PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Neural representation of objects in space: a dual coding account

Glyn W. Humphreys

Phil. Trans. R. Soc. Lond. B 1998 **353**, 1341-1351
doi: 10.1098/rstb.1998.0288

The Royal Society

Neural representation of objects in space: a dual coding account

Glyn W. Humphreys

Cognitive Science Research Centre, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

I present evidence on the nature of object coding in the brain and discuss the implications of this coding for models of visual selective attention. Neuropsychological studies of task-based constraints on: (i) visual neglect; and (ii) reading and counting, reveal the existence of parallel forms of spatial representation for objects: within-object representations, where elements are coded as parts of objects, and between-object representations, where elements are coded as independent objects. Aside from these spatial codes for objects, however, the coding of visual space is limited. We are extremely poor at remembering small spatial displacements across eye movements, indicating (at best) impoverished coding of spatial position *per se*. Also, effects of element separation on spatial extinction can be eliminated by filling the space with an occluding object, indicating that spatial effects on visual selection are moderated by object coding. Overall, there are separate limits on visual processing reflecting: (i) the competition to code parts within objects; (ii) the small number of independent objects that can be coded in parallel; and (iii) task-based selection of whether within- or between-object codes determine behaviour. Between-object coding may be linked to the dorsal visual system while parallel coding of parts within objects takes place in the ventral system, although there may additionally be some dorsal involvement either when attention must be shifted within objects or when explicit spatial coding of parts is necessary for object identification.

Keywords: within-object coding; between-object coding; visual neglect; extinction

1. INTRODUCTION

Over the past 15 years or so, one of the most influential distinctions in the field of visual information processing has been that between 'object' and 'spatial' coding. This dichotomy, first introduced by Ungerleider & Mishkin (1982), holds that the neural areas supporting object recognition are separate from those supporting location coding, as are the computational processes involved in object and spatial coding in the brain. Indeed, to recognize the same object across different viewing positions and angles, it can be argued that the varying spatial information needs to be discarded so that recognition is based on the invariant 'object' information present (see Marr (1982), for one example). Nevertheless, to direct action to objects or to remember their location for future reference, spatial information needs to be encoded and maintained. The idea that there is parallel representation of objects and space meets with the independent requirements of object recognition and action (see also Milner & Goodale 1995).

The distinction, between object and spatial coding, has influenced work varying from computational modelling of pattern recognition (see, for example, Rueckl *et al.* 1989; Jacobs & Jordan 1992) to theories of selective visual attention (for reviews, see, Humphreys & Bruce 1989; Styles 1997). Given the numerous stimuli that may be available in any scene, forms of selective attention are necessary both to ensure that recognition is successful (e.g. so that the parts of different objects are not linked together) and that action is directed to objects in a coherent way. Space-based theories propose that stimuli are selected for recognition and for action by attention being directed to their spatial locations (see, for example, Posner 1980; Eriksen & Yeh 1985; Treisman 1988). In contrast, object-based theories hold that the elements of objects are selected together, even when they overlap spatially with elements of other objects and when the elements of the object are no closer to one another than they are to the elements of other objects (Duncan 1984; Baylis & Driver 1993). These accounts are not mutually exclusive, however, and in some proposals spatial and object-based selection may be coupled by interactions between object- and space-based systems (see, for example, Farah *et al.* 1993; Humphreys & Riddoch 1993). For example, features coded during early stages of visual processing may be activated by both directed spatial attention and (top-down) by activated object representations, so that spatial attention affects object selection (biasing selection towards objects in the attended locations) and object properties affect spatial selection (so that spatial attention becomes locked onto objects). Such accounts allow for forms of interaction between 'object' and 'spatial' processing streams, so that coherent behaviour results (for similar arguments about coupling between brain areas, see Duncan (1996) and Duncan *et al.* (1997)).

Now, for coupling between the object and spatial systems to be effective, it may be useful for each system also to 'know' something about processing in the other stream: for some form of spatial information to be incorporated into object representations and for forms of object information to be incorporated into our representations of space. Top-down feedback from object to early visual representations may bias spatial selection most efficiently if the parts of objects are spatially coded; for example, without explicit coding of the spatial relations between parts, it may be difficult to facilitate selection for objects where parts are repeated (bodies, chairs, tables and so on). (An example of this can be found in connectionist models of visual recognition, where explicit coding of spatial information provides one way of representing multiple, repeated parts—as in the interactive activation model of reading (McClelland & Rumelhart 1981).) Likewise, all things being equal, it should be more effective to bias spatial attention to occupied rather than empty regions of space.

In this paper, I examine the nature of object and spatial coding in the brain and argue for the existence of at least two forms of visual representation: one in which elements are coded as parts of a single object (a within-object representation) and one in which elements are coded independently (a between-object representation). Both forms of representation can incorporate information about the spatial relations between visual elements, and I argue that the two forms of representation are realized in parallel by the visual system. The within-object representation serves object recognition, and so may be linked to the 'what' pathway, in Ungerleider & Mishkin's (1982) terms. Between-object representations may serve for spatial navigation and action, and so form part of the 'where' pathway suggested by Ungerleider & Mishkin. However, I also show that, aside from these forms of visual representation, the coding of visual space is extremely limited; there seems to be no representation of space devoid of objects. Thus 'where' codings itself involves forms of object representation. I discuss evidence for these arguments, and then review the implications of the suggestions for theories of selective visual attention.

2. THE NATURE OF OBJECT CODING: PARALLEL OBJECT REPRESENTATIONS

(a) *Bottom-up and top-down factors in object coding*

Since the time of the Gestalt psychologists onwards, it has been clear that object descriptions can be derived in a bottom-up manner based on forms of grouping between visual elements (for a summary, see Bruce *et al.* 1996). Furthermore, grouped elements can be selected together for a response. For example, Donnelly and co-workers (1991) showed that response times to detect changes in part-elements were unaffected by the number of parts presented to subjects, provided the parts grouped into a familiar shape (for example, the corners making up a square; see also Baylis & Driver 1994; Humphreys *et al.* 1994*a*, for similar evidence). In contrast, if elements group into separate object descriptions there is serial selection of each description, even if discrimination of the same part-changes are required (Donnelly *et al.* 1991; Baylis & Driver 1998).

In addition, independent (ungrouped) visual elements can be selected together provided that the elements activate a single, stored object representation. For example, word naming times are relatively immune to increases in the numbers of letters present, for words up to about six letters long (see, for example, Frederiksen & Kroll 1976; see figure 3); this suggests that naming can be supported by parallel coding of the visual information within such stimuli. This parallel coding also seems to involve independent representation of the letters present. Thus, words do not necessarily show stronger effects of CaSe MiXiNg than pronounceable non-words, though this should disrupt word codes at a supra-letter level (McClelland 1976; Adams 1979; Mayall & Humphreys 1996). In such cases the visual elements (the letters) may be encoded and selected together not solely on the basis of low-level grouping but by parallel activation of a stored word representation. This parallel activation of a stored representation enables the independent parts to act as a single object description.

I conclude that, whether based on bottom-grouping or on activation of stored representations, parts within objects can be selected in a spatially parallel manner. This evidence for parallel selection of parts within objects indicates that object descriptions can be formed without the application of focal visual attention to each visual element. (Mack and co-workers (1992) have argued that visual elements do not group under conditions of complete inattention, although contrary data have been reported by Moore & Egeth (1997). Rather than straying onto this topic, I confine my discussion to cases where attention is paid to the general region of visual displays but in a distributed manner, and the discussion is focused on whether object descriptions are coded only when focal attention is paid to each display element.)

Some of the most striking evidence for object descriptions being coded without focal attention comes from neuropsychological studies, where patients who fail to attend to areas of space nevertheless show evidence that object coding in those regions has taken place. For instance, in the syndrome of unilateral neglect, patients fail to react to stimuli presented on the side of space contralateral to their lesion (Heilman & Valenstein 1979). Marshall & Halligan (1988) reported that right hemisphere lesioned patients with marked neglect of the left side of stimuli in same–different discrimination tasks revealed implicit left-side processing on preference judgements made to the same stimuli (see also Bisiach & Rusconi 1990). Other investigators (McGlinchey-Berroth *et al.* 1992; Ladavas *et al.* 1993) report semantic priming from objects in the neglected field, suggesting that unattended objects can even be processed to the level of meaning.

Experiments on the neuropsychological phenomenon of visual extinction further reveal the existence of both bottom-up and top-down factors in encoding object descriptions without focal attention. Extinction occurs when a patient is able to detect the presence of a single stimulus in the contralesional field, but fails to detect the same item when it is presented simultaneously with another item in the ipsilesional field (see, for example, Karnath 1988). This appears to be an attentional effect, because detection of the contralesional item fails only when it is placed under conditions of attentional competition with the ipsilesional stimulus. Ward and colleagues (1994) first reported that extinction could be reduced if the elements in the contralesional field grouped with items in the ipsilesional field. They used bracket stimuli, which could group by collinearity between end segments or by symmetry. In addition, they showed that there was recovery from extinction for the elements of horizontal

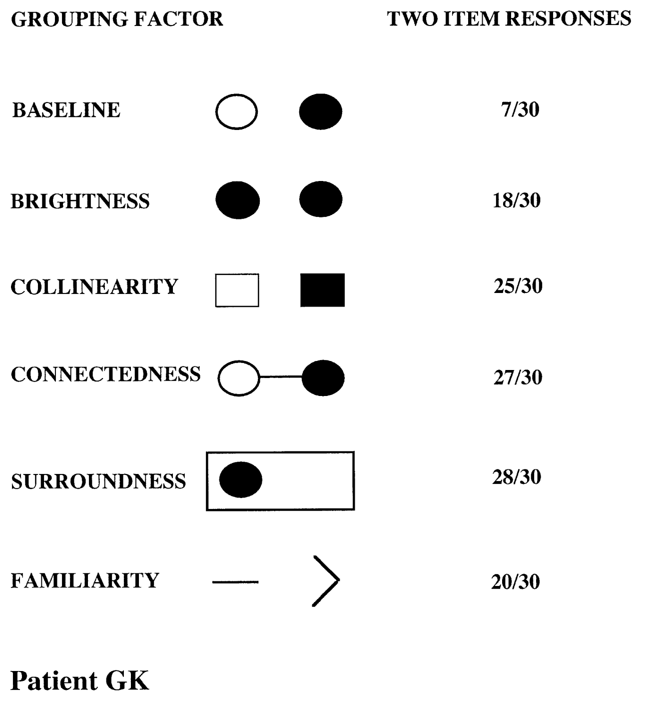

Figure 1. Recovery from extinction. Results on a detection task from two-item displays. The task required discrimination between displays with none, one or two targets. Performance in no- and one-item trials was near ceiling.

arrow stimuli, even when the line and angle components were widely separated in the field and unlikely to group by local Gestalt factors. In this last case, the elements may be recovered together by activating a stored object description. Similar converging data on extinction have been reported by Gilchrist and co-workers (1996) and by Mattingly and co-workers (1997). Gilchrist *et al.* showed that grouping effects were as strong between items presented in the contralesional field as between items in the ipsilesional field, although detection performance overall was better in the ipsilesional field. Thus, although patients may fail to attend to the contralesional field under conditions of attentional competition, grouping can be shown still to operate there.

Recent data from my laboratory on the variety of relations that can lead to recovery from extinction are presented in figure 1. Data here are from the patient, G.K. studied by Gilchrist and co-workers. G.K. suffered bilateral lesions, affecting the right parieto-occipital and parieto-temporal regions and the left tempero-parietal region, resulting in symptoms of Balint's syndrome in which he often fails to perceive more than one object at a time and shows mis-reaching under visual guidance (Balint 1909; for reports, see Humphreys *et al.* 1994*b*; Gilchrist *et al.* 1996). Due to the lesions in his right hemisphere being more severe, G.K. also manifests extinction to left-side targets. In this study, G.K. had to detect whether none, one or two stimuli were presented. Single stimuli appeared 1° of visual angle from fixation in either the left or right visual fields, and on two-stimulus trials, one item was in the left and the other in the right field. Performance in no- and one-item trials was near ceiling and data are presented only for two-item trials. In a baseline condition with minimal grouping, the items were circles of opposite contrast polarity (white and black) which were presented against a grey background. These items lack bottom-up support for grouping based on either collinearity common surface features. Reporting of two items was impaired (see also Gilchrist *et al.* 1996). However, the two-item report improved as cues were added to displays to enable the items in the ipsi- and contralesional fields to group. There was recovery from extinction if the elements had: the same brightness (two white or two black circles), collinear edges (with aligned squares, even though they had opposite contrast polarities, and preventing grouping by common brightness), a connecting line (joining circles with opposite contrast polarities), and inside–outside relations (e.g. a left-field circle appearing within a surrounding rectangle). In addition to this, as reported by Ward and co-workers (1994), performance improved if the elements formed a familiar figure (an arrow), although there were then few bottom-cues to group the elements together. These results indicate that, within the same patient, both bottom-up grouping factors and stored knowledge can be used to recover elements within the contralesional field, even though the patient fails to attend to (and even detect items in) the contralesional field when items do not group (in the baseline condition).

(b) *Parallel representations: a single case*

Neuropsychological evidence not only indicates that object descriptions can be coded without focal attention, but also that separate forms of representation are generated in parallel—in the sense that one representation does not form the input for the other one. These separate forms of representation are revealed by dissociations in which patients seem impaired at using one but not the other form of representation. Patient J.R. was studied by myself and Jane Riddoch (Humphreys & Riddoch 1994, 1995). He had suffered bilateral lesions affecting the left parieto-occipital region and the right fronto-parietal regions. On a first screening he showed an unusual pattern of performance. When asked to read words and nonwords scattered randomly around a page, he made substantial numbers of 'left neglect' errors when identifying each string (typically making letter substitution errors, such as pitch → ditch) but he also made 'right neglect' errors in which responses to strings on the right side of the page were omitted. These different patterns of spatial error did not reflect the positions of elements in the visual field, as they could be demonstrated with the same stimulus depending on the way in which it was coded for the task. We gave J.R. A4 sheets of paper with a large word or non-word written across each sheet. In one set of trials he was asked to read each string as a whole (light → light, nitch → nitch). In another set of trials his task was to read aloud each letter present (light → l, i, g, h, t; nitch → n, i, t, c, h). Marked differences emerged. When asked to read each string as a whole J.R. manifested left neglect (light → night; nitch → pitch). When asked to name each letter, right neglect was apparent. J.R. then read aloud the previously neglected left-side letters but he often omitted right-side letters that previously had been read (light → l, i, g, h, - ; nitch → n, i, t, c, -). The left neglect errors in reading aloud the whole strings tended to occur more with non-words than with words. In contrast, there

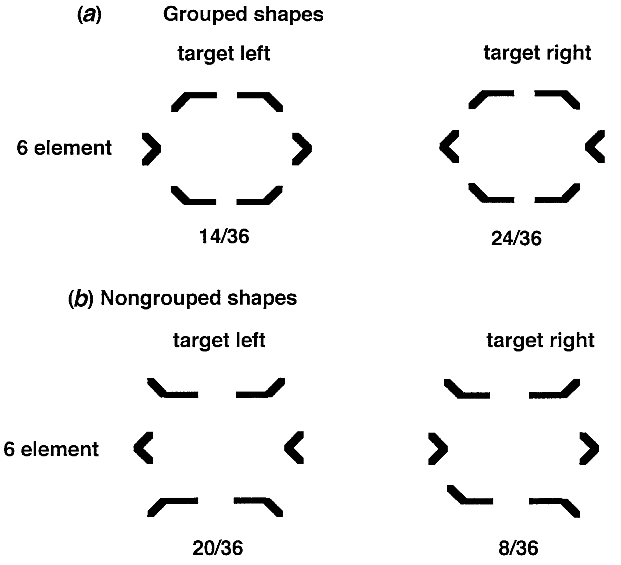

Figure 2. Number of trials on which targets were detected by J.R. as a function of whether the target was on the left or right side of the display. (*a*) Grouped shapes; (*b*) non-grouped shapes. Performance was good when the target was absent.

was no effect of the lexical status of the string on right neglect omissions when reading aloud each letter. The differential effects of lexical status provide converging evidence for the visual information being treated in different ways in the two tasks.

Similar results were also found in studies of shape perception. Donnelly and co-workers (1991) had used displays such as those used in figure 2 with normal subjects, with the task being to detect a target element that faced the opposite way with regards to fixation, relative to the other elements present. Fixation fell at the centre of each shape. In the grouped displays (figure 2*a*, where the elements had collinear ends and formed a closed shape), reaction times (RTs) were unaffected by the number of elements present, consistent with the parts being assimilated in parallel. In the non-grouped displays (when elements were reflected to no longer form a closed figure; figure 2*b*), RTs increased as the number of elements increased. These results provide an operational definition of when object descriptions were formed (in grouped but not in non-grouped displays). When given the same task, J.R. made left neglect errors with grouped displays (missing left-side targets) and right-side errors with non-grouped displays (figure 2). Thus this contrast between left- and right-side neglect was not confined to reading tasks.

We (Humphreys & Riddoch 1994, 1995) proposed that J.R. manifested neglect on either the left or right side of space according to the way in which visual information was represented for the task. For example, when letter strings were read as wholes, the letters were coded as parts of each 'object' (the letter string). When they were named separately, the letters were treated as independent objects. Owing to his bilateral lesions, J.R. may have suffered two forms of neglect. There was left neglect within representations where parts were coded relative to the whole (within-object neglect) and right neglect when elements were coded as independent objects for response (between-object neglect). The evidence demonstrates that neither form of representation serves as input for the other; were that the case, then the earlier form of neglect would have been apparent even if a second form was subsequently added. For instance, consider what should have been the case if elements were first coded as separate objects and only subsequently as parts within objects. Then, when strings were read as wholes, there should have been left neglect on top of earlier right neglect of the end letters (e.g. reading light as nigh-); instead of this, right-end letters were fully represented in left neglect responses (light → night). This supports the argument for parallel coding of the two forms of representation.

In J.R.'s case, the anatomical evidence links within-object neglect to damage to the right hemisphere (producing left-side errors), and between-object neglect to damage to the left hemisphere (producing right-side omissions). It may be that the two forms of representation are selectively localized within each hemisphere and J.R. happened to have different forms of rerepsentation damaged on different sides. Alternatively, each hemisphere may be specialized for attention to a particular form of representation. This argument, for hemispheric specialization in attention to within- or between-object representations, is supported by group studies with patients with unilateral parietal lesions. Egly and co-workers (1994) found that right parietal patients were particularly impaired at switching attention from the ipsi- to the contralesional side of space but showed no extra cost when switches were made between objects on the left and right side. Left-hemisphere damage, however, produced additional costs in shifting attention between objects to detect contralesional targets (see also Buck *et al.* (1998), for converging evidence from single photon omission computed tomography analyses on patients with degenerative posterior atrophy).

(c) *Parallel representations: a dissociative group study*

J.R.'s case provides evidence that within- and between-object codes can dissociate within a single patient. Humphreys & Heinke (1998) demonstrate similar dissociations, but between rather than within patients. They used a dissociative approach in which patients were given sets of tasks in which they either had to attend to parts within objects or in which they had to attend to parts within each of two separate objects. All patients manifested unilateral neglect on simple drawing tasks. In one experimental test of within-object processing, patients received chimeric faces formed by aligning the left and right halves of two faces of either the same or opposite gender. The task was to decide whether each face was male, female or a mixture. In a test of between-object coding, sets of new, whole faces were presented onto piles placed either to the left or right of the patient's body. On each trial either a single new face was presented onto the left or right pile, or two faces were presented onto the left and right piles. The task was to decide if the new face on each trial was male, female or whether there was one male and one female face (on two item trials). With chimeric faces, patients had to discriminate the presence of different sexes within a single object (the face). Previous work suggests that neglect within objects can occur under these conditions, with patients failing to discriminate the left half of the face (Young *et al.* 1992). In the between-object task with separate whole faces, patients could respond correctly even if they attended only to the right side of each face, but they had to attend to each of the two faces present. We (Humphreys & Heinke 1998) found that some patients only neglected the chimerics but showed no neglect in the two-face task (showing neglect only within-objects); other patients demonstrated the opposite pattern (showing no neglect of the chimerics but failing to detect the left- or right-side face when two faces were present). These last patients manifested neglect between separate objects but not within each object (discriminating both sides of chimeric faces). There is thus a double dissociation between the two forms of neglect, consistent with there being two independent forms of representation.

(d) *Parallel representations: reading and counting*

It may be that the dissociation between neglect within and between objects itself reflects the hierarchical coding of visual stimuli for different tasks. For example, tasks of within-object coding, such as whole word reading, may rely on more global visual descriptions than tasks of between-object coding, such as reading the letters within words, where more local descriptions may be involved. Within-object neglect may occur in tasks such as whole word reading because patients are impaired at attending to one side of global representations. Between-object neglect may be linked to spatial biases when local spatial representations are used. This would be consistent with some of the evidence for hemispheric specialization linked to within- and between-object neglect. It has long been known that right hemisphere lesions (associated with within-object neglect) can impair attention to global visual descriptions; likewise, left hemisphere lesions (associated with neglect between objects) may disrupt attention to local visual descriptions (Robertson *et al.* 1988; see Fink *et al.* (1996) for converging evidence using positron emission tomography). (One problem for this argument is that, in one of the patients documented by Humphreys & Heinke (1998), there was left neglect between objects (though right neglect would be expected if left hemisphere damage were crucial). However, this patient had bilateral lesions, making interpretation of the link between the lesion and side of neglect difficult.)

To test further the idea that there are separate within- and between-object representations, not simply differences in hierarchical representation of stimuli, Soren Kyllingsbaek, Andrew Olson and I assessed reading and counting in five patients with lesions affecting either posterior dorsal (parietal) or ventral (occipito-temporal) brain regions. It has long been known that damage to the posterior ventral cortex (particularly in the left hemisphere) can impair visual processing, leading to abnormal effects of the numbers of elements present in the field. Classicially this has been associated with serial letter identification in reading (Dejerine 1892; see papers in Riddoch 1991), but it can also be found in other discrimination tasks with form elements (see Friedman & Alexander 1984; Rapp & Caramazza 1991; Humphreys *et al.* 1992). In contrast to this, the ability of the patients to count visual stimuli has been reported as good (see, for example, Kinsbourne & Warrington 1962)—although it has rarely been tested formally using reaction time (RT) measures (but see Humphreys *et al.* 1985). We reasoned that such patients may have difficulty encoding parts within objects in a parallel manner (for reading), but not in encoding separate objects for counting. The opposite pattern of performance may occur in patients with dorsal (parietal) lesions. Previous studies have shown deficits in visual counting tasks in patients with posterior dorsal lesions (see, for example, Dehaene & Cohen 1994), although the recognition of single objects may be intact (see, for example, Baylis *et al.* 1994). Here one might suspect that the patients are able to encode parts within-objects in a preserved (spatially parallel) manner but the lesion disrupts the ability to assemble or maintain spatial representations between-objects, disrupting counting. A failure specifically in assembling such representations may be assessed by testing a patient's ability to count small numbers of items (less than four), which normal observers seem able to enumerate in parallel (to 'subitize'), without requiring serial scanning or maintenance of previous counted items (see, for example, Trick & Pylyshyn 1993).

One dorsal patient was G.K., who, as noted here, had suffered bilateral parietal lesions. The other two patients, M.B. and M.P., had sustained unilateral right hemisphere damage affecting the parietal lobe. M.B. and M.P. both had mild neglect. Of the two ventral patients, one (D.M.) had unilateral left hemisphere damage affecting the medial occipito-temporal region (see Humphreys *et al.* (1997), for details of the lesion); the other (H.J.A.) had bilateral lesions affecting the occipito-temporal area (Humphreys & Riddoch 1987). Both D.M. and H.J.A. showed abnormally strong effects of string length on reading in clinical tests (a characteristic of letter-by-letter reading). We required patients to either name or count the letters in sets of frequency-matched words having from two to six letters (the conditions were presented in an ABBA design for each patient). In addition, a set of single letters was added, to provide data for string length one. Stimuli were presented on an IBM personal computer and responses triggered a voice key for RT recording. There was a minimum of 12 trials for each string length in each task. Note that the global shape of words does not alter with word length, so counting cannot be based on global shape information; rather it requires the individuation of the letters present.

The data for four patients, M.B. and M.P. (dorsal patients) and D.M. and H.J.A. (ventral patients) are given in figure 3. G.K. was unable to perform the counting task under unconstrained conditions (performing at chance level). Accordingly he was given a forced-choice version of both tasks in which he was given a choice between two stimuli on each trial, for both counting and reading. One alternative was correct, the other was another number either one more or one less than the target number (for the counting task; for example, light: four or five letters?) or a word changed by one letter from the target (for the reading task; for example, light: light or night?).

The results revealed a form of double dissociation between reading and counting in the two sets of patients. The dorsal patients M.B. and M.P. were able to read and showed no effects of string length on naming times, consistent with there being parallel identification of the letters for word recognition. However, they showed effects of

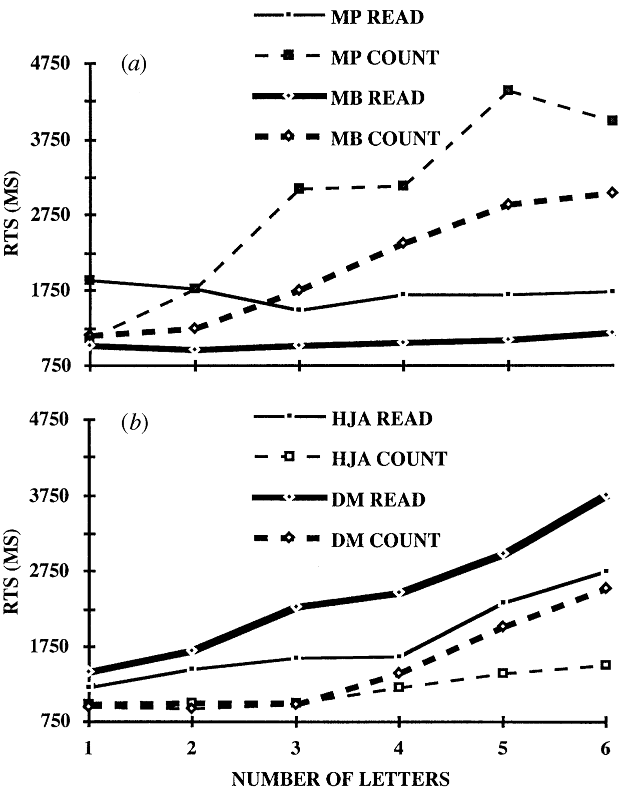

Figure 3. Reading and counting responses by patients with either (*a*) dorsal (parietal) or (*b*) ventral (occipito-temporal) lesions. Very few errors were made.

string length on counting, with counting times increasing directly as a function of the number of letters present. G.K. demonstrated a similar advantage and evidence for parallel word reading, although in his case an accuracy rather than a RT measure was used and reading was contrasted against an inability to count. He scored 32 and 28 out of 36 when asked to identify strings of one to three and four to six letters respectively (overall $\chi^2(1) = 16.53$, $p < 0.01$, relative to chance), but only 22 and 16 out of 36 when asked to count the numbers of letter present (no different from chance (18 out of 36) for string lengths one to three and four to six respectively.

The ventral patients showed the opposite pattern. For strings containing up to three letters there were no increases in counting time relative to the number of elements present, and R.T. increases for strings with more than three letters were relatively slight. The patients were able to enumerate a small number of items in a spatially parallel manner. However, reading times increased with string length, even for strings with few letters (less than three). Thus the ventral patients were impaired at assimilating in parallel the letters needed for word identification, although they were able to respond to information coded in parallel in counting. Interestingly, single-letter identification times were equally fast for the ventral and dorsally lesioned patients so the problem was not simply one of letter identification for the ventral patients. Humphreys *et al.* (1992) have reported converging evidence that parallel assimilation of form information for identification is impaired in patient H.J.A.

These data show for the first time a direct contrast between the effects of dorsal and ventral lesions on

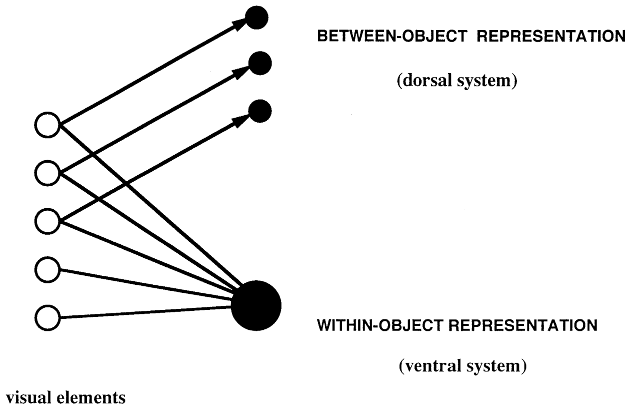

Figure 4. Schematic framework for a dual coding account of visual processing, illustrating independent forms of parallel processing supporting within- and between-object representations.

reading and counting tasks with the same stimuli and where the numbers of items have been varied to assess whether performance depends on parallel spatial coding of visual information. For both sets of patients, the preserved ability relies on spatially parallel coding, as there were no effects of string length on performance (on reading by the dorsal patients, and on counting small numbers of items by the ventral patients). There is, however, a breakdown in parallel coding for the affected task (counting, for the dorsal patients, and reading, for the ventral patients). The effects cannot easily be attributed to a contrast between local and global operations as performance in all cases relies on parallel spatial coding, which might be thought characteristic of 'global' visual processing. Instead the data fit with the proposal that there can be independent forms of visual coding: (i) parallel coding of parts within objects; and (ii) parallel coding of a small number of separate visual objects (between-object coding). These processes are limited by damage to different brain regions. Parallel coding of parts within objects is disrupted by ventral lesions; parallel coding of separate objects is affected by dorsal lesions (figure 4).

If parallel coding of parts within objects takes place within the ventral visual system, however, we are left with explaining why neglect within objects is associated with more dorsal lesions (e.g. in patients such as J.R.). I return to this point in § 3.

3. OTHER FORMS OF VISUAL REPRESENTATION OF SPACE

I have argued that independent visual descriptions are coded in parallel with one another; there is one description in which stimuli are treated as parts within whole objects (within-object coding), and one in which they are treated as separate objects (between-object coding). Apart from these representations, however, is visual space coded in any other way? In particular, is visual space represented in some cartesian manner irrespective of whether objects are present or not? This is a kind of 'blackboard' view of vision, in which objects are depicted on a spatial canvas. On such a view it should be possible to represent the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

canvas even when objects are not present—for example, when we have to remember the location of where an object fell after it has been removed. Also, if object coding involves depiction on some form of internal blackboard, then effects of spatial distance on the blackboard should be primary and not modulated by factors to do with object coding. In fact, neither of these assertions seems to hold, casting doubt on any simple blackboard view. I will go on to argue that any coding of space devoid of objects is, at best, extremely limited.

(a) *Coding empty space*

To assess the question of whether spatial information can be coded even when objects are not present, Luis Fuentes, Derrick Watson, Kevin O'Regan and I evaluated memory for the location of a small target when subjects either maintained fixation on the location or made an eye movement between the presentations of the target. Previous studies have shown that memory for location can improve when visual landmarks are present in the environment (see, for example, Matin 1976), but they have not specified the nature of the visual codes that are employed. Here subjects saw for a variable duration an initial display containing randomly positioned white dots and one (target) green dot. The initial display then disappeared for a short interval (400 ms or longer, to minimize apparent movement) and was replaced by a match display, presented for an unlimited exposure. In the match display the green dot was either kept in the original location or it was shifted a small amount (0.25°). The task was to decide whether the green dot was in the same or a different location. In one condition, subjects were allowed to keep fixating on the target during the interval. In another, a small letter was presented 7° into the periphery during the interval; subjects had to make an eye movement to report this letter and then decide whether the target dot had moved when the match display was subsequently presented. In this condition, subjects had to maintain their memory for the target's location across an eye movement. There were also two variations in the density of the background (white) dots, which were either sparse (20 dots in the field) or dense (40 dots in the field). Data averaged over five subjects are shown in figure 5. The displacement of the target (on 'different location' trials) was set so that subjects scored about 62% correct even when the first display was presented for too brief a time to ensure that subjects could fixate the target dot.

When no eye movement was made during the inter-stimulus interval, memory for location improved as the exposure of the first display increased to about 300 ms and then it asymptoted close to ceiling. This held both for sparse and dense displays. The improvement in performance over the first 300 ms is likely to be owing to subjects taking some time to fixate the target dot in the initial display; but following its fixation, memory for the dot's location can be maintained across an interval and does not depend on the number (or proximity, given that the dots were presented within a limited area) of the background dots. In marked contrast to this are the data when an eye movement was made. For sparse displays there was then no improvement in memory no matter how long subjects were allowed to view the first display for. For dense displays there was an improvement over time.

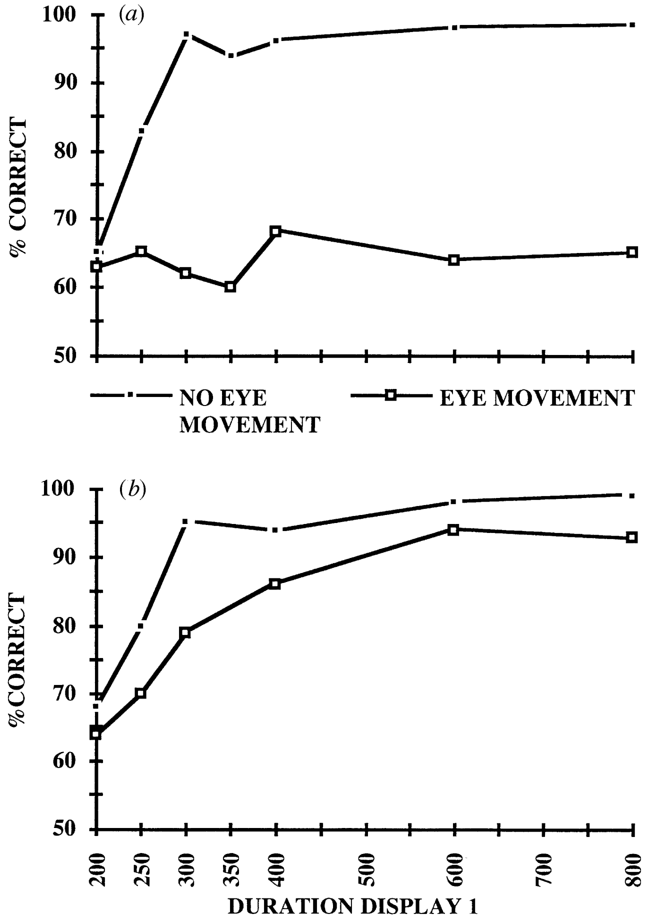

Figure 5. Data from five subjects on the dot localization task, as a function of the number of dots present in the displays. (*a*) Sparse displays: one target, field 20, ISI 400 ms. (*b*) Dense displays: one target, field 40, ISI 400 ms.

These data suggest that memory for visual location can depend on several factors. First, a position memory can be maintained for a dot in an otherwise sparsely populated space, provided fixation does not shift from the location. Here performance depends either on a retinotopic code for the dot location or on some form of proprioceptive memory for eye position (e.g. subjects can judge whether the dot has shifted by keeping eye position constant and judging whether there is a change in acuity when the dot is re-presented—just as one might discriminate whether a shift has occurred by pointing a finger at the location and judging whether the dot reappears where the finger is). Second, a fine position memory for a single dot cannot be maintained across an eye movement (in the condition with sparse displays). Third, position memory across fixation is based on the coding of the relative locations of the dots—hence position memory is found with dense displays but not with sparse displays.

Our results indicate that it is extremely difficult to make fine spatial memory judgements across fixation in a sparsely coded visual world. The data are consistent with other studies on the integration (rather than memory, as here) of visual information across eye movements, which show the use of abstract information about the relative positions of visual elements rather than some form of buffer that codes visual space (see Irwin 1993). Hence we do not appear to encode some form of blackboard memory for empty space,

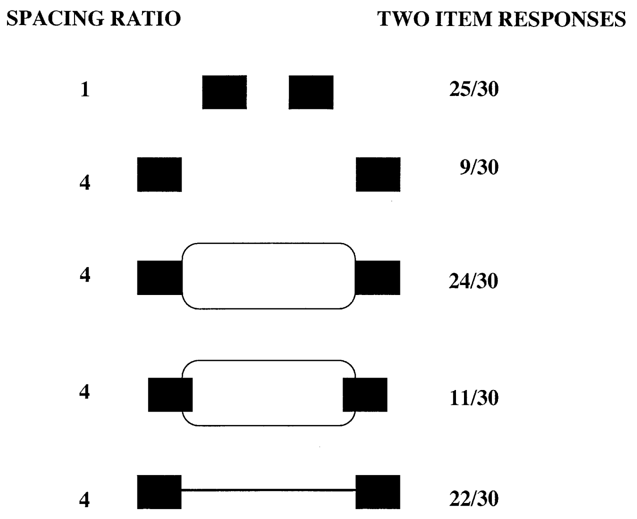

Figure 6. Data on extinction from patient G.K. The task required discrimination between displays with none, one or two items present; performance on no- and one-item displays was good and data are shown here only for two-item trials. There is extinction with two squares with a wide spacing ratio (size of spacing to size of elements) and when the squares maintain this ratio and occlude the background shape.

which can be maintained when the eyes move. Across fixation, memory for location depends on coding the relative positions of objects. Given that the dots were randomly presented here, and so tended not to form any learned configuration, we suggest that performance was determined by a between-object spatial code.

(b) *Spatial effects on visual selection are modulated by object coding*

If visual elements are coded in terms of their positions on a spatial blackboard, then we would expect the distance between elements on the blackboard to affect both their coding into objects and how attention may then be applied to the elements. Certainly grouping by proximity should decrease as inter-element distance increases, as may other forms of grouping. Strong neuropsychological evidence for this comes again from studies of visual extinction. As we have already noted, Gilchrist *et al.* (1996) showed that extinction could be reduced when elements had both the same brightness and collinear edges (with collinearity producing an effect greater than that due to common brightness alone; figure 1). However, when the relative distance between the elements increased so that the distance was more than three times the size of each stimulus, extinction was restored (see figure 6). Now neither collinearity nor common brightness was sufficient to group the contra- and ipsilesional items together. Distance effects have also been shown to modulate selection in normal observers. For example, the time it takes normal subjects to switch attention from one object to another increases with their distance (Vecera 1994) and interference between incongruent items is larger when they are close than when they are distant, even when they are part of the same object (Kramer & Jacobson 1991).

Yet, other evidence from my laboratory shows that the coding of distance itself is modulated by grouping between stimuli. Again by using the extinction paradigm, we have found that the effects of distance on extinction can be eliminated if the space between elements is filled with an occluding figure. By using pairs of squares, we showed that there is re-grouping over distances that are otherwise sufficient to introduce extinction if the elements are coded to be grouped are separated by an occluder (figure 6)—an effect that may be owing to amodal completion. When the elements are moved slightly closer, so that they now occlude the background shape, extinction occurs again. Quite similar effects of occlusion on extinction have been reported by Mattingly *et al.* (1997). The present data extend them by showing that grouping between the occluded elements modulates distance effects in visual selection. Note that such effects run counter to a simple blackboard model of spatial coding; spatial representations deal with the relations between objects and are modulated by grouping between parts.

4. SPATIAL REPRESENTATION AND ATTENTION

I have argued that: (i) independent representations of space within and between objects are encoded in parallel; (ii) the object representation that mediates performance can depend on the task; and (iii) forms of spatial representation other than those that code space within or between objects are limited.

(a) *Implications for selective visual attention*

These proposals have implications for understanding selective visual attention. I discuss three here.

1. The evidence indicates that both within- and between-object codes are assembled in a spatially parallel manner, without the need for focal attention to be paid to the elements. For example, consider studies of reading and counting. The data here demonstrate that each task can be unaffected by the number of elements present (i.e. each task is supported by parallel visual coding) at least up to a certain limit (three to four items, for counting), although the tasks also dissociate following different brain lesions (figure 3).
2. Although each representation is encoded in parallel, there are limits. For example, between-object codes seem constrained to just a small number of objects at a time, as indicated by the limit on subitization (four items or less). I suggest that the parallel coding of a small number of objects helps to provide useful information for future actions (for example, for navigation between separate objects, for bimanual actions and so on), without overloading the systems that control action. There are also limits on the number of within-object representations that can be used simultaneously for a response. Were there no limits at this level, we would be able to identify several stimuli presented simultaneously; however, there is considerable evidence against this (Duncan 1984; Baylis & Driver 1993). Elements encoded in parallel still need to be selected serially for identification, as one within-object description at a time. Interestingly, one consequence of preventing serial selection of numerous stimuli is that elements from each stimulus can be amalgamated into a single response (McClelland & Mozer 1986). Serial selection of within-object representations may thus be

computationally useful to prevent cross-talk during the identification process. The neuropsychological phenomenon of extinction may also arise because of the limitation in the number of within-object codes that can selected at one time. The data on recovery from extinction by grouping indicates that elements can be processed in parallel even when they are subsequently extinguished—when grouping does not take place. Selection is limited by whether elements can be grouped into a single within-object spatial representation. However, if there is dual coding of within- and between-object representations, why do patients seem unable to use between-object codes to at least detect the presence of numerous items? This may be because the lesion, typically to the dorsal visual system, disrupts the use of between-object codes; this will make such patients reliant on within-object codes for detection and identification. We consider this point further.

3. There need to be forms of task-based selection, that determine whether within- or between-object codes guide responses (as in the reading and counting data). Task-based selection may operate by priming one form of coding so that it is selected before the other (cf. Chelazzi *et al.* 1993). Alternatively, within a system in which spatial attention and object recognition systems are coupled by re-entrant connections to early visual processing (cf. Humphreys & Riddoch 1993), attentional activation of one rather than another between-object representation could bias selection towards the within-object code at an associated location—much as envisaged by space-based theories of visual selection (Treisman, this issue). This may then control whether whole stimuli or their parts are selected for the response. What should be noted here, however, is that in such a model attention is not paid to empty space but only to an occupied location (signalled by the between-object code).

The account I have suggested differs from other current theories of visual selection in several respects. For example, unlike feature integration theory (Treisman, this issue), I propose that form elements are bound together without focal attention. On the account I have proposed, attention may be involved in selecting between several within-object descriptions for object identification, but it is not necessary either to encode the elements or to group them into within-object descriptions in the first place. This leaves to one side the question of how form descriptions are linked to surface properties of objects, which may require additional processes to those discussed here. Also, unlike the theory of visual attention (Bundesen, this issue), I suggest that there are dual forms of perceptual categorization, one to form within-object representations and one to form between-object representations. These dual forms of categorization can be dissociated after brain damage.

(b) *Brain mechanisms*

The evidence on reading and counting is consistent with the proposal that within-object codes are assimilated in parallel within the ventral visual system, whereas between-object codes are assimilated within the dorsal visual stream. Indeed, it may be the massive impairment of between-object codes in cases of bilateral parietal damage that causes the devastating loss of visual information for spatial guidance of action (see, for example, in Balint's syndrome; see also Treisman, this issue), although object recognition *per se* is relatively preserved (performed via the ventral stream). As I have already pointed out though, this still leaves unexplained why there is neglect within objects after putative dorsal lesions (Young *et al.* 1992; Humphreys & Riddoch 1994, 1995; Humphreys & Heinke 1998). There are two possibilities, as follows. It may be that the inferior parietal lobe actually forms part of a ventral or overlapping ventral-dorsal processing area that deals with aspects of object coding, and so lesions to this area affect object recognition (see Milner & Goodale 1995). Alternatively, dorsal processing areas may be recruited when attention needs to be switched from one part of an object to another (cf. Egly *et al.* 1994) or when the spatial relations between parts are important for identification (see Humphreys & Heinke 1998). Defining the brain regions mediating different forms of object representation constitutes only the first step in understanding how they may interact to determine behaviour across wide sets of circumstances.

This work was supported by grants from the Biological and Biotechnology Research Council and the Medical Research Council of the UK and from the Human Frontier Science Program. The research has benefited from collaboration and discussion with a large number of friends and colleagues whom I thank: L. Alston, A. Cooper, L. Fuentes, I. Gilchrist, S. Kyllingsbaek, A. Olson, K. O'Regan, J. Riddoch and D. Watson. Thanks also to A. Treisman and R. Ward for insightful comments on earlier drafts of the paper.

REFERENCES

Adams, M. J. 1979 Models of word recognition. *Cogn. Psychol.* **11**, 133–176.

Balint, R. 1909 Seelenlahmung des 'Schauens', optische Ataxie, raumliche Storung der Aufmerksamkeit. *Manatsschr. Psychiatr. Neurol.* **25**, 51–81.

Baylis, G. C. & Driver, J. 1993 Visual attention and objects: evidence for hierarchical coding of location. *J. Exp. Psychol. Hum. Percept. Perf.* **19**, 451–470.

Baylis, G. C. & Driver, J. 1994 Parallel computation of symmetry but not repetition within single visual shapes. *Vis. Cogn.* **1**, 377–400.

Baylis, G. C. & Driver, J. 1998 Perception of symmetry and repetition within and across visual shapes: part-descriptions and object-based attention. *Vis. Cogn.* (In the press.)

Baylis, G. C., Driver, J., Baylis, L. L. & Rafal, R. D. 1994 Reading of letters and words in a patient with Balint's syndrome. *Neuropsychologia* **32**, 1273–1286.

Bisiach, E. & Rusconi, M. L. 1990 Breakdown of perceptual awareness in unilateral neglect. *Cortex* **26**, 643–649.

Bruce, V., Green, P. & Georgeson, M. A. 1996 *Visual perception*, 3rd edn. London: Psychology Press.

Buck, B. H., Black, S. E., Behrmann, M., Caldwell, C. & Bronskill, M. J. 1998 Spatial-and object-based attentional deficits in Alzheimer's disease: relationship to HMPAOSPECT measures of parietal performance. *Brain.* (In the press.)

Chelazzi, L., Miller, E. K., Duncan, J. & Desimone, R. 1993 A neural basis for visual search in inferior temporal cortex. *Nature* **363**, 345–347.

Dehaene, S. & Cohen, L. 1994 Dissociable mechanisms of subitizing and counting: neuropsychological evidence from simultanagnosic patients. *J. Exp. Psychol. Hum. Percept. Perf.* **20**, 958–975.

Dejerine, J. 1892 Contribution a l'étude anatomoclinique et clinique des differentes variétés de cécité verbale. *Mém. Soc. Biol.* **4**, 61–90.

Donnelly, N., Humphreys, G. W. & Riddoch, M. J. 1991 Parallel computation of primitive shape descriptions. *J. Exp. Psychol. Hum. Percept. Perf.* **17**, 561–570.

Duncan, J. 1984 Selective attention and the organization of visual information. *J. Exp. Psychol. Gen.* **113**, 501–517.

Duncan, J. 1996 Co-operating brain systems in selective perception and action. In *Attention and performance XV* (ed. T. Inui & J. L. McClelland), pp. 549–578. Cambridge, MA: MIT Press.

Duncan, J., Humphreys, G. W. & Ward, R. 1997 Integrated mechanisms of selective attention. *Curr. Opin. Biol.* **7**, 255–261.

Egly, R., Driver, J. & Rafal, R. D. 1994 Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J. Exp. Psychol. Hum. Percept. Perf.* **123**, 161–177.

Eriksen, C. W. & Yeh, Y.-Y. 1985 Allocation of attention in the visual field. *J. Exp. Psychol. Hum. Percept. Perf.* **11**, 583–597.

Farah, M. J., Wallace, M. A. & Vecera, S. P. 1993 'What' and 'where' in visual attention: evidence from the neglect syndrome. In *Unilateral neglect: clinical and experimental studies* (ed. I. H. Robertson & J. C. Marshall), pp. 123–138. London: Lawrence Erlbaum.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S. J. & Dolan, R. J. 1996 Where in the brain does visual attention select the forest and the trees? *Nature* **382**, 626–628.

Frederiksen, J. R. & Kroll, J. F. 1976 Spelling and sound: approaches to the internal lexicon. *J. Exp. Psychol. Hum. Percept. Perf.* **2**, 361–379.

Friedman, R. B. & Alexander, M. P. 1984 Pictures, images and pure alexia: a case study. *Cogn. Neurol.* **1**, 9–23.

Gilchrist, I., Humphreys, G. W. & Riddoch, M. J. 1996 Grouping and extinction: evidence for low-level modulation of selection. *Cogn. Neurol.* **13**, 1223–1256.

Heilman, K. & Valenstein, E. 1979 Mechanisms underlying hemispatial neglect. *Ann. Neurol.* **5**, 166–170.

Humphreys, G. W. & Bruce, V. 1989 *Visual cognition: computational, experimental and neuropsychological perspectives*. London: Lawrence Erlbaum.

Humphreys, G. W. & Heinke, D. 1998 Spatial representation and selection in the brain: neuropsychological and computational constraints. *Vis. Cogn.* (In the press.)

Humphreys, G. W. & Riddoch, M. J. 1987 *To see but not to see: a case study of visual agnosia*. London: Lawrence Erlbaum.

Humphreys, G. W. & Riddoch, M. J. 1993 Interactions between object and space systems revealed through neuropsychology. In *Attention and performance XIV* (ed. D. E. Meyer & S. Kornblum), pp. 143–162. Cambridge, MA: MIT Press.

Humphreys, G. W. & Riddoch, M. J. 1994 Attention to within-object and between-object spatial representations: multiple sites for visual selection. *Cogn. Neurol.* **11**, 207–242.

Humphreys, G. W. & Riddoch, M. J. 1995 Separate coding of space within and between perceptual objects: evidence from unilateral visual neglect. *Cogn. Neurol.* **12**, 283–312.

Humphreys, G. W., Riddoch, M. J. & Quinlan, P. T. 1985 Interactive processes in perceptual organization: evidence from visual agnosia. In *Attention and performance XI* (ed. M. I. Posner & O. S. M. Marin), pp. 301–318. Hillsdale, NJ: Lawrence Erlbaum.

Humphreys, G. W., Riddoch, M. J., Quinlan, P. T., Price, C. J. & Donnelly, N. 1992 Parallel pattern processing in visual agnosia. *Can. J. Psychol.* **46**, 377–416.

Humphreys, G. W., Keulers, N. & Donnelly, N. 1994*a* Parallel visual coding in three dimensions. *Perception* **23**, 453–470.

Humphreys, G. W., Romani, C., Olson, A., Riddoch, M. J. & Duncan, J. 1994*b* Non-spatial extinction following lesions of the parietal lobe in humans. *Nature* **372**, 357–359.

Humphreys, G. W., Riddoch, M. J. & Price, C. J. 1997 Top-down processes in object identification: evidence from experimental psychology, neuropsychology and functional anatomy. *Phil. Trans. R. Soc. Lond.* B **352**, 1275–1282.

Irwin, D. E. 1993 Perceiving an integrated world. In *Attention and performance XIV* (ed. D. E. Meyer & S. Kornblum), pp. 121–143. Cambridge, MA: MIT Press.

Jacobs, R. A. & Jordan, M. I. 1992 Computational consequences of a bias towards short connections. *J. Cogn. Neurosci.* **4**, 323–336.

Karnath, H. O. 1988 Deficits of attention in acute and recovered visual hemi-neglect. *Neuropsychologia* **26**, 27–43.

Kinsbourne, M. & Warrington, E. K. 1962 A disorder of simultaneous form perception. *Brain* **85**, 461–486.

Kramer, A. F. & Jacobson, A. 1991 Perceptual organization and focused attention: the role of objects and proximity in visual processing. *Percept. Psychophysiol.* **50**, 267–284.

Ladavas, E., Paladini, R. & Cubelli, R. 1993 Implicit associative priming in a patient with left visual neglect. *Neuropsychologia* **31**, 1307–1320.

McClelland, J. L. 1976 Preliminary letter recognition in the perception of words and nonwords. *J. Exp. Psychol. Hum. Percept. Perf.* **2**, 80–91.

McClelland, J. L. & Mozer, M. C. 1986 Perceptual interactions in two-word displays: familiarity and similarity effects. *J. Exp. Psychol. Hum. Percept. Perf.* **12**, 18–35.

McClelland, J. L. & Rumelhart, D. E. 1981 An interactive activation model of context effects in letter perception. 1. An account of basic findings. *Psychol. Rev.* **88**, 375–407.

McGlinchey-Beroth, R., Milberg, W. P., Verfaillie, M., Alexander, M. & Kilduff, P. T. 1992 Semantic processing in the neglected visual field: evidence from a lexical decision task. *Cogn. Neurol.* **10**, 79–108.

Mack, A., Tang, B., Tuma, R., Kahn, S. & Rock, I. 1992 Perceptual organization and attention. *Cogn. Psychol.* **24**, 475–501.

Marr, D. 1982 *Vision*. San Francisco: W. H. Freeman & Co.

Marshall, J. C. & Halligan, P. W. 1988 Blindsight and insight into visuo-spatial neglect. *Nature* **336**, 766–767.

Matin, L. 1976 Saccades and extraretinal signal for visual direction. In *Eye movements and psychological processes* (ed. R. A. Monty & J. W. Sanders), pp. 205–219. Hillsdale, NJ: Lawrence Erlbaum.

Mattingly, J. B., Davis, G. & Driver, J. 1997 Pre-attentive filling-in of visual surfaces in parietal extinction. *Science* **275**, 671–674.

Mayall, K. A. & Humphreys, G. W. 1996 Case mixing and the task sensitive disruption of lexical processing. *J. Exp. Psychol. Learn. Mem. Cogn.* **22**, 278–294.

Milner, D. A. & Goodale, M. A. 1995 *The visual brain in action*. Oxford University Press.

Moore, C. M. & Egeth, H. 1997 Perception without attention: evidence of grouping under conditions of inattention. *J. Exp. Psychol. Hum. Percept. Perf.* **23**, 339–352.

Posner, M. I. 1980 Orienting of attention. *Q. J. Exp. Psychol.* **32**, 3–25.

Rapp, B. C. & Caramazza, A. 1991 Spatially determined deficits in letter and word processing. *Cogn. Neurol.* **8**, 275–311.

Riddoch, M. J. 1991 *Neglect and the peripheral dyslexias*. London: Lawrence Erlbaum.

Robertson, L. C., Lamb, M. R. & Knight, R. T. 1988 Effects of lesions of temporal-parietal junction on perceptual and attentional processing in humans. *J. Neurosci.* **8**, 3757–3769.

Rueckl, J. G., Cave, K. R. & Kosslyn, S. M. 1989 Why are 'what' and 'where' processed by separate cortical visual systems? A computational investigation. *J. Cogn. Neurosci.* **1**, 171–186.

Styles, E. 1997 *Attention*. London: Psychology Press.

Treisman, A. 1988 Features and objects: the Fourteenth Bartlett Memorial Lecture. *Q. J. Exp. Psychol.* A **40**, 201–237.

Trick, L. & Pylyshyn, Z. 1993 What enumeration studies can show us about spatial attention: evidence for limited capacity preattentive processing. *J. Exp. Psychol. Hum. Percept. Perf.* **101**, 80–102.

Ungerleider, L. G. & Mishkin, M. 1982 Two cortical visual systems. In *Analysis of visual behavior* (ed. J. Ingle, M. A. Goodale & R. J. W. Mansfield), pp. 549–586. Cambridge, MA: MIT Press.

Vecera, S. P. 1994 Grouped locations and object-based attention: comment on Egly, Driver and Rafal 1994. *J. Exp. Psychol. Gen.* **123**, 316–320.

Ward, R., Goodrich, S. & Driver, J. 1994 Grouping reduces visual extinction: neuropsychological evidence for weight-linkage in visual selection. *Vis. Cogn.* **1**, 101–130.

Young, A. W., Hellawell, D. J. & Welch, J. 1992 Neglect and visual recognition. *Brain* **112**, 51–71.