to the size of an object for which he was reaching [28]. This dissociation is puzzling: How can visual form agnosics act on an object that they cannot perceive?

[1] Because 'apperceptive agnosia' is used to refer to different patient groups (contrast Ref. [14] with Ref. [51], for example), I will refer to the syndrome as 'visual form agnosia' to avoid confusion. My use of visual form agnosia corresponds to Farah's [14] narrow definition of apperceptive agnosia.

2. Two explanations of the perception–action dissociation in visual form agnosia

How can the dissociation between visual perception and visuomotor control (action) in visual form agnosia be explained? Answers to this question are important because they could inform us about the organization of the primate visual system. The existence of the perception–action dissociation automatically rules out certain architectures for the visual system. For example, the 'what–how' dissociation rules out a two-stage architecture in which form perception processes provide the sole input to visuomotor processes; in this architecture, damage to form processes also would impair visuomotor processes, and this is a result inconsistent with results from visual form agnosics.

There are two straightforward accounts of the 'what–how' dissociation, both of which stem from insights based on neuropsychological inference (see Ref. [19] for a review). In what follows, I critically examine each of these explanations for the 'what–how' dissociation in visual form agnosia and find weaknesses with each. Because of the inadequacies with the two standard accounts, I propose a third explanation of the 'what–how' dissociation observed by Goodale and coworkers.

Both of the standard accounts assume that there are two cortical visual streams, the ventral stream which mediates visual perception and the dorsal stream which mediates visuomotor control. An additional shared assumption of these explanations is that the two cortical pathways receive shared input from early-level visual areas in the occipital lobe (see Ref. [19] for a review). The shared assumptions of these explanations are shown in the simplified visual architecture shown in Fig. 1.

The first account is the *ventral disconnection account*, depicted in Fig. 1A. In this account, visual form agnosia occurs because of damage along the ventral processing pathway (the 'what' pathway). Because the neural damage in visual form agnosia is diffuse, the circumscribed lesion depicted in Fig. 1A represents the functional consequence of brain damage. Damage along the ventral pathway impairs the connection between early-level vision and shape perception, preventing adequate shape representation. However, the inputs from early-level vision to the dorsal action pathway are intact, thus sparing visuomotor skills (see Refs. [19, p. 190; and 30, pp. 424–425]).

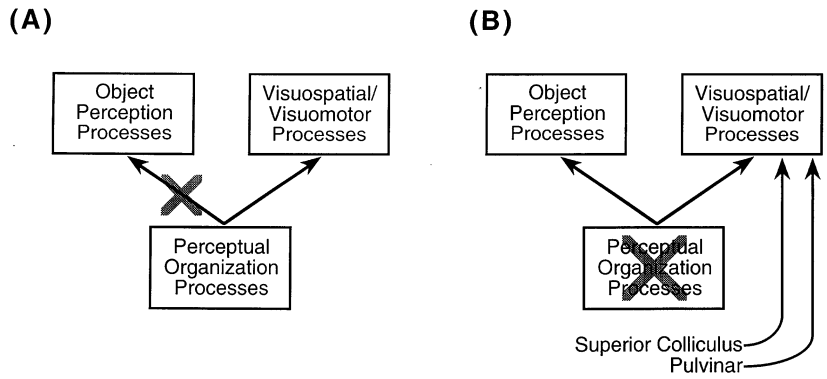

Fig. 1. Two possible explanations for the 'what–how' dissociation observed in patients with visual form agnosia. (A) The ventral disconnection account, in which damage occurs along the ventral 'what' pathway, sparing processing along the dorsal 'where–how' pathway. (B) The spared input account, in which damage occurs to an input common to both pathways. Undamaged inputs to the dorsal pathway, possibly from the superior colliculus and pulvinar, spare processing in this pathway.

Although the ventral disconnection account is straightforward, there are two shortcomings with the account. The first shortcoming is based empirically: the ventral disconnection account would have difficulty explaining some recent data on perceptual organization in patients with visual form agnosia. Early cortical visual areas play a role in gestalt perceptual organization, including figure–ground organization (for recent neurophysiological evidence see Refs. [26,39,52]). Therefore, if early-level vision was intact and provided inputs to visuomotor processes, early-level visual processes such as segregation should also be intact. But early-level perceptual organization processes are disrupted in visual form agnosics (see Refs. [30,48]). The second shortcoming of the ventral disconnection account is based theoretically: this account implicitly makes the locality assumption [15] in which (1) the cognitive architecture is modular and (2) brain damage has local effects (i.e. effects on a module and the outputs of that module). As Farah [15] has pointed out, the locality assumption in neuropsychology is unlikely to be correct, and more biologically plausible approaches based on neural network models may lead to a better understanding of normal cognitive processes.

The second account of the dissociation in visual form agnosia is the *spared input account*, depicted in Fig. 1B. Under this account, the primary damage occurs to early-level visual areas which provide the input to both pathways. Visuomotor function is preserved because additional inputs exist to the dorsal system (see Refs. [19, p. 190;21, p. 605;29]). There is ample evidence that the superior colliculus projects to parietal lobe visual areas via the pulvinar, and these preserved inputs may allow visually guided action to remain intact. There are at least two shortcomings with this account, however. First, there are undoubtedly intact subcortical projections to ventral stream visual areas (e.g. the lateral geniculate nucleus projects to extrastriate visual areas that lie within the ventral stream; [11,18,38,44]). Thus, a simple spared input account also would predict that ventral processing (i.e. object perception) could be relatively intact. Second, and more important, although appealing to spared anatomical inputs is entirely reasonable, appealing to such spared inputs does not inform visual theorists about the function of these inputs. Additional information, such as single-unit recording data that demonstrate the effects of pulvinar or collicular inputs on the dorsal stream, would strengthen the spared input account.

In the next section, I propose a third account of the 'what–how' dissociation in visual form agnosia. It is not my intent to rule out the ventral dissociation and spared input accounts; rather, I intend my account to provide another means by which dissociated 'what–how' performance could arise from a damaged neural system. My account is aimed at addressing the shortcomings of the ventral disconnection and spared input accounts.

3. A third explanation: the 'pathway information' account

In my *pathway information account*, I hypothesize that a common input to the ventral and dorsal streams has been damaged, as in the spared input account. However, unlike the shared input account, I suggest that differences between perceptual and visuomotor representations may explain the dissociation observed in visual form agnosia. The input–output transformations that occur in the dorsal 'where' pathway may be computationally easier than the input–output transfor-

mations that occur in the ventral 'what' pathway. Further, the *separability* of the input representations may lead to the difference in the input–output transformations. Separability refers to the degree to which a set of patterns' features overlap with one another; patterns that share many overlapping features are less separable than patterns that share fewer features. Object identity may be less separable than spatial location because an individual object will share many features with other objects. For example, handles can appear on many objects — cups, suitcases, and buckets — which prevent handles from being a unique feature. In contrast, an individual location will share features with few locations (only those locations that are nearby). The consequence of greater separability for representations of object identity than for spatial location is that object identity is a more computationally complex task than is spatial location. The greater separabilty of location inputs than identity inputs may allow the dorsal 'where' pathway to be more resilient to neural damage than the ventral 'what' pathway. Thus, the neurological damage seen in patients with visual form agnosia may degrade the inputs to both the ventral and dorsal pathways, with more detrimental consequences for the ventral stream than for the dorsal stream because the neural damage may further reduce the separability of the input representations.

The pathway information account explains the 'what–how' dissociation as follows: the degraded inputs along the 'what' pathway prevent shape perception; the same degraded inputs along the 'where–how' pathway disrupt but do not prevent localization because this pathway may operate relatively well in the face of noisy or limited input. This account predicts that a lesion to an input shared by both the 'what' and 'where' pathways would allow 'what' and 'where' tasks to be dissociated, precisely the dissociation observed in some patients with visual form agnosia.

Another consequence of this account is that it may explain the impaired gestalt organization observed in visual form agnosics, results that the ventral disconnection account may not easily explain. Perceptual organization would be impaired because of the disruption of early-level cortical visual areas and processes, those areas involved in perceptual organization and figure–ground segregation, as noted earlier. However, elementary visual functions, such as acuity or color perception, could remain relatively intact in these patients because these processes may rely on neural structures that operate before the level of gestalt organization processes.

Note that the pathway information account goes beyond a simple 'performance threshold' analysis, which proposes that the object recognition simply is more difficult than spatial localization. This difference between recognition and localization may be obscured in neurologically normal adults because of extensive practice with both tasks. The important point made by my pathway information account is that the structure of the neural representations of each task may differ, a point that provides a specific reason that identification may be more difficult than localization. The object identity task may require a higher-dimensional representational space to distinguish objects that are highly similar compared to the localization task. These representational differences may emerge only under large amounts of neural noise, such as neurological damage or brief, masked presentations in normal control subjects.

It is difficult to test the pathway information account with neuropsychological patients because of the difficulties associated with measuring the amount of information required by the 'what' and 'where–how' processing streams. The plausibility of the pathway information approach can, however, be examined with neural network simulations in which the separability of identity information and location information can be determined within the network's simulated environment. For example, a network could be trained on an environment that contains more similarity among objects' shapes than among spatial positions. This environment would place greater demands on the 'what' task than on the 'where' task. Performance following a lesion could then be assessed.

In what follows, I present four simulations that test the idea that the behavior of visual form agnosia results from differential similarity structure required by the 'what' and 'where–how' pathways. I used a simple neural network, originally tested by Rueckl et al. [40], that learns to correctly identify objects along one pathway (the 'what' pathway) and learns to correctly localize those objects along a second pathway (the 'where–how' pathway). In Simulation 1, I ask if the model exhibits any phenomena that would suggest differences between the 'what' and 'where' pathways. If such a difference existed, it could provide a computational mechanism for differential performance between the two pathways following a lesion. In Simulation 2 I conduct a lesion analysis with the model to measure the consequences on the two pathways of lesioning their shared input. The pathway information account predicts that a lesion to the shared input of the two pathways should impair object identification more than spatial processing. In Simulation 3 I show that equating the similarity structure of the two pathways abolishes the post-lesion differences between the two pathways. Finally, in Simulation 4 I replicate the results of Simulation 2 using a 'where–how' task that more closely approximates the tasks used with neuropsychological patients. The results of these simulations provide an 'existence proof' of the pathway information account, suggesting that this account should be added to the list

of possible explanations of the 'what–how' dissociation in visual form agnosia.

4. Simulation 1: exploration of a split-pathway model

For convenience, in the remainder of the paper I refer to the network's pathways as the 'what' and 'where' pathways. As noted in Section 1, Goodale [19] has argued that the dorsal pathway can be viewed as having a central role in the control of movements within visual space. Thus, the dorsal pathway may be better described as a 'how' pathway, responsible for the coordination of visually guided actions. My use of 'where' in place of 'how' to describe the dorsal pathway follows the terminology of Rueckl et al. [40] and is a term generally applied to the dorsal processing pathway. I would not disagree that the dorsal pathway is involved in visuomotor functions, as implied by numerous single-unit recording studies (for reviews see Refs. [25,42]). The dorsal pathway likely underlies a variety of visual functions, including spatial attention (e.g. Refs. [12,37]), metric information about the size of an object or the center of mass of an object (e.g. Refs. [6,21]), and information about the orientation of surfaces [43]. These visual functions may be used to guide visuomotor functions (e.g. Ref. [19]).

My goal in Simulation 1 was to determine if the 'what' and 'where' pathways differ in the necessary amount of network resources (e.g. units) when a network learns the same number of objects and locations. If the pathways do differ in their processing requirements, then the 'what–how' dissociation in visual form agnosia could be the result of this difference. The network I used to study visual form agnosia is shown in Fig. 2. There were five layers in this network, and these five layers were divided into two pathways with a shared input to the pathways. The network takes an object in a location as input. From this input, the model learns to identify the object irrespective of its location (the 'what' task) and to localize the object irrespective of its identity (the 'where' task).

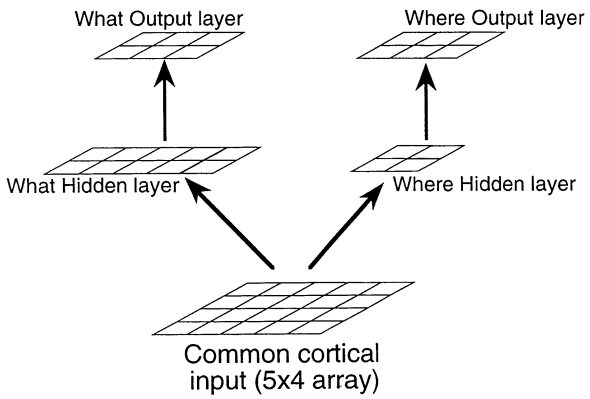

Fig. 2. The general architecture of the split-pathway neural network used in the present simulations. See text for the details of each layer and patterns of connectivity.

The operation of the model is similar in spirit to Goodale and Milner's [19,29] 'transformation account,' in which the 'what' and 'where' pathways receive the same input but transform this input differently. The focus of the transformation account is on the output systems to which the 'what' and 'where' pathways project. The operation of the model described below can be described in similar terms. Knowing the input to the model is insufficient for understanding the functional differences between the 'what' and 'where' pathways; one also must know the input-to-output transformation that is performed by each pathway. As will be evident from the simulation results, the transformation along the model's 'what' pathway is more computationally demanding than the transformation along the model's 'where' pathway.

Because Rueckl et al. [40] did not explore differences between the 'what' and 'where' pathways in detail, in Simulation 1, I investigated two ways in which the 'what' and 'where' pathways may differ from one another. First, following Rueckl et al. [40], I manipulated the resources (units) allocated to each of the two pathways. This manipulation amounts to changing the configuration of the network to determine the critical number of units required to perform each of the two tasks. Second, I examined the time course of learning along the two pathways. There is some evidence from my own preliminary simulations and from Rueckl et al.'s [40] report that the 'where' task may be learned faster than the 'what' task, although both tasks are learned equally well by the end of training. Both the resource allocation and the initial learning advantage for the 'where' task over the 'what' task could underlie the relative preservation of function along the 'where' pathway observed in visual form agnosia.

4.1. Network architecture

The network architecture is shown in Fig. 2. There were five different layers of units, beginning with the Input layer that presented input to both pathways. The Input layer consisted a 5×4 array. The 'what' pathway contained the What Hidden layer, which contained units that received input from the Input layer and projected to the What Output layer. There were a variable number of units in the What Hidden layer, depending on the network's configuration. The What Output layer contained six units, each of which coded the identity of one of six different objects.

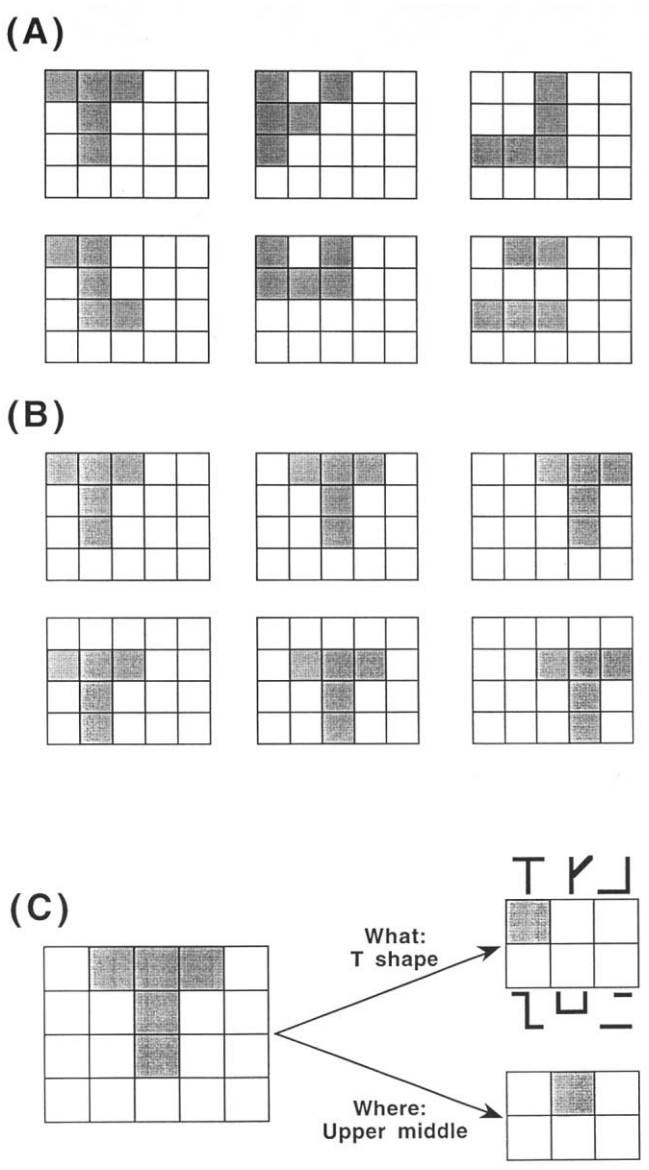

Fig. 3. A representation of the 'what' and 'where' tasks learned by the network. (A) The six objects (each appearing in the same location). (B) The six locations (shown with the same object). (C) Each pathway receives a 5×4 unit array as input, which contains an object in a location. The network learns to identify the object irrespective of where it appears (the 'what' task) and to localize the object irrespective of what it is (the 'where' task).

The 'where' pathway contained the Where Hidden layer, which received input from the Input layer. The Where Hidden layer also had a variable number of units depending on the network's configuration. The Where Output layer contained six units, each of which represented one of six locations in which an object could appear. The coding scheme of the What Output and the Where Output layers is shown in Fig. 3. The network architecture consisted of feedforward projections only. There was full interconnectivity between two connected layers (i.e. every unit in one layer projected to every other unit in the following layer).

In Simulation 1 I examined the performance of five network configurations on 'what' and 'where' tasks. The configurations consisted of a pool of 14 hidden units; these 14 units were differentially allocated to the 'what' and 'where' tasks. The network configurations are referred to by the number of hidden units in the 'what' and 'where' pathways, respectively. For example, in one network configuration, 8 hidden units were allocated to the 'what' pathway and 6 hidden units were allocated to the 'where' pathway; this network architecture formed the 8/6 configuration. The other configurations were 9/5 (i.e. 9 'what' hidden units and 5 'where' hidden units), 10/4, 11/3, and 12/2. All of these network configurations give more units to the 'what' pathway because, as reported by Rueckl et al. [40], the 'what' pathway needs more units to learn as effectively as the 'where' pathway. The network configurations I chose will allow me to equate 'what' and 'where' performance, which would be analogous to comparable identification and localization in neurologically normal adults.

4.2. Network training

All network configurations were given the same two tasks of identifying the object in the input image and localizing that object to a specific spatial region in the image. These 'what' (or identification) and 'where' (or localization) tasks were performed concurrently. Each network was trained on 36 images, one for each object/location pairing. To correctly identify an object, the network needed to activate the one output unit that represented that object while simultaneously not activating (or 'turning off') the other five output units. To correctly localize the object, the network needed to activate the one output unit that represented that location while simultaneously not activating the other five 'where' output units.

The networks' task was to take an image and correctly produce the identity and location of that object. Any discrepancy between the output generated by the network and the correct representation of identity and location is scored as an error termed the sum of squared error (SSE). For each of the input patterns, the difference between the networks' output and the correct ('target') output was calculated. This difference was then squared, and these squared differences were summed across the 36 input patterns, resulting in the SSE. The SSE was computed separately for the 'what' and 'where' tasks. The networks were trained with the backpropagation learning algorithm (also known as the generalized delta rule; see Ref. [41]). The weights in each network were randomized initially; the value of each weight in the network was set by randomly selecting from a Gaussian distribution with a mean weight value of 0 and a standard deviation of 0.5.

During a training epoch, a network was shown the 36 input patterns individually; for each pattern, the network's output was then compared with the correct representation, and the SSE was computed. The networks' weights were adjusted with the backpropagation algorithm to reduce the SSE (for details see Refs. [3,41]) The parameters used in the present simulations are standard, and in preliminary simulations I found that minor changes in these parameters did not alter the qualitative performance of the networks.

For each of the five network configurations, ten different networks with different random weight configurations were trained to ensure that the results were not due to a specific set of weights. The averages across the ten networks are presented in the next section.

4.3. Results and discussion

For the ten networks at each configuration, the SSE from the last 50 training epochs was averaged together to minimize random noise in the SSE measure. Bonferroni corrected t tests were used to compare the performance of the 'what' and 'where' tasks to control for multiple comparisons (one comparison for each of the five network configurations). The Bonferroni correction for five tests results in a critical probability of 0.01, instead of the standard 0.05 level.

Fig. 4 depicts the networks' performance. Each graph presents the performance averaged across ten different weight initializations for each of the hidden unit configurations. As is evident in the graph, by the end of training the 8/6 and 9/5 configurations resulted in significantly larger 'what' SSE compared to 'where' SSE, $t(9) = 4.13$, $P < 0.005$ for the 8/6 configuration and $t(9) = 5.69$, $P < 0.0005$ for the 9/5 configuration. These results indicate that the 'what' pathway may not have been allocated sufficient hidden units. There was no difference between the 'what' and 'where' SSE for the 10/4 and 11/3 networks, $t(9) = 1.46$, $P > 0.10$ for the 10/4 configuration and $t(9) = 2.73$, $P > 0.02$ for the 11/3 configuration. (Note that there was a trend toward a significant difference in the 11/3 configuration; this comparison did not reach statistical significance because the Bonferroni correction required a P value of 0.01 in order to claim statistical significance.) Finally, the 12/2 configuration showed significantly larger 'where' error than 'what' error, $t(9) = 29.59$, $P < 0.0001$. Two 'where' hidden units were insufficient to perform the localization task, so the 12/2 networks could not reduce the 'where' SSE to fully trained levels in which the SSE was near zero.

These results replicate those of Rueckl et al. [40] in that the 'what' and 'where' error depends on the hidden unit configuration of the model. Specifically, in the present simulations either 8 or 9 'what' hidden units appear to be insufficient for the network to perform the 'what' identification task. Although the networks reduced the error associated with this task, the 'what' error remained significantly larger than the 'where' error. The same interpretation holds for the 'where' hidden units, in which 2 'where' hidden units are insufficient to perform this task. Networks with only 2 hidden units devoted to the 'where' task have significantly larger 'where' error compared to 'what' error.

A striking pattern that emerges from Fig. 4, and from Rueckl et al.'s [40] results, is that the 'where' task appears to be learned much more rapidly than the 'what' task. This pattern emerges even for network configurations that do not permit low asymptotic learning for the 'where' task, such as the 12/2 configuration in the present simulations. I investigated this early training advantage for the 'where' task further because this training difference between the two tasks may provide a mechanism to explain the dissociation between 'what' and 'where' tasks in visual form agnosia.

To examine the apparent early training advantage for the 'where' task, the average SSE was computed for training epochs 1–40 across all network configurations. The first 40 epochs were chosen because the majority of the learning (i.e. reduction in SSE) occurs within these epochs. Across the first 40 epochs of training, the 'where' error is was significantly smaller than the 'what' error, $t(9) = 16.31$, $P < 0.0001$, with an average SSE of 20.69 for the 'what' task and an average SSE of 9.21 for the 'where' task. This indicates that the 'where' pathway in the model initially learned the localization task more rapidly than the 'what' pathway learned the identification task. Because three of the network configurations led to less-than-optimal performance for the two tasks (the 8/6, 9/5, and 12/2 configurations), the same analysis was performed separately for the 10/4 and 11/3 configurations because these configurations resulted in comparable performance between the 'what' and 'where' tasks at the end of training. For the 10/4 networks, the 'where' error was significantly smaller than the 'what' error across the first 40 training epochs, $t(9) = 29.16$, $P < 0.0001$, with an average SSE of 21.01 for the 'what' task and an average SSE of 7.03 for the 'where' task. Similarly, for the 11/3 networks, the 'where' error was significantly smaller than the 'what' error in the first 40 training epochs, $t(9) = 13.36$, $P < 0.0001$, with an average SSE error of 20.14 for the 'what' task and an average SSE of 10.1 for the 'where' task.

The conclusion that the 'where' task was learned more rapidly than the 'what' task was corroborated by an additional analysis of the 10/4 and 11/3 networks. Across the first 20 training epochs, a simple regression was calculated between the epoch number and the SSE for the two tasks. The first 20 epochs were chosen because the reduction in SSE was approximately linear

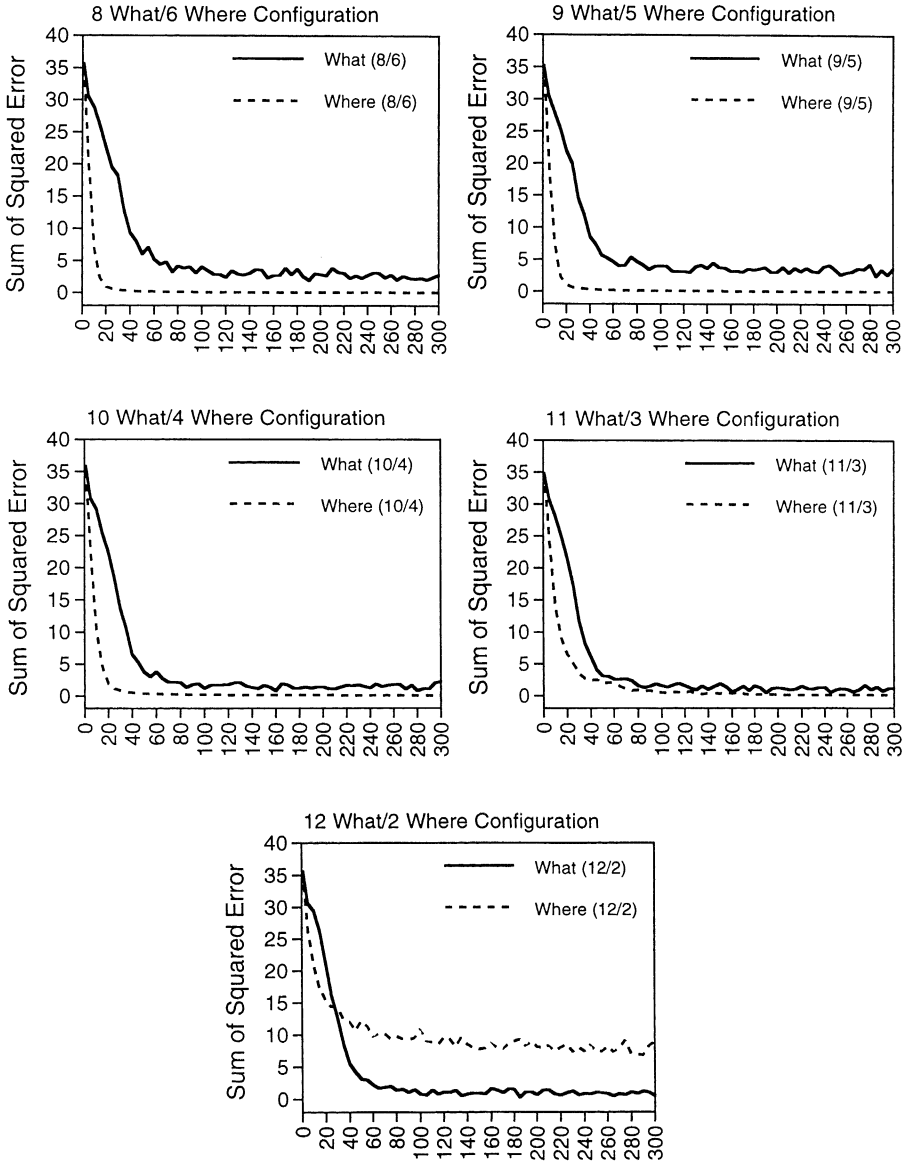

Fig. 4. Learning functions for the five different network configurations tested in Simulation 1. Each panel presents the SSE averaged across ten networks with different initial weights.

within this range, thus enabling description of these data with a simple linear regression. The slope of the regression equation measures the learning rate for each of the two tasks, with larger negative slopes indicating faster learning than smaller negative slopes. A separate regression was computed for each of the ten networks in both the 10/4 and 11/3 configurations. The slopes of the ten regression equations were compared between the 'what' and 'where' tasks.

For the 10/4 network configuration, the slope from the regression equation was significantly larger for the 'where' task than for the 'what' task as indicated by a Wilcoxon signed rank test, $P < 0.01$ (slope of -0.59 for the 'what' task and -1.72 for the 'where' task).[2] The 11/3 network configuration exhibited similar performance, with significantly larger slopes associated with the 'where' task than with the 'what' task, $P < 0.01$ (slope of -0.622 for the 'what' task and -1.487

[2] The Wilcoxon signed rank test, a nonparametric statistic, was used for this comparison because the slopes of regression equations may not be normally distributed. Indeed, inspection of the ten slopes from the 10/4 and 11/3 configurations indicated that the slopes were approximately uniformly distributed, not normally distributed.

for the 'where' task). These results indicate that the initial learning of the 'where' task was faster than the initial learning of the 'what' task.

The differences between the 'what' and 'where' tasks in early training can be understood in terms of how each task influences the learning algorithm and the update of the weights and the SSE term. The 'where' task is, apparently, an easier task computationally than the 'what' task because 'where' task is linearly separable and the 'what' task is nonlinearly separable [24]. Linearly separable problems have mathematically simpler input–output contingencies than do nonlinearly separable problems. Task difficulty provides the basis for the learning rate differences between the two pathways. Because the learning algorithm seeks to minimize SSE across training epochs, weight changes that produce larger reductions in error will be made, which allows early weight changes to reduce the 'where' error more rapidly than the 'what' error.

The results of Simulation 1 are important for two reasons. First, replicating the findings of Rueckl et al. [40] is important to ensure that the 'what' and 'where' pathways have sufficient numbers of hidden units to perform the identification and localization tasks. These tasks must be performed accurately before lesions of the model can be tested in Simulation 2. Any differences between 'what' and 'where' performance could be due to the choice of an idiosyncratic network configuration. Demonstrating that 'what' and 'where' can be learned to the same degree after 300 training epochs in an optimal configuration, such as the 10/4 configuration, reduces the role of network configuration as a trivial explanation of the lesion results. Second, Simulation 1 demonstrated that there is an initial training advantage for the 'where' task over the 'what' task. This result is important because this early 'where' advantage may provide the basis of dissociating identification and localization tasks in visual form agnosia. Importantly, the initial training advantage for the 'where' task is observed across all network configurations, including those that allow both 'what' and 'where' tasks to be learned to similar levels. Because an optimal network configuration can be chosen based on the results of Simulation 1, an informed lesion analysis can be performed.

In Simulation 2, a lesion analysis was performed with the 10/4 network configuration. This configuration was chosen because the ten Simulation 1 networks trained with this configuration showed no difference between the 'what' and 'where' error after 300 training epochs and, therefore, had no biases toward either the 'what' or 'where' task. The critical insight with this split-pathway network is that under noisy (i.e. damaged) situations, the 'where' task performance may remain more stable than the 'what' task performance based on the early-learning differences along the two pathways in the model. More stable 'where' performance than 'what' performance would mirror the results from visual form agnosics, thereby providing a dissociation between 'what and 'where' tasks under damaged conditions. This hypothesized dissociation was examined in Simulation 2, in which a model was 'lesioned' by adding noise to the activations of the units in the input layer. This lesion has the potential to influence processing along both pathways equally because the lesion occurs to an input common to both pathways. However, if the performance of the 'where' task is more robust than the 'what' task, as suggested by the initial learning differences between the two pathways, then the two pathways may be differentially influenced by an early-level lesion.

5. Simulation 2: lesioned performance

The goal of Simulation 2 was to determine if 'what' and 'where' performance could be dissociated following a lesion to a processing layer that lies before the separation of the two pathways. The ultimate goal, of course, is to try to understand the puzzling phenomenon in visual form agnosia whereby these patients are profoundly impaired at object recognition but perform spatial tasks, including visuomotor tasks, reasonably well.

How would a lesioned model look like a patient with visual form agnosia? The present simulations are not intended to mimic an individual patient's performance from specific tasks; instead, these simulations investigate the relative performance differences between the 'what' and 'where' tasks that are typical of the aggregate population of visual agnosics. To effectively simulate visual form agnosia, the model needs to show that 'what' tasks are performed more poorly than 'where' tasks when the model is lesioned in a manner that simultaneously influences both the 'what' and 'where' pathways.

As noted previously, the 10/4 network configuration was lesioned by adding noise to the activations of the units in the input layer. This type of lesion was chosen because it influences the two pathways in a similar manner; that is, both pathways receive degraded, noisy input. Several methods exist for lesioning neural network models (see Refs. [16,22,32,33,35,36]), such as adding noise to units' activations, adding noise to weights, removing weights, and removing units. Performance across different lesion types (e.g. weight removal vs adding noise to weights) results in similar impairments (see Ref. [22]). In the present simulations, however, some lesion methods could produce an artificial advantage favoring the 'where' pathway over the 'what' pathway. For example, removing input units would impact the 'what' pathway more than the 'where' path-

way result because each input unit is connected to 10 'what' hidden units but only to 4 'where' hidden units. Thus, removing an input unit would remove more weights one the 'what' pathway than on the 'where' pathway.

5.1. Lesion methodology

The random noise added to the input layer was generated with a Gaussian distribution having a mean of 0 and a variable standard deviation. On average, no noise was added to an input unit's activation; however, the standard deviation of the Gaussian influenced the range of the noise added, with distributions with larger standard deviations having the potential to add more noise than distributions with smaller standard deviations. There were four levels of lesions in Simulation 2. The noise added to the input units' activations had a standard deviation of 0.1, 0.2, 0.35, or 0.5. Larger standard deviations add more variability to units' activation and, therefore, reduce the reliability of the information that a unit represents. Adding noise with a larger standard deviation acts as a larger lesion of the network.

A single trained network was tested with all lesion levels. This network had a 10/4 hidden unit configuration and learned the 'what' and 'where' tasks to similar levels (the difference between performance on these two tasks did not differ significantly). Ten different random noise lesions were generated at each of the four levels of noise. For each input pattern, a different random noise lesion was generated. The network was tested on 360 random noise lesions for each of the four amounts of noise. Testing involved presenting the network with a noisy input pattern and comparing the networks' output to the correct output. The SSE was computed, and the results reported are the average SSE across the 360 different random lesions for the four levels of noise.

5.2. Results and discussion

The SSE across each random noise lesion was computed separately for the two tasks. The average lesioned performance appears in Fig. 5. The best-fitting line for the 'what' and 'where' tasks is also shown in Fig. 5. There are three patterns apparent in these data. First, as the lesion increases (i.e. as more noise is added to the input level), the SSE increases. Second, the 'what' error is greater than the 'where' error for all but the smallest lesion level. Third, the 'what' error increases more rapidly than does the 'where' error.

The three patterns that appear in Fig. 5 were examined by analyzing the SSE with a two-factor analysis of variance (ANOVA), with lesion amount (0.1, 0.2, 0.35, and 0.5) and task ('what' vs 'where') as factors. The increase in SSE differed among the four lesion amounts, supported by a main effect for lesion amount, $F(3, 36) = 327.0$, $P < 0.0001$. The SSE was greater for the 'what' task than for the 'where' task (SSE of 15.66 and 7.37, respectively), $F(1, 36) = 350.6$, $P < 0.0001$. Finally, lesions increased SSE more rapidly for the 'what' task than for the 'where' task as indicated by the interaction between lesion amount and task, $F(3, 36) = 55.67$, $P < 0.0001$.

These lesion results indicate that 'what' and 'where' are influenced unequally by a lesion to an earlier common processing level. Specifically, Gaussian noise lesions proved more detrimental to the identification task than to the localization task. Not only did the 'what' task show more error than the 'where' task for all lesion sizes, but as the lesion size increased, the 'what' error increased more rapidly than did the 'where' error.

Why does the model perform 'what' more poorly than 'where' when lesioned? An analysis of the signal-to-noise ratio of the 'what' and 'where' tasks may provide an answer. When lesioned, the signal of an individual pattern (i.e. 'what' the pattern is and where it is located) becomes degraded and less reliable. But, the location of an object can be more readily discerned in the presence of noise than can the identity of the object. Thus, if the 'what' and 'where' pathways are faced with input that contains the same amount of noise, the 'what' task will be hindered more than the 'where' task.

The noise tolerance difference between the identification and localization tasks also explains the faster initial learning of the 'where' task compared to the 'what' task. At the outset of training, the weight values are

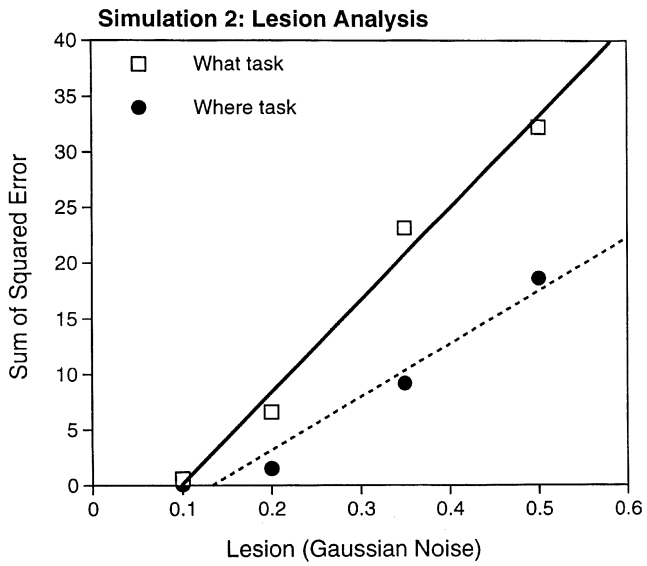

Fig. 5. Results from Simulation 2. This graph shows a lesioned network's performance following the addition of Gaussian noise to the units in the Input layer. The addition of noise causes the network to produce more errors, as indicated by the increase in SSE. Most important, the addition of noise causes more errors for the 'what' task than for the 'where' task.

random, and these random weight values produce activation patterns that are random. As training progresses, however, the weights begin to change to reduce the SSE term. As the weights change, the patterns of activation across both pathways still contain noise. Because the spatial location of an object is preserved more under noise than is the identity of that object, the error produced by the 'where' task will be smaller than that produced by the 'what' task. Thus, early weight changes are more likely to favor the 'what' pathway, allowing the localization task to be learned more rapidly than the identification task. This interpretation explains the 'what' and 'where' differences by appealing to the input stimuli on which the network is trained. No reference need be made to the specific learning algorithm or network configuration. Indeed, in Simulation 1 the same training pattern was observed for all network configurations, including the 12/2 configuration, in which the 'where' task was performed more poorly overall than the 'what' task. 'Where' is an easier task computationally than is 'what' because localizing an object requires less information than identifying that object. As a consequence, when a system is lesioned, 'where' performance is better than 'what' performance.

The results from Simulation 2 are consistent with the pathway information account of visual form agnosia. Although the present simulations are simple in nature, the lesioned model's performance bears a striking resemblance to the behavior of patients with visual form agnosia. Like patients JW and DF discussed in Section 1, the network shows poorer identification performance than visuomotor/visuospatial performance following an early-level lesion.

Although Simulation 2 supports the pathway information account, additional evidence is required to argue convincingly that 'what' and 'where' task differences may lie at the heart of the dissociation observed in visual form agnosia. One prediction of the pathway information account that stems from Simulations 1 and 2 is that equating the difficulty of 'what' and 'where' tasks should abolish the dissociation between 'what' and 'where' performance following an early-level lesion. I test this prediction in Simulation 3 by reducing the difficulty of the 'what' task. Specifically, in Simulation 3 the network is required to learn the identification of fewer patterns than in the previous simulations but the same number of locations. Here, the network was trained to classify three shapes instead of the six shapes. These objects could appear in six locations, as in the previous simulations. The number of hidden units to perform the 'what' and 'where' tasks was now identical, providing a degree of equivalence. The pathway information account makes two predictions in Simulation 3: (1) that learning the 'what' and 'where' tasks should be roughly equivalent; and (2) that following an early-level lesion the impairment of 'what' and 'where' performance should be roughly equivalent. These predictions can be understood in terms of linear separability. When a large corpus of objects is used to train the model, as in Simulation 2, the 'what' task is nonlinearly separable and, consequently, computationally more difficult than the 'where' task. Reducing the number of objects that the network must learn allows this task to become linearly separable, because the objects now have less overlap. The objects in the reduced corpus are more orthogonal to one another than the objects in the full corpus used in Simulation 2. The reduced overlap among the shapes should allow the 'what' task to be learned more quickly with a fewer number of hidden units. The critical issue is whether the reduced overlap in the 'what' task also will allow the model to perform this task in the face of damage.

6. Simulation 3: equating 'what' and 'where' performance

6.1. Simulation methodology

The training and testing procedures were identical to those described for Simulation 2, with two exceptions. First, the number of patterns the network learned to identify was reduced from six to three. The networks in Simulation 3 learned to recognize patterns two, three, and four are shown in Fig. 3. Second, the number of 'what' hidden units was reduced to 4, a number that allowed the 'what' pathway to recognize the three patterns and that allowed the hidden units to be equated between the 'what' and 'where' pathways. The training performance of this 4/4 network configuration was tested across ten different networks, with each network beginning training with a different random weight configuration.

The lesion methodology was similar to that used in Simulation 2. However, instead of testing a range of lesion magnitudes, only the largest lesion was tested (0.5 Gaussian noise added to the activation values of the input patterns) because the post-lesion performance in Simulation 2 was largest for this lesion.

6.2. Results and discussion

6.2.1. Training performance

The training performance of the ten networks is shown in Fig. 6. To determine if equating the difficulty of the 'what' and 'where' tasks allowed the unlesioned network to learn the two tasks at similar rates, I computed the two measures of learning rate used in Simulation 1. For each of the ten networks, I calculated the average SSE across the first 40 training epochs for each task. Across the initial learning, these networks showed larger error terms for the 'where' task than for

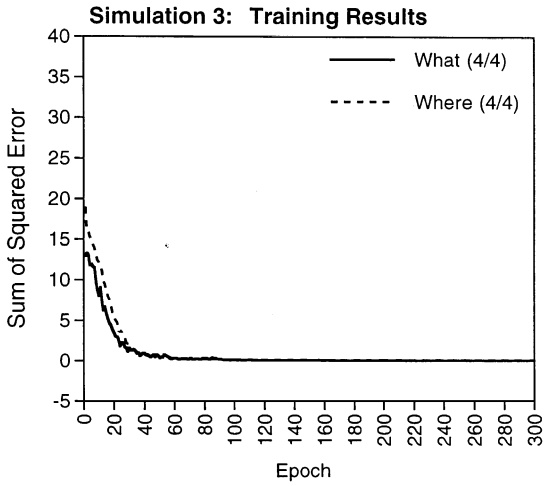

Fig. 6. Average training performance from ten networks tested in Simulation 3. These networks learned 'what' and 'where' tasks that required the same numbers of hidden units along each pathway. Equating the two tasks reduces the learning differences between the 'what' and 'where' tasks.

the 'what' task (5.02 'what' error and 6.89 'where' error). This difference was statistically significant, $t(9) = 11.41$, $P < 0.0001$, and it demonstrates a reversal of the training data from the 10/4 network configuration (see Simulation 1).

Second, to characterize further the learning in the 4/4 network configuration, I computed, as in Simulation 1, a simple linear regression across the first 20 training epochs for each task. Recall that the slope of the regression equations measures the learning rate for each of the two tasks, with larger negative slopes indicating faster learning than smaller negative slopes. The slopes of the ten regression equations were compared between the 'what' and 'where' tasks. The average slope for the 'what' task was -0.56; for the 'where' task the average slope was -0.64. The difference in average slopes was significant between the two tasks as indicated by a Wilcoxon signed rank test, $P < 0.03$. Thus, although there was less error associated with the 'what' task than with the 'where' task during the initial learning, the 'where' task was learned slightly faster than the 'what' task. In summary, the training results are equivocal: On the one hand, one measure of learning, the SSE across the first 40 epochs, suggests that the 'what' task is easier than the 'where' task. On the other hand, the other measure of learning, the slope of the learning curve, suggests that the 'where' task is learned more rapidly than the 'what' task. Thus, unlike in Simulation 1, there is little evidence from Simulation 3 that indicates a distinct learning advantage for the 'where' task over the 'what' task. The two tasks appear to be approximately equivalent to one another in terms of the networks' performance.

6.2.2. Lesioned performance

As in Simulation 2, the 'what' and 'where' SSE terms were computed across ten noise lesions for each of the input patterns in one network. There was a small, nonsignificant difference between the error terms, with a marginally larger error associated with the 'what' task than with the 'where' task (11.96 vs 9.97, respectively), $t(9) = 2.05$, $P > 0.05$. In addition, this difference in the 4/4 network configuration was significantly smaller than the difference observed in Simulation 2 (10/4 configuration) for the same magnitude of lesion, $t(18) = 9.42$, $P < 0.0001$. This latter result indicates that reducing the difficulty of the 'what' task relative to the 'where' task produces a corresponding reduction in the network's post-lesion performance. Stated simply, when the overall difficulty of the 'what' and 'where' tasks is similar, the post-lesion performance of the two tasks is similar. Thus, the relative difficulty of the two tasks corresponds directly to the magnitude of the dissociation in performance following a lesion to a shared input, consistent with a task difficulty account of the 'what–where' dissociation observed in visual form agnosia.

Qualitatively similar results were obtained by lesioning other 4/4 network configurations following training. Some of these networks demonstrated a significant difference between 'what' and 'where' performance following a lesion with larger error on the 'what' task than on the 'where' task. However, the difference between the two tasks was always smaller in the 4/4 configuration than in the 10/4 configuration, again indicating that reducing the relative difficulty of the two tasks reduces the magnitude of the dissociation following a lesion to the shared input to the two pathways.

Overall, the results of Simulation 3 are consistent with the pathway information account of the 'what–how' dissociation in visual form agnosia. When the 'what' and 'where' tasks are equated with one another in the split-pathway model, both tasks are learned at similar rates. By reducing the number of patterns that the network is required to identify, 4 'what' hidden units can be used to perform the 'what' task. Four hidden units also are required along the where pathway in order to correctly distinguish six different locations.

More important, following a lesion to the inputs to the 'what' and 'where' pathways, networks that have learned comparable 'what' and 'where' tasks show little, if any, difference in performance on the two tasks. The early-level lesion disrupts the network's performance compared to an unlesioned network, but the disruption is similar for the 'what' and 'where' tasks. Thus, visual form agnosic patients may show impaired visual shape perception and relatively preserved visuomotor function because damage to an input to both of these processes has been disrupted. The consequence of this damaged input process is poorer performance for 'what' tasks than for 'where' or 'how' tasks.

The results from the previous three simulations suggest that the transformations performed by the 'what' and 'where' pathways and the linear/nonlinear separability of the two tasks may explain the perception and action dissociation in visual form agnosia. However, one valid criticism of the foregoing simulations is that the 'where' task is very crude and is not representative of the tasks that neuropsychological patients perform. Patients with visual form agnosia not only localize objects, they also calibrate their thumb and forefinger appropriately in picking up objects, indicating that some size processing is preserved (see Ref. [19]). These patients also grasp the positions on objects that allow for the most stable grasp [21].

In the final simulation, I examine whether a dissociation between the 'what' and 'where' tasks can be obtained if the model is trained on a more realistic 'where' task. In Simulation 4, the split-pathway model is given the 'what' task from the previous simulations and a 'where' task that requires both localization and depth information. The model is presented with the standard object-in-a-location input and a parallel depth input. The depth input represents the depth plane of the object (a crude 'near,' 'intermediate,' or 'far' distance). The depth input projects to the Where Output units, as shown in Fig. 7A. The 'where' task is to conjoin the object's location and depth plane. That is, the model must learn about two 'where' components. There are now 18 Where Output units, one for each of the six locations by three distance combinations.

The depth input in Simulation 4 would allow the model to provide a motor system with information that could be used to take the retinal extent of the object into account when computing a grasp aperture. Because the objects' retinal size does not change (all objects are a 3×3 pattern; see Fig. 3), the depth input would allow the 'where' pathway to know if the object was relatively small or relatively large, which could, in turn, guide grip scaling. There is neurophysiological evidence that neurons in parietal lobe areas are depth selective and represent stimuli in terms of egocentric distance (e.g. Refs. [42,43]), which makes a depth input to the 'where' pathway a plausible refinement to the model. If the results of Simulation 4 replicate those of the previous simulations, then the pathway information account may explain the behavior of visual form agnosics who are tested with more complex action tasks.

7. Simulation 4: an extended 'where' task

7.1. Simulation methodology

The training and testing procedures were identical to those described for Simulation 2, with one exception: a Depth Input layer was added; this layer projected directly to the Where Output units. There were three units in the Depth Input layer, and each unit represented a different distance that the object was from the model. As described above, for the 'where' task the model was trained to combine correctly the object's location and its distance. There were 18 Where Output units, one for each location-by-depth conjunction. The 'what' task was to identify the object irrespective of its position or distance. There were 6 What Output units, as in Simulations 1 and 2.

The lesion methodology was similar to that used in Simulation 2. Importantly, both the standard object-in-a-location input (Fig. 3) and the depth input were lesioned by adding random Gaussian noise to the activation values of the input patterns. The Depth Input layer was lesioned because to prevent giving the 'where'

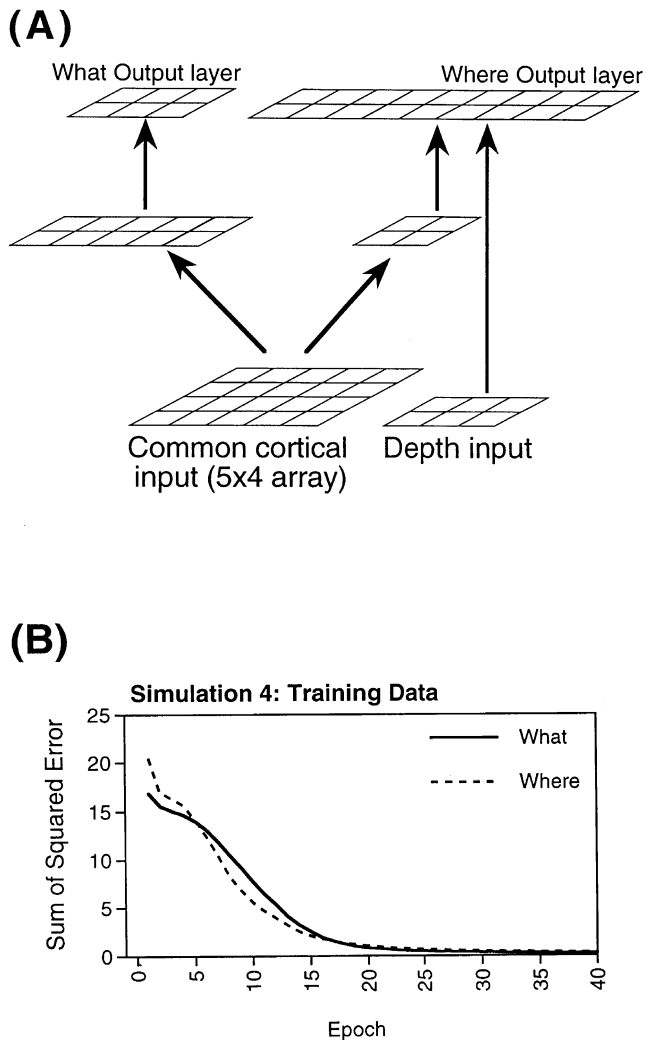

Fig. 7. (A) Network architecture; and (B) training results from Simulation 4. Training results are averaged across ten networks each having a different random weight configuration before training. The training results are shown only for the first 40 epochs to highlight the differences between the 'what' and 'where' tasks. The networks' performance reached asymptote after the first 40 epochs shown.

pathway an unfair advantage: If the Depth Input layer went unlesioned, the 'where' pathway would receive some intact inputs. The lesion amounts were those used in Simulation 2.

7.2. Results and discussion

7.2.1. Training performance

The training performance of the ten networks is shown in Fig. 7B. To assess the learning of the two tasks, I computed the two measures of learning rate used in Simulation 1. For each of the ten networks, I calculated the average SSE across the first 40 training epochs for each task. Across the initial learning, these networks showed no difference in the error terms for the 'where' task and the 'what' task (4.09 'what' error and 3.99 'where' error), $t(9) < 1$, n.s. However, the learning rate, measured by the slope of a linear regression across the first 20 training epochs, did differ between the two tasks. The networks learned the 'where' task faster than the 'what' task. The average slope for the 'what' task was -0.95; for the 'where' task the average slope was -1.03. This difference between the slopes was significant as indicated by a Wilcoxon signed rank test, $P < 0.01$. The added complexity of the 'where' task minimally altered the networks' learning. The 'where' task is learned faster than the 'what' task, although the two tasks' error terms converge rapidly within the first 40 training epochs. The two tasks are learned equally well by the end of 300 training epochs, as the 'what' and 'where' error terms do not differ in the final 50 epochs of training, $t(9) < 1$, n.s.

7.2.2. Lesioned performance

The 'what' and 'where' SSE terms were computed across ten noise lesions for each of the four lesion levels used in Simulation 2. As is evident in Fig. 8, the 'what' task is affected more by the lesions than the 'where' task, despite the 'where' task requiring a more complex output. The data depicted in Fig. 8 were analyzed with a two-factor ANOVA, with lesion amount and task as factors. There was a main effect for lesion amount, $F(3, 36) = 940.43$, $P < 0.0001$, indicating that the increases in SSE differed among the four lesion amounts. There also was a main effect for task, $F(1, 36) = 1429.56$, $P < 0.0001$, with greater error in the 'what' task than in the 'where' task (SSE of 9.08 and 2.47 for 'what' and 'where,' respectively). Finally, there was a statistically significant interaction between lesion amount and task, $F(3, 36) = 231.38$, $P < 0.0001$. This significant interaction indicates that the 'what' error showed larger increases than did the 'where' error across these four lesion levels.

The results of Simulation 4 replicate the results of Simulation 2, but use a 'where' task that more closely approximates the actions that patients with visual form agnosia are asked to perform. When lesioned, the model was better able to localize and perceive the distance and size of objects than to recognize the objects' identities. As suggested by the previous simulations, 'where' tasks, even those that require a conjunction of location and depth, may be computationally easier (i.e. more linearly separable) than 'what' tasks, and thus may provide an explanation of the dissociation between perception and action in visual form agnosia. Of course, the dorsal pathway is involved in processing many different types of visual and motor information (e.g. orientation; see Ref. [43]), and future simulations could focus on exploring increasingly complex 'where' tasks.

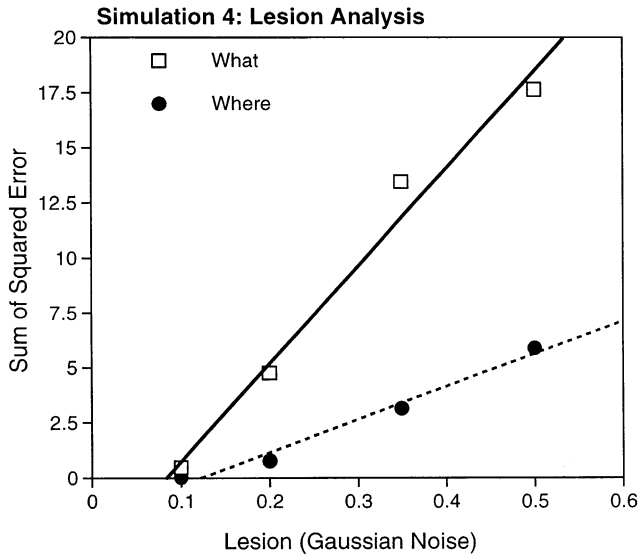

Fig. 8. Lesion results from Simulation 4. The addition of noise disrupts performance more for the 'what' task than for the 'where' task, despite the 'where' task learning a conjunction of spatial location and retinal disparity inputs.

8. General discussion

Patients with visual form agnosia show an interesting neuropsychological dissociation: they can act upon and localize visually presented shapes without being able to visually perceive those shapes. Several theoretical accounts have been proposed to explain this 'what–how' dissociation in these patients. I presented a third explanation of the dissociation, the pathway information account, and tested that explanation with neural network simulations. In four simulations, processing differences between the 'what' and 'where' pathways can allow a single lesion to a common input to disrupt differentially the performance of these two pathways. One does not need to hypothesize separate perception and action modules that are damaged selectively. Instead, a lesion to the common input to the pathways

impairs object identification ('what') more than it impairs spatial localization ('where'). The different performance of the two pathways is caused by the complexities of the 'what' and 'where' tasks and the information required by each pathway. Because 'what' is computationally more challenging than 'where', the what task is learned slower and is learned with higher error rates initially than the 'where' task. When the two tasks are equated with one another, the network no longer produces the dissociation observed in visual form agnosics. These results support the pathway information account of the 'what–how' dissociation in visual form agnosia. There are several issues arising from the pathway information account and the present simulations that warrant discussion.

The most specific issues for discussion pertain to theoretical explanations of visual form agnosia. Although the present simulations potentially explain the 'what–how' dissociation observed in some of these patients, several caveats are in order. As I discussed earlier, my pathway information account is not intended to replace other accounts, such as the accounts discussed by Goodale and coworkers. My pathway information account can be viewed as providing a computational mechanism for the transformation account that is discussed by Goodale and Milner [20,29]. The split-pathway model I have used also relies on the input–output transformations to create two different pathways. My pathway information account provides a computational mechanism that allows perception and action to be dissociated without the need to hypothesize a 'disconnection' between modules or 'spared inputs' to one module. Thus, the pathway information account offers a new perspective on the 'what–how' dissociation, but a perspective that is consistent with the transformation account.

One potential difficulty that was raised by a reviewer was the apparent circularity of the arguments I have made: the 'what' and 'where' are dissociable because the information required for each task is different, and we know the information must be different because the processed underlying two tasks are dissociable. However, I have tried to define 'task difficulty' in more objective, noncircular terms by appealing to the overlap among representations (i.e. linear and nonlinear separability). The 'what' and 'where' tasks differ not because performance on these two tasks is dissociable, but because of the statistical structure of the model's world. In the model's limited world, objects have more overlap with one another, making object identification a nonlinearly separable task. Spatial locations overlap less, making gross localization a linearly separable task. The statistical structure of the world influences the model's internal representations through the training process. The internal representations are influenced differently by damage to a common input, which appears as a dissociation between perception and action.

Another apparent shortcoming of the present simulations suggested by a reviewer is that the results are trivial because the two tasks differ in complexity; it is no surprise that the damaged model performs better on the less-complex task. However, the same reasoning could be applied to the studies with visual form agnosics. These patients may show better performance on visuomotor tasks than on recognition tasks simply because the recognition tasks are computationally more demanding. Object recognition often is measured using seemingly simple same–different matching tasks in which patients report whether two stimuli are physically identical. It is tempting to speculate that recognition is computationally easy because of the apparent simplicity of same–different matching in neurologically normal subjects. However, one cannot assess the ease or difficulty of a task reliably by using intuition. A same–different matching task requires invariant object representations, so that two stimuli that differ in location or size can be matched against one another. Invariant object recognition is a computationally demanding task, which indicates that same–different matching, although apparently easy, may involve difficult computations, such as forming invariant object representations.

In the present simulations, I have not simulated the specific behavior of individual patients. There is a tightrope between generality and specificity in modeling. I have opted for a more general approach, although this approach can be criticized because the details of individual studies are lost. The strength of a general approach, however, is that it can explain a range of behaviors across a range of patients. For example, some patients with visual form agnosia show preserved spatial attention [48] and some preserved perception of spatial relations [34]. The general split-pathway model would suggest that action, spatial relation perception, and some forms of spatial attention may require dorsal-stream processes, which, in the model, can remain relatively intact following damage. Developing a specific model of a particular task or patient might cause one to miss the commonalties among some processes.

An advantage of a general computational approach is that it can make connections with other data. For example, what is the consequence of the neural damage in visual form agnosia? That is, what do these patients 'see', and how do object recognition deficits follow? Some theories of visual form agnosia have stated explicitly that these patients perceive the world as if looking through a 'peppery mask' [8,9]. Presumably this peppery mask acts to add noise to the visual image. On the surface, the noise lesion used in my simulations appears consistent with the peppery-mask hypothesis of visual form agnosia. However, adding masking noise to a visual image does not explain all of the degraded

performance observed in these patients. For example, some visual form agnosics fail to attend to objects (i.e. they have disrupted object-based attention; [48]). Vecera and Gilds [49,50] demonstrated that random masking noise did not abolish object-based visual attention in neurologically normal observers, a result inconsistent with the peppery-mask hypothesis. How can my use of a random noise lesion be reconciled with the evidence against the peppery-mask hypothesis?

One must consider the *locus* of the noise. In the lesioned model, the noise added to the Input layer is not equivalent to random masking noise present in an image. The Input layer in the model is taken as an early cortical level of processing; noise added to this layer of the model therefore corresponds to noise in the early stages of the cortical visual system. In contrast, the masking noise used in tests of the peppery-mask account of visual form agnosia is added to the image viewed by a subject (i.e. noise is superimposed on a visual image); noise added to an image is noise in the external environment. In visual form agnosics, 'noise' would correspond to noisy visual features (e.g. edges) and perceptual groups because the cortical visual areas that correspond to the Input layer in the model represent features and perform perceptual grouping (e.g. Refs. [26,39,52]). Visual form agnosics may see features flickering in and out of view, and they appear unable to organize these features in accordance with gestalt grouping principles. Consistent with the idea of noisy features and groups, Vecera and Gilds [50] simulated visual form agnosia in neurologically normal subjects by removing visual features important for performing perceptual organization. Although adding masking noise to an image may not explain the visual experiences of visual form agnosics, adding feature and grouping noise might.

The present simulations can be linked to other neuropsychological syndromes in two ways. First, there is a *double* dissociation between 'what' and 'how' that must be explained by the network. Second, there are other neuropsychological dissociations that can be explained with different pathway or task information requirements in a neural network model. I discuss these issues in turn.

Visual form agnosics show only a single dissociation between 'what' and 'how' performance. Patients with optic ataxia exhibit the opposite pattern of behavior by demonstrating impaired visuomotor behavior with intact visual shape perception. How do the present simulations explain this double dissociation and the performance of patients with optic ataxia? The split-pathway model explains the performance of optic ataxics by hypothesizing damage that is confined to the dorsal ('where/how') processing stream. Consistent with this prediction, the object ataxic patient tested by Goodale et al. [21] had large bilateral lesions to her occipitoparietal areas. Lesioning the 'where' pathway in the model (e.g. the Where Hidden layer or the Where Output layer) would produce a pattern opposite to the results of Simulation 2 — that is, poor 'where' performance compared to 'what' performance.[3] The presence of a double dissociation between 'what' and 'how' does not appear problematic for the model or for the pathway information account of visual form agnosics.

My present results also are consistent with interpretations of other neuropsychological syndromes. In many neuropsychological syndromes, the presence of a dissociation leads to the inference of two modules that have been disconnected from one another, as in the ventral disconnection account of visual form agnosia. However, my simulation results suggest that disconnections are not always required to explain a neuropsychological dissociation. Simulations of other neuropsychological syndromes have lead to the same conclusion. Consider the case of covert face recognition in prosopagnosia. Patients with prosopagnosia are unable to perform face recognition. These patients show an interesting dissociation between overt (or explicit) face recognition (i.e. assigning the correct name to a face) and covert face recognition (i.e. a 'feeling of knowing' a face). Covert face recognition is measured through indirect tasks or measures, such as performing a simple same versus different discrimination on two faces. Some have explained the overt–covert dissociation in prosopagnosia as the result of a disconnection between an intact face processing module and the conscious awareness of facial identity (see Ref. [16] for a review). However, there is a simpler explanation for the dissociation observed in these patients: Covert recognition tasks may be more sensitive than overt recognition tasks in detecting the residual functioning of a partially damaged face recognition system. As a consequence, when the face recognition system is partially damaged, overt recognition fails but covert recognition is relatively intact. Farah and coworkers [16] addressed this account of prosopagnosia by lesioning a neural network model that named representations of faces. Damage to the face layers in this model produced an overt–covert dissociation. This model did not require a disconnection between intact face processing and conscious awareness of faces to obtain this dissociation.

[3] I have confirmed this result by removing 50% of the connections between the 'where' hidden units and the 'where' output units. Following the lesion, I compared performance on the 'what' and 'where' tasks. A lesion along the 'where' pathway causes greater impairment to the 'where' task than to the 'what' task. Eight lesions to the Where Hidden layer result in an average SSE of 32.37 for the 'where' task and 0.059 for the 'what' task, which differed significantly from one another, $t(7) = 14.91$, $P < 0.0001$. Thus, a lesion confined to the 'where' pathway exhibits performance similar to patients with optic ataxia — poorer 'where-how' performance than 'what' performance.

The pathway information account I have developed for visual form agnosia is similar to the task-sensitivity account that explains the overt–covert dissociation in prosopagnosia. The 'where' pathway, like a covert face recognition task, can tolerate degraded information that arises following a lesion. The 'what' pathway and an overt face recognition task cannot tolerate the degraded information. The result is a difference in 'what–where' performance in visual form agnosics and a difference in overt–covert recognition in prosopagnosics. Thus, the same computational principle (less degradation of performance for some tasks/pathways) can be used to explain dissociations observed in different neuropsychological syndromes.

The present simulation results also may have implications for animal models of extrastriate processing following lesions to primary visual cortex (V1). Single units in the dorsal 'where–how' visual pathway remain responsive following either permanent or reversible damage to V1, but single units in the ventral 'what' pathway are less likely to remain responsive (see Refs. [7,45] for reviews). These neurophysiological results could be due to a preserved pathway from subcortical areas to the dorsal stream, as hypothesized by the spared input account that I discussed in Section 1. However, the pathway information account is consistent with the neurophysiological data: following early-level damage, response properties of ventral stream neurons are disrupted more than dorsal stream neurons because of the computational requirements of the two pathways and the learning that occurs in each pathway. Indeed, the reason that my model fails object identification is because units in the 'what' pathway are less tolerant of noise than units in the 'where' pathway, and these tolerance differences are due to the input–output transformations and receptive field characteristics acquired by the two pathways.

Finally, the pathway information account has implications for understanding the visual environments in which we function. One reason that identification may be more difficult than localization is because there is greater variability within the shape category than within the space category. Also, the shape category is likely larger than the space category — there are more potential objects than there are locations, and new objects can be created. These observations are often used by object recognition theorists to emphasize the difficulty of object recognition (e.g. Ref. [5]). Consequently, object recognition may be a computationally more difficult, nonlinearly separable task and localization a computationally easier, linearly separable task. These intuitive differences between the shape and space categories appear to have important consequences for the representations in the visual system, from the division of processing into two relatively separate pathways to the pattern of behavior following neurological damage.

Acknowledgements

The author wishes to thank Gregg Oden, Maureen Marron, and three anonymous reviewers for comments on this research. This research was supported in part by grants from the National Science Foundation (BCS 99-10727) and the National Institute of Mental Health (MH60636-01).

References

[1] Adler A. Disintegration and restoration of optic recognition in visual agnosia: analysis of a case. Archives of Neurology and Psychiatry 1944;51:243–59.
[2] Alexander MP, Albert ML. The anatomical basis of visual agnosia. In: Kertesz A, editor. Localization in neuropsychology. New York: Academic Press, 1983:393–415.
[3] Bechtel W, Abrahamsen A. Connectionism and the mind: an introduction to parallel processing in networks. Oxford: Basil Blackwell, 1991.
[4] Benson DF, Greenberg JP. Visual form agnosia. Archives of Neurology 1969;20:82–9.
[5] Biederman I. Recognition-by-components: a theory of human image understanding. Psychological Review 1987;94:115–47.
[6] Biederman I. Visual object recognition. In: Kosslyn SM, Osherson DN, editors. Visual cognition: an invitation to cognitive science, 2nd ed. Cambridge, MA: MIT Press, 1995:121–65.
[7] Bullier J, Girard P, Salin P-A. The role of area 17 in the transfer of information to extrastriate visual cortex. In: Peters A, Rockland KS, editors. Cerebral cortex, vol. 10. New York: Plenum Press, 1994:301–30.
[8] Campion J. Apperceptive agnosia: the specification and description of constructs. In: Humphreys GW, Riddoch MJ, editors. Visual object processing: a cognitive neuropsychological approach. Hove, UK: Erlbaum, 1987:197–232.
[9] Campion J, Latto R. Apperceptive agnosia due to carbon monoxide poisoning: an interpretation based on critical band masking from disseminated lesions. Behavioral Brain Research 1985;15:227–40.
[10] Cohen JD, Servan-Schreiber D, McClelland JL. A parallel distributed processing approach to automaticity. American Journal of Psychology 1992;105:239–69.
[11] Cowey A, Stoerig P. The neurobiology of blindsight. Trends in Neurosciences 1991;14:141–5.
[12] Driver J, Mattingly J. Parietal neglect and visual awareness. Nature Neuroscience 1998;1:17–22.
[13] Efron R. What is perception? Boston Studies in the Philosophy of Science 1968;4:137–73.
[14] Farah MJ. Visual agnosia. Cambridge, MA: MIT Press, 1990.
[15] Farah MJ. Neuropsychological inference with an interactive brain: a critique of the 'locality' assumption. Behavioral and Brain Sciences 1994;17:43–61.
[16] Farah MJ, O'Reilly RC, Vecera SP. Dissociated overt and covert recognition as an emergent property of a lesioned neural network. Psychological Review 1993;100:571–88.
[17] Farah MJ, Wallace MA, Vecera SP. 'What' and 'where' in visual attention: evidence from the neglect syndrome. In: Robertson IH, Marshall JC, editors. Unilateral neglect: clinical and experimental studies. Hove, UK: Erlbaum, 1993:123–37.
[18] Fries W. The projection from the lateral geniculate nucleus to the prestriate cortex of the macaque monkey. Proceedings of the Royal Society of London B 1981;213:73–80.

[19] Goodale MA. The cortical organization of visual perception and visuomotor control. In: Kosslyn SM, Osherson DN, editors. Visual cognition: an invitation to cognitive science, 2nd ed. Cambridge, MA: MIT Press, 1995:167–213.
[20] Goodale MA, Milner AD. Separate visual pathways for perception and action. Trends in Neurosciences 1992;15:20–5.
[21] Goodale MA, Meenan JP, Bülthoff HH, Nicolle DA, Murphy KJ, Racicot CI. Separate neural pathways for the visual analysis of object shape in perception and prehension. Current Biology 1994;4:604–10.
[22] Hinton GE, Shallice T. Lesioning an attractor network: investigations of acquired dyslexia. Psychological Review 1991;98:74–95.
[23] Humphreys GW, Riddoch MJ. Interactive attentional systems and unilateral visual neglect. In: Robertson IH, Marshall JC, editors. Unilateral neglect: clinical and experimental studies. Hove, UK: Erlbaum, 1993:139–67.
[24] Jacobs RA, Jordan MI, Barto AG. Task decomposition through competition in a modular connectionist architecture: the what and where vision tasks. Cognitive Science 1991;15:219–50.
[25] Jeannerod M, Arbib MA, Rizzolatti G, Sakata H. Grasping objects: the cortical mechanisms of visuomotor transformation. Trends in Neurosciences 1995;18:314–20.
[26] Lamme VAF. The neurophysiology of figure–ground segregation in primary visual cortex. Journal of Neuroscience 1995;15:1605–15.
[27] Lissauer H. Ein Fall von Seelenblindheit nebst einem Beitrage zur Theorie derselben. Archiv für Psychiatrie und Nervenkrankheiten 1889;21:222–70.
[28] Marotta JJ, Behrmann M, Goodale MA. The removal of binocular cues disrupts the calibration of grasping in patients with visual form agnosia. Experimental Brain Research 1997;116:113–21.
[29] Milner AD, Goodale MA. The visual brain in action. Oxford, UK: Oxford University Press, 1995.
[30] Milner AD, Perrett DI, Johnston RJ, Benson PJ, Jordan TR, Heeley DW, Bettucci D, Mortara F, Mutani R, Terazzi E, Davidson DLW. Perception and action in 'visual form agnosia'. Brain 1991;114:405–28.
[31] Mishkin M, Ungerleider LG, Macko KA. Object vision and spatial vision: two cortical pathways. Trends in Neurosciences 1983;6:414–7.
[32] Mozer MC, Behrmann M. On the interaction of selective attention and lexical knowledge: a connectionist account of neglect dyslexia. Journal of Cognitive Neuroscience 1990;2:96–123.
[33] Mozer MC, Halligan PW, Marshall JC. The end of the line for a brain-damaged model of unilateral neglect. Journal of Cognitive Neuroscience 1997;9:171–90.
[34] Murphy KJ, Carey DP, Goodale MA. The perception of spatial relations in a patient with visual form agnosia. Cognitive Neuropsychology 1998;15:705–22.
[35] Plaut DC, Shallice T. Perseverative and semantic influences on visual object naming errors in optic aphasia: a connectionist account. Journal of Cognitive Neuroscience 1993;5:89–117.
[36] Plaut DC, Shallice T. Deep dyslexia: a case study of connectionist neuropsychology. Cognitive Neuropsychology 1993;10:377–500.
[37] Rafal RD. Neglect. In: Parasuraman R, editor. The attentive brain. Cambridge, MA: MIT Press, 1998:489–525.
[38] Rodman HR, Moore T. Development and plasticity of extrastriate visual cortex in monkeys. In: Rockland KS, Kaas JH, Peters A, editors. Cerebral cortex, vol. 12. New York: Plenum, 1997:639–72.
[39] Roelfsema PR, Lamme VAF, Spekreijse H. Object-based attention in the primary visual cortex of the macaque monkey. Nature 1998;95:376–81.
[40] Rueckl JG, Cave KR, Kosslyn SM. Why are 'what' and 'where' processed by separate cortical visual systems? A computational investigation. Journal of Cognitive Neuroscience 1989;1:171–86.
[41] Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL, editors. Parallel distributed processing: explorations in the microstructure of cognition. In: Foundations, vol. 1. Cambridge, MA: MIT Press, 1986:318–64.
[42] Sakata H, Taira M. Parietal control of hand action. Current Opinion in Neurobiology 1994;4:847–56.
[43] Sakata H, Taira M, Kusunoki M, Murata A, Tanaka Y. The parietal association cortex in depth perception and visual control of hand action. Trends in Neurosciences 1997;20:350–7.
[44] Sorenson KM, Rodman HR. A transient geniculo-extrastriate pathway in macaques? Implications for 'blindsight'. NeuroReport 1999;10:3295–9.
[45] Stoerig P, Cowey A. Blindsight in man and monkey. Brain 1997;120:535–59.
[46] Ungerleider LG, Haxby JV. 'What' and 'where' in the human brain. Current Opinion in Neurobiology 1994;4:157–65.
[47] Ungerleider LG, Mishkin M. Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield RJW, editors. Analysis of visual behavior. Cambridge, MA: MIT Press, 1982:549–85.
[48] Vecera SP, Behrmann M. Spatial attention does not require preattentive grouping. Neuropsychology 1997;11:30–43.
[49] Vecera SP, Gilds KS. What is it like to be a patient with apperceptive agnosia? Consciousness and Cognition 1997;6:237–66.
[50] Vecera SP, Gilds KS. What processing is impaired in apperceptive agnosia? Evidence from normal subjects. Journal of Cognitive Neuroscience 1998;10:568–80.
[51] Warrington EK. Agnosia: the impairment of object recognition. In: Vinken PJ, Bruyn GW, Klawans HL, editors. Handbook of clinical neurology. In: Clinical neuropsychology, vol. 1. Amsterdam: Elsevier, 1985.
[52] Zipser K, Lamme VAF, Schiller PH. Contextual modulation in primary visual cortex. Journal of Neuroscience 1996;16:7376–89.