Detection by action: neuropsychological evidence for action-defined templates in search

Glyn W. Humphreys and M. Jane Riddoch

Behavioural Brain Sciences Centre, School of Psychology, University of Birmingham, Birmingham, B15 2TT, UK
Correspondence should be addressed to G.W.H. (g.w.humphreys@bham.ac.uk)

How do we detect a target in a cluttered environment? Here we present neuropsychological evidence that detection can be based on the action afforded by a target. A patient showing symptoms of unilateral neglect following damage to the right fronto-temporal-parietal region was slow and sometimes unable to find targets when they were defined by their name or even by a salient visual property (such as their color). In contrast, he was relatively efficient at finding a target defined by the action it afforded. Two other patients with neglect showed an opposite pattern; they were better at finding a target defined by its name. The data suggest that affordances can be effective even when a brain lesion limits the use of other properties in search tasks. The findings give evidence for a direct pragmatic route from vision to action in the brain.

Visual search tasks typically require us to find a particular target presented among set of distractors. An example is looking for a hammer in a box of tools. Current neurophysiological, psychological and computational models[1–5] hold that search can be directed in a top-down manner, by means of a memory template for the target. This template modulates processing, so that the competition for selection between the distractors and the target is biased for the target. Physiological evidence comes from a previous study[1] in which monkeys were cued with the features of a forthcoming target (such as a red circle) and were required to make an eye movement to this target when it was subsequently presented along with a distractor stimulus. Cells in the inferotemporal cortex responsive to the features of the target remained active during the interval after the cue. This also further boosted the activation of 'target' cells when the target and distractor appeared, enabling the target to be selected for the eye movement response. The activity of cells responsive to target features before the critical display provided a memory template for the target.

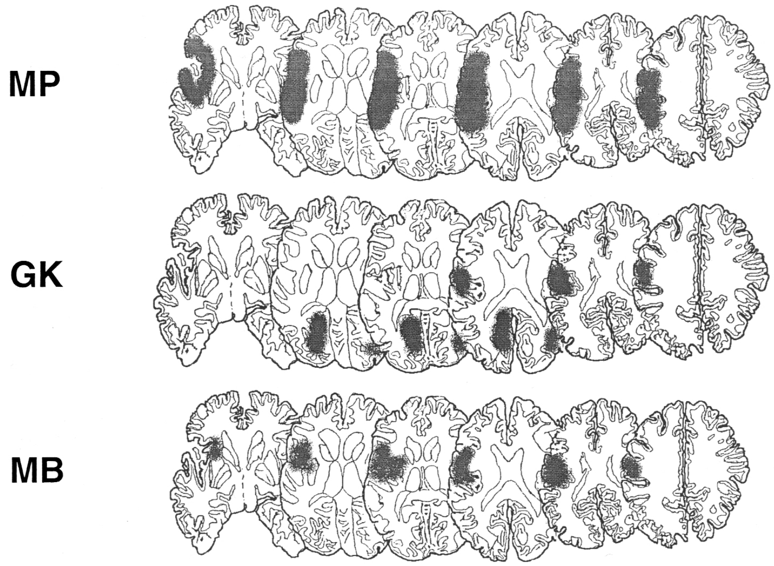
In most studies, the memory template of the target has been specified in perceptual terms, based on features such as the target's shape, color, size or pattern of motion. We do not know whether templates can take other forms. For example, can templates be defined not in specific perceptual terms, but in terms of the goals of an action? In everyday circumstances, we find ourselves in situations in which we do not have a particular target in mind, but must nevertheless find an object that will do a given task. For example, we might look for something to hammer a nail when there is no hammer available. Can we use a template based on the action needed, so that the search is directed to appropriate objects (for example, those large enough and with a surface hard enough to act as a hammer)?

Psychologists such as J.J. Gibson[6] have argued that objects can 'afford' an action, based on the overlap between the perceptual features of the objects and the goal of the actor. A rock may 'afford' hammering because it is sufficiently large, it can be gripped in an appropriate manner, and it has a hard surface. Memory templates set by the goal of an action may be responsive to affordances, biasing selection to one object in a scene. (We pick up a rock rather than a stick, for instance.)

Here we examined whether memory templates while searching are influenced by affordances. The study focused on a patient with symptoms of unilateral neglect following damage to fronto-temporal-parietal regions of the right hemisphere, who was markedly impaired at finding targets defined by their perceptual features ("find the red object") or their names ("find the cup"). The patient, MP, remarked that sometimes he could find objects if he thought of what to do with them. We evaluated whether this was indeed the case, and whether a dissociation existed between his ability to use a perceptual template ("find the red object" or "find the cup") and a template for action ("find the object you could drink from"). In experiments 1–3 and 5, we presented MP with arrays of 10 everyday objects oriented with their handles toward him; his task was to point to a target defined in various ways. In experiment 4, the objects were replaced with words written on cards, and in experiment 6, we used pictures of objects and plausible non-objects. MP's performance in experiment 1 was contrasted with the performance of two other patients who also showed symptoms of visual neglect, but who never reported this ability to find things by thinking of what to do with them. Clinical details on the lesions sustained by the patients are shown in Fig. 1.

Results

In experiment 1, for MP, the target was defined on separate trials either by its name (cup), its color (red) or by an associated and/or plausible action (an item you would drink from). MB and GK performed in just the name and action conditions. Each object was a different color, each had a unique name, and each had a definable function (an object you would use to drink from, to write with, *etc.*). MP, MB and GK were each able to name all the objects and colors in single-item displays.



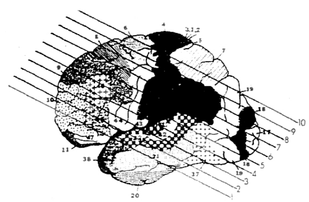


Fig. 1. Lesion reconstructions in the patients, from MRI scan. Lesions have been drawn onto standard slices[7]. Bottom, the 10 slices used. Only slices three to eight are depicted here. Left of each slice represents the right hemisphere.

For MP, target definition affected search performance (for accuracy, $F_{2,18} = 7.94$, $p < 0.001$; for reaction time (RT), $F_{2,18} = 37.13$, $p < 0.001$). He was better in the action condition than in the name condition (for accuracy, $F_{1,9} = 11.25$, $p < 0.01$; for RT, $F_{1,9} = 22.49$, $p < 0.01$) and the color condition (for accuracy, $F_{1,9} = 5.0$, $p < 0.05$; for RT, $F_{1,9} = 51.62$, $p < 0.001$). Any advantage in accuracy for the color condition relative to the name condition was canceled by the slower RTs for color search (**Fig. 2**). On error trials, MP reported that he could not find the target (on 13 of 18 error trials, the target was on MP's left; this bias held across all the current experiments).

In contrast to MP, MB and GK were better in the name than in the action condition. MB responded promptly but made two left-side omissions across trials (39/40 for each condition). Her mean RTs in the name and action conditions were 1.1 and 1.5 seconds, respectively ($F_{1,9} = 5.82$, $p < 0.05$). GK was slower overall but showed the same pattern; his RTs for the name and action conditions were 9.7 and 14.3 seconds, respectively ($F_{1,9} = 5.18$, $p < 0.05$). He made only five misses (37/40 and 38/40 correct for action and name-defined search), all to left-side targets.

In experiment 2, we examined MP's search for targets defined by their name, a verbal definition of an appropriate action, or a physically different exemplar of the same basic type of object (for example, a different kind of cup). Performance was again affected by the way the target was defined (for accuracy, $F_{2,18} = 8.53$, $p < 0.01$; for RTs, $F_{2,18} = 56.5$, $p < 0.001$). There was an advantage for search defined by action, and no differences for 'name' and 'exemplar'-defined targets. For action versus name, $F_{1,9}$ was 74.32 for RTs and 14.88 for accuracy, and for action versus exemplar, $F_{1,9}$ was 83.16 for RTs and 10.29 for accuracy (all $p < 0.01$; **Fig. 2**).

In experiment 3, targets were defined either by their name or by a gesture made by the experimenter (for example, a drinking action for a cup). MP performed better in the gesture than in the name condition (for RT, $F_{1,9} = 44.1$; for accuracy, $F_{1,9} = 32.1$, both $p < 0.001$; **Fig. 2**).

MP's search was consistently faster and more accurate for targets defined by action (either by verbal description or by gesture) than for targets defined by name (and even by their color or by perceptual features from another functionally equivalent exemplar). Experiment 4 tested the generality of this result by replacing objects with words. MP's better performance in the action conditions may have been due to higher arousal from the action cues[8]. However, in this case, similarly high arousal should occur when word rather than object stimuli are presented. In contrast, effects due to affordances should be specific to objects; words do not have shapes or textures that relate in a direct way to action. Targets were defined by a description of their associated action ("find the word for an object you would drink from") or by their name. Performance did not differ for tasks in which words were defined by action rather than by name ($F < 1.0$ for both accuracy and RT). Performance in the action condition was worse than in experiments 1 and 2, when object targets were used (experiment 1 and 2 versus experiment 4, for accuracy, $F_{1,9} = 17.99$, $p < 0.01$ for both comparisons; for RTs, $F_{1,9} = 12.42$ and 10.6, $p < 0.01$). This argues against effects due to increased arousal. Also, searching in the 'name' condition did not differ across the studies (all $F < 1.0$, for accuracy and RT), so the equivalence of the 'action' and 'name' conditions was not due to an improvement in the 'name' condition. The result is consistent with templates for action being activated by affordances from objects.

Experiment 5 varied the viewing positions of the objects. Either the objects had their handles oriented toward MP (as in experiment 1–3), or the objects were rotated so that each handle faced away from him. Objects were cued by a definition of their action or by their name. For both accuracy and RTs, there were reliable main effects of object orientation and cueing procedure (for accuracy, $F_{1,9} = 37.6$ and 4.89, $p < 0.001$ and 0.05 respectively; for RTs, $F_{1,9} = 56.18$ and 35.46, both $p < 0.001$), and there were interactions between these factors (for accuracy, $F_{1,9} = 6.88$,

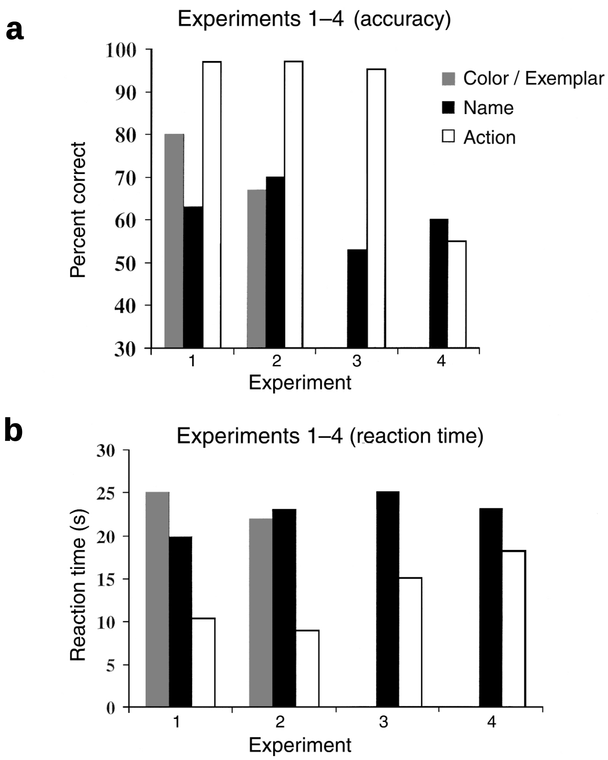


Fig. 2. MP's performance across experiments 1–4. (**a**) Percentage correct responses in the different target cueing conditions, across experiments 1–4. (**b**) Mean correct RTs (s) across experiments 1–4. Color/Exemplar, effects of color cueing in experiment 1 and cueing by a functionally equivalent exemplar in experiment 2. Name, cueing by object name. Action, cueing either by a verbal definition of the action (experiments 1, 2 and 4) or by a gesture (experiment 3). In experiments 1–3, the stimuli were objects, whereas in experiment 4, the stimuli were words.

$p < 0.03$; for RTs, $F_{1,9} = 25.62$, 0.001). When search was cued by action, and the handles of the objects faced toward MP, his performance was faster and more accurate than in the other three conditions ($p < 0.01$, Scheffe tests). However, his poorer performance when handles faced away was not because the objects were harder to recognize; their change in orientation did not affect the name condition (**Fig. 3**). However, affordances for action are reduced when objects are not oriented appropriately.

In experiment 6, we assessed performance with non-objects as well as with objects (examples, **Fig. 4**). In an 'exemplar' condition, MP was given a picture of the target that he had to find; in the 'action' condition, he was told to find the object for which he would make a 'twisting action.'

With objects, MP scored 16/32 and 22/32 in the exemplar and action conditions; with non-objects, he scored 12/32 and 22/32 respectively. There was an overall effect of task on accuracy ($F_{1,7} = 22.4$, $p < 0.01$), no effect of object type (object versus non-object), and no interaction ($F_{1,7} = 1.75$ and $F < 1.0$, both $p > 0.05$; **Fig. 5**). On RTs, there were significant main effects of task and object type ($F_{1,7} = 42.08$ and 6.96, $p < 0.001$ and 0.05, respectively), and a reliable interaction effect ($F_{1,7} = 6.08$, $p < 0.05$). For action-defined targets, there was no difference for objects and non-objects ($F < 1.0$). For exemplar-defined targets, search was faster for objects than for non-objects ($F_{1,7} = 15.86$, $p < 0.01$). However, with both object and non-objects, performance was faster in the action condition (for objects, $F_{1,7} = 15.26$; for non-objects, $F_{1,7} = 49.43$, both $p < 0.01$).

Discussion

MP showed aspects of unilateral neglect, often failing to detect targets presented on his left side. However, our data reveal that his performance was affected by how objects were cued. MP was more accurate and faster at detecting targets cued by action than targets cued by name, by color or by a functionally equivalent item. This result occurred if the stimuli were objects rather than words (experiment 4), and if objects were oriented for grasping (experiment 5); it also occurred with both objects and non-objects (experiment 6). In contrast, two other patients with unilateral neglect were better able to find target objects defined by their names rather than by action (experiment 1). The double dissociation between MP and the other patients indicates that action- and name-defined templates are functionally independent in the brain.

Our results imply that search can be based on intended actions, and not just on the perceptual properties of objects. MP could name individual objects and colors, but he had difficulty with these cues in search. This finding suggests an impairment in linking perceptual cues in memory to objects, when multiple objects compete for attention. In contrast, MP could detect the same targets when cued by action. This indicates that action templates can influence visual search and selection independently of perceptual templates of targets. Normally, search based on action may be slower than search based on perceptual templates (as in patients MB and GK), but action templates may still be useful when other memory templates are made unavailable because of brain damage.

A second implication of our results is that action templates are activated by affordances offered by objects, that is, by parts

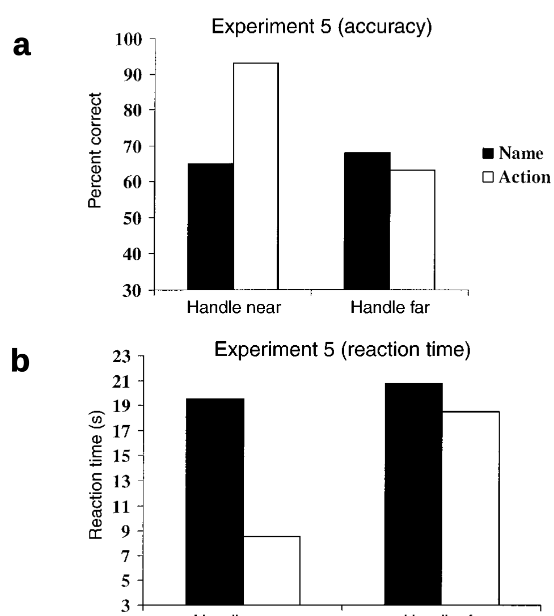


Fig. 3. RTs for MP in experiment 5. (**a**) Percentage correct responses in experiment 5. (**b**) Mean correct RTs. In the 'handle near' condition, objects were placed with their handles facing MP. In the 'handle far' condition they were placed with their handle away from MP. Targets were cued by their name or by a verbal definition of their action.

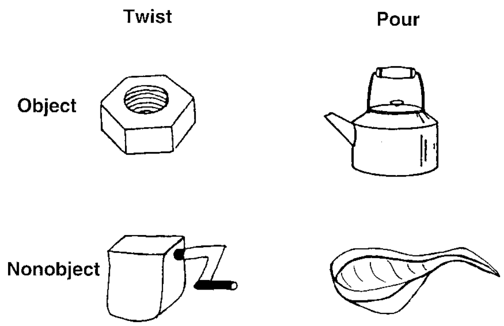


Fig. 4. Example object and non-object stimuli used in experiment 6. Targets were objects used with a 'twisting' action (left) and distractors were objects that would be used with a 'pouring' action (right).

of objects directly associated with action. Thus, the advantage for cueing by action disappeared when the objects were replaced by words, which do not afford action (experiment 4), and when the objects were oriented inappropriately (experiment 5). For example , cups are less likely to afford action when their handles face away from the subject. One argument against an affordance account is that MP performed well for action-defined targets with both pictures and real objects (in experiment 6). However, work with normal participants shows that pictures also afford action (though perhaps less strongly than real objects)[9]. The argument for affordances is supported further by the action advantage found with non-objects that could easily be discriminated from real objects (experiment 6). With non-objects, responses are unlikely to be linked to a stored memory representation. Indeed, exemplar-based, but not action-based, searching was faster for objects than for non-objects (**Fig. 5**), consistent with recognition—but not action—favoring the familiar items.

How might affordances, offered by objects and non-objects, benefit MP's search? MP's long search times suggest that he searched arrays serially, even with objects cued by action. Thus, affordances seemed not to guide search directly to a target but rather to enable targets to be detected. There seemed to be facilitation of the process of matching a stimulus to a search template, but no clear facilitation of search itself.

Physiological evidence indicates that templates for the perceptual features of targets exist in the ventral visual stream[1], though these templates may be contingent on activation in the frontal lobes determined by task instructions[10]. MP is impaired in using perceptual templates to guide search, perhaps because the damage to his right temporal lobe disrupts top-down priming of perception. This impairment, coupled with poor allocation of spatial attention to the left due to his right parietal lesion, leads to both slow and error-prone performance to left-side targets. The damage to MP's right temporal lobe is more extensive than that suffered by either MB or GK, who may thus represent templates for the perceptual features of targets (**Fig. 1**). Nevertheless, templates defined by intended actions continue to operate for MP, and they are activated by affordances from the objects present. These affordances may operate within a 'pragmatic' route to action within which neurons respond according to how an object may be used, but not according to the object's identity (a process dependent on ventral object recognition). Consistent with this idea, patients with temporal lobe damage and impaired semantic knowledge of objects can decide how an object should be used even when their semantic judgments about the objects are impaired[11]. We suggest that MP remained able to use action-templates within this pragmatic route to action. The spared action templates enabled him to detect targets on his affected side.

The third implication of the data is that, despite making errors by failing to detect targets, MP must have processed the objects enough to enable action templates to be activated by affordances. This adds to the growing body of evidence indicating that there can be implicit processing of displays in patients who show visual neglect[12].

The cases

MP, a left-handed former tool worker born on January 19, 1947, suffered an aneurysm of the right middle cerebral artery in 1992, resulting in cerebral artery occlusion and infarct. There was a mild left hemiparesis for his upper limb. Magnetic resonance imaging (MRI) and single photon emission computed tomography (SPECT) scans showed that damage occurred in fronto-temporo-parietal regions of his right hemisphere, including the inferior frontal gyrus, the superior temporal gyrus, the supramarginal and angular gyri and the post-central gyrus. MP exhibited unilateral left neglect, particularly in tasks requiring him to scan across cluttered visual arrays. On the standardized Behavioural Inattention Test (BIT)[13] he scored 94/146, showing particularly poor performance on line crossing, star cancellation and letter cancellation tasks, where there were omissions of stimuli on both the far left and right sides of the page. We found evidence of neglect in line bisection, particularly when the task involved pointing to the center of the line[14]. MP made few errors in reading single words, but whole words could be omitted from the left side of the page when reading text (12 errors across 5 pages). He also had problems in detecting both the facial identity and gender of the left side of chimeric faces (scoring 0/20 at identifying the left-side face and 10/22, chance, at discriminating its gender). There was a mild perceptual impairment on the Visual Object and Space Perception (VOSP) test[15], where he was impaired at the 5% level on the dot counting, position discrimination and number location sub-tests. He had no major deficits on picture naming (70/76 on the full naming test from Birmingham Object Recognition Battery; BORB[16]) and no major problems in dealing with overlapping perceptual features (15/18 items named correctly with both overlapping and non-overlapping items from BORB). On the Warrington Recognition Memory test, he performed better for words rather than faces (35/50 versus 25/50), consistent with right hemisphere pathology. MP's verbal short-term memory was impaired (digit span 4), and he also

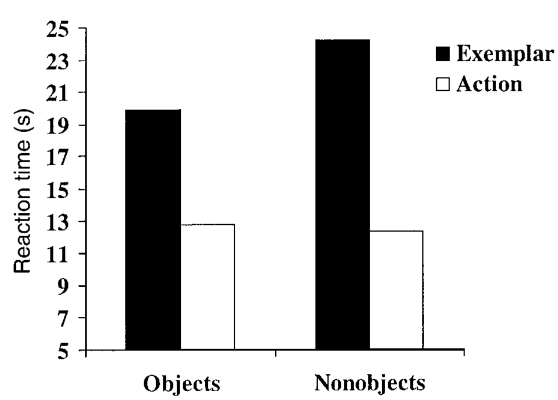


Fig. 5. Mean correct RTs for MP in experiment 6.

had problems in mental arithmetic. These last two problems are unusual after right hemisphere damage, but may reflect some degree of crossed laterality in MP's case.

GK, born on January 8, 1939, suffered two consecutive strokes in 1986. These produced bilateral lesions affecting the right medial occipito-parietal region (including the cuneus and precuneus), the right tempero-parietal region and the left tempero-parietal region. Subsequently, GK had a number of neuropsychological deficits, most notably, Balint's syndrome[17], which causes optic ataxia (visual misreaching) and simultanagnosia (seeming to 'see' only one object at a time). In addition, GK manifests aspects of unilateral neglect. He can make left-end errors in reading words, he bisects lines to the left of true center, and he shows left-side extinction with laterally presented stimuli[18,19].

MB, born on February 5, 1942, suffered a stroke in 1995 that affected several regions within her right hemisphere: the inferior frontal and superior temporal gyri (minimally), the inferior parietal lobule (affecting the supramarginal gyrus but sparing the angular gyrus) and the ventral putamen. We found no field defect on confrontation testing. However, we found left-neglect on initial clinical testing; she missed targets from the left third of the field on a cancellation test, and she showed a right bias in line bisection. At the time of testing, the more marked aspects of her neglect had decreased, though neglect could be demonstrated using brief visual presentations[20].

Methods

All experiments were approved by the Psychology Ethical Committee, Birmingham University. Patients sat at a table 120 cm wide by 60 cm deep. Before each trial, a cardboard shield was placed in front of the array. In experiment 1, ten objects were arranged in preset locations, five to the right and five to the left of midline. On each side of midline, 3 objects were at a depth of 40 cm, and 2 were at depths of 50 cm. Before each trial, the objects were rearranged and randomly assigned to new locations relative to the previous trial, with the constraint that the target had to fall no more than once in a given location across the block. Thus, each location was sampled once across a block. Each object was also chosen as a target once per block, and each condition (cue by color, name or action, for MP) was presented in a block of 10 trials. The order of the blocks was then counterbalanced across test sessions. For MP, the color, name and action cues were presented in an order of ABC, BCA and CAB over three sessions. Consequently, there were 30 trials per condition, with each object being cued as a target 3 times. For MB and GK, the name and action conditions were also presented across three blocks, using the orders AB, BA, AB for MB and BA, AB, BA for GK.

The procedure for experiment 2, done only by MP, was similar. In this experiment, cueing conditions were by name, verbal definition of action, or the showing of a functionally equivalent exemplar to the target. In experiments 3 and 4, there were 2 cueing conditions presented in an ABBA, BAAB design across 2 sessions (to create 40 trials per condition). In experiment 5, there were four cueing conditions (by name and by verbal definition of action, with objects' handles facing toward or away from MP). These were presented in an ABCDDCBA design in one session and a DCBAABCD design in a second session (again creating 40 trials per condition). The following objects were used: cup (red), hammer (wooden), hairbrush (blue), pen (black), key (brass), scissors (light silver), teapot (brown), screwdriver (yellow), hacksaw (dark metal), stapler (green). For experiment 6, we used line drawings of real objects and non-objects that a set of 40 independent subjects had judged would be used for either 'pouring or twisting.' Stimuli were included only if there was 90%+ agreement for the 'pour'/'twist' action over judges. The non-objects were designed to make it difficult to judge if they were derived from any particular 'parent' object, but they contained parts about which the independent subjects nevertheless made consistent action decisions. Seven real object distractors and seven non-object distractors ('pouring' stimuli) were used, along with eight different object and eight non-object targets ('twisting' stimuli). Stimuli were about 4×4 cm in size, each presented on a 7×7 cm card. The cards were arranged horizontally on a table, centered on MP's midline, at a distance of 30 cm. On each trial, MP saw pictures of either eight objects or eight non-objects (seven distractors, one target). A different target was presented on each trial in a block, and targets appeared in all spatial positions across a block. There were two cueing conditions. In the exemplar cue condition, MP was given a picture of the target, placed directly in front of him at his midline, and was asked to find this item. In the action cue condition, he was asked to find a target for which you would make a 'twisting' action. The four conditions (exemplar objects, action objects, exemplar non-objects, action non-objects) were presented across four sets of trial blocks in the following order: ABDC, CDBA, DCAB, BACD. An independent group of 10 control subjects had no difficulty in discriminating the real objects from the non-objects (100% correct performance).

Each trial began with the lifting of the divide between the patient and the stimuli, and search was timed by stopwatch. There was a time-out after 60 s. The data were analyzed by summing (for accuracy) or averaging (for RT) the results for each target across the test blocks (so, for accuracy there was a maximum score of three or four per object, depending on the study). Analyses of variance were done with target object as a random factor.

Acknowledgements

This work was supported by grants from the MRC and the Wellcome Trust. We thank MP, MB and GK for their participation in these studies and V. Chavda for help with scan reconstruction.

Received 15 August; accepted 20 November 2000

1. Chelazzi, L., Miller, E. K., Duncan, J. & Desimone, R. A neural basis for visual search in inferior temporal cortex. *Nature* **363**, 345–347 (1993).
2. Duncan, J. & Humphreys, G. W. Visual search and stimulus similarity. *Psych. Rev.* **96**, 433–458 (1989).
3. Wolfe, J. M. Guided search 2.0: a revised model of visual search. *Psych. Bull. Rev.* **1**, 202–238 (1994).
4. Humphreys, G. W. & Muller, H. M. SEarch via Recursive Rejection (SERR): A connectionist model of visual search. *Cognit. Psych.* **25**, 43–110 (1993).
5. Humphreys, G. W. & Heinke, D. Spatial representation and selection in the brain: neuropsychological and computational constraints. *Vis. Cognit.* **5**, 9–48 (1998).
6. Gibson, J. J. *The Ecological Approach to Perception* (Houghton Mifflin, Boston, 1979).
7. Gado, M., Hanaway, J. & Frank, R. Functional anatomy of the cerebral cortex by computed tomography. *J. Comput. Assist. Tomogr.* **3**, 1–19 (1979).
8. Robertson, I. H. & Manley, T. in *Attention, Space and Action: Studies in Cognitive Neuroscience* (eds. Humphreys, G. W., Duncan, J. & Treisman, A.) 297–310 (Oxford Univ. Press, Oxford, 1999).
9. Rumiati, R. I. & Humphreys, G. W. Recognition by action: dissociating visual and semantic routes to action in normal observers. *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 631–647 (1998).
10. Miller, E. K. in *Attention, Space and Action: Studies in Cognitive Neuroscience* (eds. Humphreys, G. W., Duncan, J. & Treisman, A.) 251–272 (Oxford Univ. Press, Oxford, 1999).
11. Hodges, J. R., Spatt, J. & Patterson, K. "What" and "how": evidence for the dissociation of object knowledge and mechanical problem-solving skills in the human brain. *Proc. Natl. Acad. Sci. USA* **96**, 9444–9448 (1999).
12. Driver, J. in *Attention* (ed. Pashler, H.) 279–310 (Taylor Francis, London, 1998).
13. Wilson, B., Cockburn, J. & Halligan, P. W. *Behavioural Inattention Test* (Thames Valley Test Company, Titchfield, Hants., 1987).
14. Edwards, M. G. & Humphreys, G. W. Pointing and grasping in unilateral visual neglect: effect of on-line visual feedback in grasping. *Neuropsychologia* **37**, 959–973 (1999).
15. Warrington, E. K. & James, M. *The VOSP Test of Visual Function* (NFER, Berkshire, 1991).
16. Riddoch, M. J. & Humphreys, G. W. *Birmingham Object Recognition Battery (BORB)* (Psychology Press, London, 1993).
17. Balint, R. Seelenlahmung des "Schauens": optische ataxie, raumliche Storung der Aufmerksamkeit. *Monatschrift fur Neurologie und Psychiatrie* **25**, 51–81 (1909).
18. Hall, D., Humphreys, G. W. & Cooper, A. Reading using supra-letter units: Evidence from attentional dyslexia. *Quart. J. Exp. Psychol.* (in press).
19. Gilchrist, I., Humphreys, G. W. & Riddoch, M. J. Grouping and extinction: evidence for low-level modulation of selection. *Cognit. Neuropsychol.* **13**, 1223–1256 (1996).
20. Shalev, L. & Humphreys, G. W. Biased attentional shifts associated with unilateral left neglect. *Cognit. Neuropsychol.* **17**, 339–364 (2000).