



CrossMark
click for updates

Review

Cite this article: Humphreys GW, Kumar S, Yoon EY, Wulff M, Roberts KL, Riddoch MJ. 2013 Attending to the possibilities of action. *Phil Trans R Soc B* 368: 20130059. <http://dx.doi.org/10.1098/rstb.2013.0059>

One contribution of 17 to a Theme Issue 'Attentional selection in visual perception, memory and action'.

Subject Areas:

cognition, neuroscience

Keywords:

visual affordance, action, attention

Author for correspondence:

Glyn W. Humphreys

e-mail: glyn.humphreys@psy.ox.ac.uk

Attending to the possibilities of action

Glyn W. Humphreys¹, Sanjay Kumar¹, Eun Young Yoon², Melanie Wulff³, Katherine L. Roberts⁴ and M. Jane Riddoch¹

¹Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK

²Korean NeuroTraining Center, Apsan-soonhwan Road 736, Nam-gu, Daegu, South Korea

³School of Psychology, University of Birmingham, Birmingham B15 2TT, UK

⁴Department of Psychology, University of Warwick, Coventry CV4 7AL, UK

Actions taking place in the environment are critical for our survival. We review evidence on attention to action, drawing on sets of converging evidence from neuropsychological patients through to studies of the time course and neural locus of action-based cueing of attention in normal observers. We show that the presence of action relations between stimuli helps reduce visual extinction in patients with limited attention to the contralesional side of space, while the first saccades made by normal observers and early perceptual and attentional responses measured using electroencephalography/event-related potentials are modulated by preparation of action and by seeing objects being grasped correctly or incorrectly for action. With both normal observers and patients, there is evidence for two components to these effects based on both visual perceptual and motor-based responses. While the perceptual responses reflect factors such as the visual familiarity of the action-related information, the motor response component is determined by factors such as the alignment of the objects with the observer's effectors and not by the visual familiarity of the stimuli. In addition to this, we suggest that action relations between stimuli can be coded pre-attentively, in the absence of attention to the stimulus, and action relations cue perceptual and motor responses rapidly and automatically. At present, formal theories of visual attention are not set up to account for these action-related effects; we suggest ways that theories could be expected to enable action effects to be incorporated.

1. Introduction

Imagine sitting in a tea shop with friends, with several cups and a teapot on the table in front of you. As you are holding a conversation, one friend reaches across and pours tea into your cup but their aim is not correct and tea falls to the side of the cup. You rapidly turn to move the cup so that the tea now falls into it. An everyday example similar to this suggests that the action event—the friend pouring tea from the teapot to the cup—can be processed even when you are attending to elsewhere (looking at your friend as you hold your conversation) and that you are sensitive to (your attention can be cued to) the relations between the objects being used (whether the teapot is angled correctly in relation to the cup). Is there psychological evidence to support these assertions? Do we code the spatial relations between objects being used together in action? Do we do this even when we are not fully attending to the stimuli? Though there has been a large amount of work on understanding human attention over the past 50 years, the degree to which attention is determined by action information in the environment has received little consideration—rather theories have tended to stress how selection depends on the perceptual properties of stimuli, not on how the stimuli relate to action. In this paper, we will reconsider this, reviewing evidence for the role of action-based information on attention. We will argue that there is 'pre-attentive' coding of action events (i.e. that such events are coded even when they are not both attended) and that there are at least two components to these effects: (i) a perceptual response based on the familiarity of the action-related stimulus, and (ii) a motor-based simulation of the action. The results suggest that action-based

links to attention play an important role in attentional operations in real-world environments. We go on to discuss ways in which these effects may be taken into account within current theories of visual attention.

(a) Cueing attention by higher order properties of stimuli

The great majority of studies of human visual attention have used relatively simple stimuli and examined the effects of what we might think of as low-level properties of these stimuli—for example, their brightness and contrast relative to local neighbours, the presence of particular types of motion (e.g. looming versus receding) and whether the stimuli share local grouping properties, as set out originally by the Gestalt psychologists. Theories converge in suggesting that there are both bottom-up ‘drivers’ of attention to salient visual signals [1] and top-down factors that guide attention (e.g. expectancies for particular items, the task-based relevance of stimuli, information held in working memory; [2,3]). The experiments on these bottom-up and top-down factors have furnished us with considerable knowledge about attentional operations, but typically they are not concerned with whether higher order properties of stimuli (not reducible to differences in low-level features) can influence what we attend to and how such effects may come about. To give one example, Moores *et al.* [4] asked participants to search for a target object (e.g. a motorcycle) and presented as distractors objects that were associatively but visually related to the target (e.g. a motor cycle helmet). They found that attention (e.g. the initial saccade made to the display) often went to the associatively related distractor, suggesting that the semantic properties of the distractors could be coded and attracted attention if they had some match to the expected target. Given that the distractors were associatively and not visually related to targets in this study, it is difficult to attribute the effect to attentional guidance to low-level features.

Converging evidence comes from studies using event-related potentials (ERPs), which highlight how rapidly particular effects arise. Telling *et al.* [5] used the same paradigm as Moores *et al.* [4] and measured ERPs. They documented effects of semantic distractors on the so-called N2pc component. The N2pc is based on the difference in evoked activity over the hemisphere contralateral to a stimulus relative to the hemisphere ipsilateral to the stimulus. This difference arises about 200 ms after the onset of the stimulus and has been taken to reflect the ease of attentional selection for the item [6]. Telling and colleagues found that the N2pc varied according to whether the distractor fell in the same or the opposite visual field relative to the search target. This suggests that the presence of the associated distractor affected the ease of selecting the target. The data from both eye movements and ERPs converge in indicating that there are relatively fast-acting responses to the higher order properties of stimuli, including whether they are semantically linked to the targets we are looking for. Belke *et al.* [7] also reported that the ‘semantic guidance’ effect was equally large for displays for few [4] or many [8] items. This last result is interesting because it suggests that, at least up to the display sizes used, attention can be guided efficiently based on higher order associative relations between a distractor and the item being searched for. That is, there may not necessarily be strong limits on the number of objects over which these

relations can be computed. We will return to this point about the limitations on processing higher order properties of stimuli.

One way to conceive of the results reported by Moores *et al.* [4] is that they arise through top-down expectancies, which feed back to modulate earlier processes. For example, the expectancy for a motorbike might activate the representations of all related items, so that the visual system is ‘primed’ to respond to these related items. This may help to overcome processing limitations as the number of items in the field increases. However, not all effects of higher order factors can be attributed to such expectancies. Rappaport *et al.* [8] asked participants to search for a target object (e.g. corn) that appeared in a complex display with several other objects similar in shape (aubergines and carrots) and carrying different colours (yellow, purple and orange). The target could be in its correct (learned) colour or in a colour that was not strongly linked to the object (yellow versus purple corn). These colours had no differential influence on search when they fell in geometric shapes without learned colour associations. However, quite different results emerged when the objects were assigned the same colours. Rappaport *et al.* found that the target object was efficiently selected when it was in its learned colour (yellow corn), while selection was quite inefficient when the target did not appear in its learned colour (purple corn). The authors then varied the probability with which correctly and incorrectly coloured targets were present. The advantage in selecting correctly coloured targets occurred even when these items appeared in only a low proportion of trials and when the target was more likely to appear in an incorrect colour. Measures were also taken of participants’ eye movements as they searched for the targets. Of interest here were the eye movements on trials when the target was absent, since where participants looked on these trials is informative about which target they were expecting. Interestingly, on these target-absent trials, eye movements were directed to distractors with the predicted (incorrect) target colour (e.g. purple distractors) rather than the learned but low probability target colour (yellow distractors). This provides strong evidence for expectations about the likely target colour (purple) directing search. However, when a target was present, attention was still most efficiently directed to the target carrying its learned colour (yellow corn ‘popped out’; purple corn did not). In this case, the learned properties of the target stimuli determined the ease of selection, even though this went against the top-down expectation being held by observers. Such results dissociate top-down guidance (based on the probability of a given target occurring) from bottom-up guidance of attention (based on the learned relationship between the colour and the shape of the target). The data suggest that learning itself can ‘tune’ bottom-up processing, so that attention is drawn to the target carrying the appropriate learned properties. As with the study of Belke *et al.* [7] noted earlier, search for a target carrying the learned properties was not strongly modulated by the numbers of distractors present.

2. Action and attention

This evidence indicates both that there is learning of associative relations between the features of objects (yellow and corn) and between independent objects (motorbike and crash

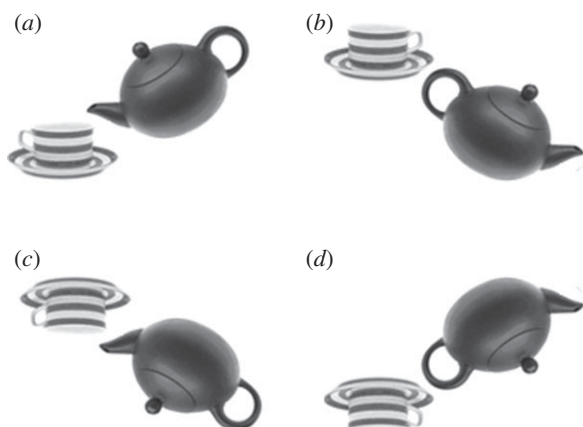


Figure 1. Stimuli used by Riddoch *et al.* [12]. (*a,b*) Upright objects (*a*) objects positioned for action, (*b*) objects not positioned for action). (*c,d*) Inverted objects (*c*) derived from action-related objects; (*d*) derived from action-unrelated objects).

helmet), and that these effects of associative learning can modulate attention. This point is important when we think about how action might modulate attention. Consider again the everyday example we started with. Several objects may be on the table—a spoon, sugar, a pen, a pair of glasses, etc. We [9] have previously argued that there is not an equal likelihood of forming associative relations between these stimuli, partly because the objects will not co-occur equally often but also because objects that are used in a joint action will have a ‘special relationship’. In particular, objects used in a joint action will be part of a common event, and their representations will be more active than objects that occupy background positions during the action event. Based on a process such as Hebbian learning [10], an associative link will be formed between the active object representations, so that in future, activation of one representation will lead to co-activation of the other. This should in turn increase the chances of both objects being selected together. As we review next, there is evidence consistent with these proposals.

(a) Action relations and visual selection

We first examined the issue of whether interacting objects are selected together in the context of visual extinction in neuropsychological patients. In the phenomenon of extinction, patients can fail to note stimuli presented on the side of space contralesional to their lesion when an ipsilesional item is presented simultaneously. On the other hand, the same contralesional stimulus can be detected when presented alone, suggesting there is a limit in attention (when stimuli compete for selection) rather than perception (assessed in the unilateral condition, when a competing ipsilesional item is not present). We [11] presented extinction patients with stimuli that were depicted interacting together (e.g. a teapot pouring into a cup) or in non-interacting positions (e.g. where the teapot faced the wrong way for any interaction; figure 1). When the objects were non-interacting, the patients showed strong extinction—typically reporting only the item on the ipsilesional side. However, when the objects were depicted interacting with other (figure 1*a*), then extinction reduced and the patients were often able to detect and even identify contralesional as well as ipsilesional items. That is, the presence of the interacting objects reduced extinction. We argued that

interacting objects are grouped as a single ‘action event’ and selected together rather than competing for attention.

This effect of action events was not because interacting objects have stronger local ‘Gestalt-type’ relations between their parts. First, note that the objects were not connected and they were not positioned to align their edges or to produce any other local Gestalt property. Second, in an extension of our original study, patients were tested either with upright pairs of interacting or non-interacting objects (as in [11]) or they were shown the same objects but inverted (figure 1*c,d*). In this case, any local Gestalt relation between the elements should be as strong when the objects were inverted as when they were upright. As before, upright, interacting objects led to recovery from extinction. By contrast, the benefit from the interaction was eliminated when objects were inverted—then extinction occurred irrespective of whether the objects were or were not in interacting co-locations [13]. These data indicate that the ability of the patient to attend to both of the objects present depended on the objects falling in familiar (upright) co-locations. Riddoch *et al.* [13] proposed that the perception of action relations between objects was similar to the perception of configural relations between the features of faces, being stronger when objects are viewed in their normal orientations (upright) [14]. Notably, any learned visual association between the stimuli would be formed when the objects are seen in upright orientations appropriate to object use. It was also found that the effects of action relation were contingent on the objects having the appropriate relative sizes, again consistent with how the objects appear when used together in the world.

There are several pieces of converging evidence from normal participants that match these neuropsychological data. For example, Green & Hummel [15] have reported consistent data with healthy young participants, who were better at confirming that a target picture matched a written label (e.g. ‘glass’) when a ‘distractor’ object (a jug) was positioned to interact with the target. This occurred when there was a short (50 ms) interval between the target and distractor, but not with a longer interval (200 ms), suggesting that the benefit arose from perceptual integration of the two stimuli. Similarly, Yoon *et al.* [16] reported that normal participants were better able to discriminate if two objects would be used together when the objects were placed in their familiar positions in relation to the actor (e.g. aligned with the hands the participant would normally use to act on the objects—teapot on left, cup on right—rather than vice versa). This occurred even though hand alignment was irrelevant to the decision.

One interesting aspect of these results on recovery from extinction is that the patients can be unaware of the presence of the contralesional stimulus unless it groups with its action-related partner. This in turn indicates that the action relation is coded even though the patient has reduced attention to the contralesional side. That is, action-related coding is not subject to strong attentional limitations in such cases. This argument for action relations being coded without awareness is also supported by the errors such patients can make. Riddoch *et al.* [11] (see also [17]) evaluated performance on trials when extinction did occur and patients reported only one of the two objects present. When the objects were non-interacting, patients typically reported the object that fell on the ipsilesional side—the standard extinction result. However, a strikingly different pattern occurred when the objects were shown interacting together, when report was affected

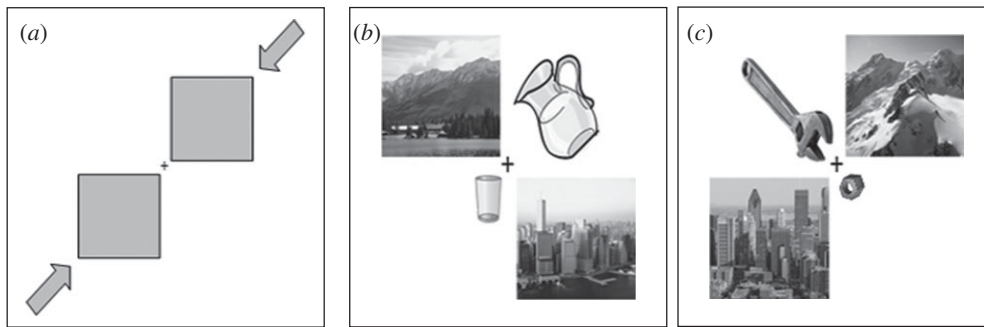


Figure 2. Stimuli used by Roberts & Humphreys [19]. Participants were cued at the start of each trial as to which pair of stimuli to attend to and they either made judgements about the pair of objects (this would be the case for display (b) following display (a)) or the scenes (this would happen for display (c) following display (a)).

by which object was ‘active’ and which was ‘passive’ in the action. The ‘active’ object is the stimulus that would be used to perform the action upon its passive partner—the teapot being ‘active’ and the cup ‘passive’ in the example shown in figure 1. With interacting objects, patients tended to report the active object rather than the passive one, and this was found even when the active object fell in the contralesional field. Note that report of the contralesional item and extinction of the ipsilesional stimulus is *opposite* to the standard pattern of extinction (ipsilesional report > contralesional report). Thus, when a contralesional teapot was shown pouring into an ipsilesional cup, patients might just report the teapot with the cup being extinguished. However, if the teapot was not positioned to interact with the cup, then the patient would typically report an ipsilesional cup with the contralesional teapot being extinguished. This reversal of the standard finding would occur if (i) the action relation between the stimuli is coded even though the patient is unaware of both items and (ii) attention is drawn first to the active member of the pair, when the action relationship is coded. This drawing of attention to the active member of an interacting pair of objects seems to be a bottom-up response to the image since it then leads to extinction of the ipsilesional item (at least on trials when both items are not recovered). The effect of action relations even on these error trials fits also with the argument that action relations are coded prior to the attentional limits of the patient affecting on performance (i.e. coding is pre-attentive).

Converging evidence for attention being ‘pushed’ to the active object, when objects are interacting together, comes from the work by Roberts & Humphreys [18]. These authors had normal participants make judgements about the temporal order in which two objects were presented, with the stimuli being presented either in interacting or non-interacting pairs. Roberts & Humphreys found that there was a bias to report the active object as appearing first, but only when the objects appeared to interact. This fits with the active object being attended earlier than the passive member of an interacting pair, and so the active object gains ‘prior entry’ to temporal order judgements. This bias to the active member of the pair disappeared when the objects were not positioned for action, consistent with the action relation being coded prior to attention going to the active member of the pair.

3. Visual and motor components

One important conclusion emerging from current work (reviewed below) is that the effects on attention of action

information in images stem from at least two sources—from a visual response (located in visually sensitive brain regions) and a motor response (located in brain regions concerned with planning motor actions).

Evidence for *both* visual and motor components of the action effects comes from several sources. Roberts & Humphreys [19] conducted an fMRI study in which they presented observers with interacting and non-interacting objects. Four stimuli appeared on a trial, one in each quadrant of the display. In one diagonal pair of quadrants, interacting or non-interacting objects were presented, whereas in the other, a pair of scenes appeared. Participants were asked either to make judgements to the objects (the task was to judge whether the objects were related or not) or to the scenes (the task was to judge whether the scenes were both indoor), with the relevant diagonal being cued on each trial (figure 2). Note that when participants make judgements about the scenes, minimal attention should be paid to the objects. Roberts & Humphreys found that there was increased activation to interacting relative to non-interacting objects, and this was observed in brain regions typically thought to reflect visual responses to object shape (activity was found in the lateral occipital complex (LOC) and the anterior fusiform gyrus, particularly in the left hemisphere—regions associated with shape coding). A further striking finding was that this enhanced response to interacting objects occurred even when participants attended and made responses to the scenes; that is, the response occurred even when minimal attention was allocated to the objects. Similar findings, of enhanced responses in the LOC to interacting objects, have been reported by Kim & Biederman [20]. In addition, Kim *et al.* [21] used the procedure of Green & Hummel [15] (see §2a) while transcranial magnetic stimulation (TMS) was applied to the LOC and to the posterior parietal cortex (PPC), a brain region implicated in the allocation of spatial attention (cf. [22]). As with Green & Hummel [15], there was facilitated responding to interacting objects after TMS to the PPC, but this benefit was eliminated after TMS to the LOC. This last result points to a causal role of the LOC in responding to pairs of interacting objects.

These data are highly consistent with the neuropsychological findings on extinction. First, the results of Roberts & Humphreys [19] show that the differential neural response to pairs of interacting objects is unaffected by whether or not participants attend to the stimuli. Second, the data indicate that interfering with brain activity in the PPC (using TMS) does not disrupt the beneficial effects of interacting objects [21]. Extinction in patients is associated with damage to the PPC [23] and such patients remain sensitive to the effects of action relations between objects [11]. Third,

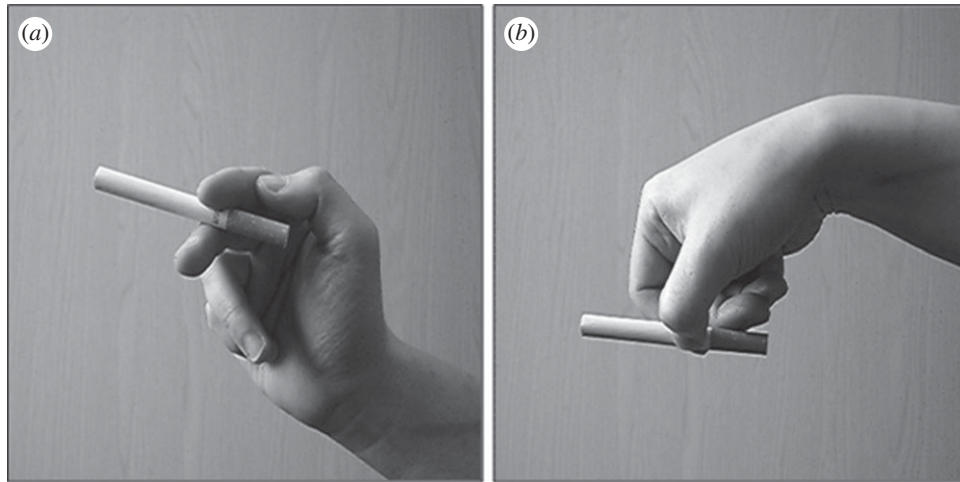


Figure 3. Examples of objects subject to a congruent grip (a) or an incongruent grip (b). The incongruent grips were selected from congruent grips to other objects (here a congruent grip to a pin).

the evidence for sensitivity in the LOC to action relations fits with this region typically being spared in patients showing extinction and with the notion that residual visual processing in extinction reflects activity within spared regions of ventral cortex [24]. We propose that, in patients, there is recovery of extinction due to mutual reinforcement of activation between interacting objects from the associative links that have developed within ventral visual regions when objects are seen being used together. This mutual reinforcement enhances the activation of the contralesional stimulus, enabling it to be reported. This enhanced ventral activation comprises the *visual component* of the effects of action relations.

Alongside this, there is also evidence for a *motor* component of our heightened response to action-related stimuli. We can think of this as a form of simulation of the action, perhaps stemming from the so-called canonical neurons (neurons associated with visuo-motor transformations of objects) that are activated when a hand shapes to grasp an object [25]. On the other hand, the ‘classic’ mirror neuron system appears not to be sensitive to how objects are grasped [26].

Neuropsychological evidence for this second, motor component comes from a study by Humphreys *et al.* [27]. They tested the effects of action relations on extinction and varied whether the objects appeared either from the patient’s own view or from a third-person perspective, and whether the position of the active object aligned with the hand that the patient would have used for the action (e.g. whether the jug shown in figure 2 is on the left or the right). They found that there was an overall advantage for interacting relative to non-interacting objects, both when the objects were depicted from the patient’s own view and from a third-person perspective. However, recovery from extinction was particularly strong when the objects appeared from the patient’s own view and the active object aligned with the hand the patient would typically use for the action. Humphreys *et al.* proposed that interacting objects generate a stronger visual response both in the first and the third-person perspectives, but, on top of this, there is also a motor response to objects seen from the patient’s own perspective and aligned with the hands used for action. This motor response could couple with the visual response to the objects to enhance the representation of an item on the contralesional side, so reducing the effects of extinction.

The study of Yoon *et al.* [16] provides converging evidence. As we described in §2a, Yoon *et al.* found that judgements

about whether objects are used together were enhanced when the objects aligned with the hands the observer would use in the action. This effect of hand alignment was stronger when objects were depicted in the participant’s own reference frame and it reduced when the objects were aligned to an actor seen from across a table.

(a) Effects of hand grip

This argument for both visual and motor components of the response to action information gains additional support from studies where action information is manipulated by presenting hands in relation to objects. Borghi *et al.* [28] had participants categorize stimuli as manipulable artefacts versus natural objects, and preceded the stimuli by photographs of hand postures with a power or precision grip. Participants were faster to respond to objects which could be grasped by a precision grip when the prime was a precision grip hand posture, and vice versa for objects that would be grasped with a power grip. The results are consistent with responses to objects being primed by the pre-activation of a motor response triggered by the hand grasp. Yoon & Humphreys [29] used stimuli where hands were depicted grasping objects, with the object and the hand grip either being congruent or incongruent (figure 3). The incongruent grips were selected from congruent grips to other objects, so that the congruent and incongruent grips themselves were equally familiar. The targets of the grips were either real objects (the cigarette, in figure 3) or non-objects created by combining the parts of two real objects. When non-objects were presented, the grips were congruent with the graspable part of the original parent object (e.g. the stem of the cigarette). The task was to decide whether the stimulus was an object or a non-object, and participants were informed that the grip was irrelevant. Despite the grip not being relevant, reaction times were strongly affected by the congruence of the grip; responses were faster for stimuli with correct relative to incorrect grips.

Using these same stimuli, we have explored the time course of this hand grip effect using ERPs. ERPs are useful because they not only provide information about the (coarse) localization of a response in the brain, but they also provide fine-grained analysis of the time when any event-related potential arises. Kumar *et al.* [30] measured ERPs when participants

performed object decision responses to objects and non-objects assigned congruent and incongruent grips. They found that the P1 component, typically taken to reflect an initial early response to a stimulus emerging after around 100 ms, was greater over motor cortex for objects shown with a congruent relative to an incongruent grip. This was followed by a stronger later component, the N1 (emerging around 150 ms), to congruently gripped objects (compared with incongruently gripped objects) over posterior brain sites. The magnitude of the increase in the two components (P1 anterior, N1 posterior) for congruently gripped objects relative to incongruently gripped objects was correlated across participants (i.e. participants showing a larger P1 enhancement tended also to have a larger N1 enhancement). The timings of these effects are of interest because they suggest that the effect of hand grip is registered first through a motor response to the stimulus, and this is followed by grip-modulation of a visual response. Furthermore, the correlation between the two effects is consistent with the motor response feeding back to modulate visual processing over posterior sites.

Along with the enhanced early P1 responses to congruently gripped objects over motor cortex, there was also evidence for a reverse effect in the P1 component over posterior brain sites—in this case, there was a greater posterior P1 response to incongruently gripped objects compared with objects assigned a congruent grip. Kumar *et al.* [30] attributed this response to incongruently gripped objects to their visual unfamiliarity. Essentially, incongruently gripped objects present the observer with conflicting information between the familiar object and the unfamiliar grip assigned to the object. The disparity between the normal linkage of the object and the grasp might require greater perceptual resources to be registered, leading to an enhanced perceptual (P1) response. The opposite direction of this effect to the P1 effect over motor cortex provides strong evidence for the two early effects being independent of one another.

The results showing a strong visual P1 response to incongruent stimuli are in line with other studies where the predictability of a sequence of actions has been manipulated. For example, Reid *et al.* [31] presented adults and infants with an action series in which an actor used an object either in a standard manner or in a non-standard way (e.g. a toothbrush might be brought to the mouth or to the forehead). They found a greater N400 response (peak amplitude about 500 ms after the stimulus) in both adults and nine-month-old infants when the object was used in a non-standard way. Our data, along with those of Reid and colleagues, indicate that incongruent stimuli can generate a strong visual response. Importantly, our results show that this visual response is separate from the motor-based simulation to stimuli congruent with the standard motor action. We suggest that this evoked motor response feeds back to enhance visual processing and, based on the results with visual extinction [27], this can enable stimuli to be selected together.

In addition to measuring the visual response to congruent and incongruently gripped objects, Kumar *et al.* [32] evaluated electroencephalogram (EEG)-based oscillatory activity in the mu frequency band (8–12 Hz). This is of interest because desynchronization of the mu rhythm over motor areas is linked to motor preparation [33–36] and has been observed in relation to both object-directed grasp responses [37] and precision grips [38]. Using stimuli such as those shown in figure 3, Kumar *et al.* [32] found that there was greater

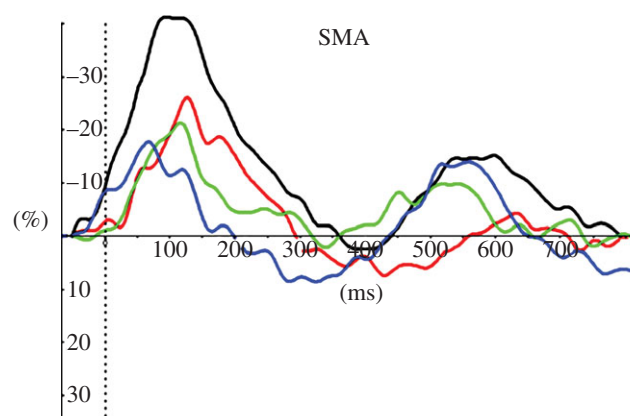


Figure 4. De-synchronization power in the mu frequency. Black, objects with congruent grip; red, objects with incongruent grip; green, non-objects with congruent grip; blue, non-objects with incongruent grip. Data here are recorded from electrodes above the supplementary motor area (SMA).

desynchronization over motor cortex and other brain regions associated with motor preparation (e.g. the supplementary motor area, a region associated with motor planning) for congruently gripped objects compared with objects assigned an incongruent grip, particularly over the left hemisphere. Similar to the P1 effect over motor cortex, this effect occurred strikingly soon after the onset of the stimulus, with the peak of the desynchronized response occurring at around 100 ms (figure 4). Kumar *et al.* [32] also noted that the magnitude of mu rhythm desynchronization to congruent objects correlated with response times (the greater the desynchronization, the faster the response to congruently gripped stimuli), supporting the argument that this early motor-related response is related to behaviour. The lateralization of the effects over the left hemisphere also matches evidence for right-hand dominance for using objects. These results converge with the ERP data in indicating that there is a fast-acting motor response to an object assigned a congruent grip.

Several other investigators have also reported evidence for an early motor as well as visual response to action-related information in images. For example, Hoening *et al.* [39] found the largest negativity in the P1 time window over the motor cortex when participants made action rather than visual verification judgements about stimuli. Similarly, Kiefer *et al.* [40] found effects of action priming on responses to objects over motor areas in the P1 time window. Both these data and the data we report fit with the idea that action-related information is processed rapidly in pre-frontal, motor-related brain regions, as well as in more posterior (visually driven) areas, with the pre-frontal activity modulating the visual regions at a later time window.

4. Attentional orienting

In all of the above studies where objects are depicted with a correct or incorrect hand grip, the stimuli fell at the centre of the field at the focus of attention (FOA). Is there evidence that hand grip can affect how attention itself is directed? Recently, we [41] have extended this work by examining whether congruency between a grip and an object modulates visual orienting to stimuli that are initially not at the FOA. We cued participants with an action name (e.g. drink) followed by a bilateral display with images of objects assigned either

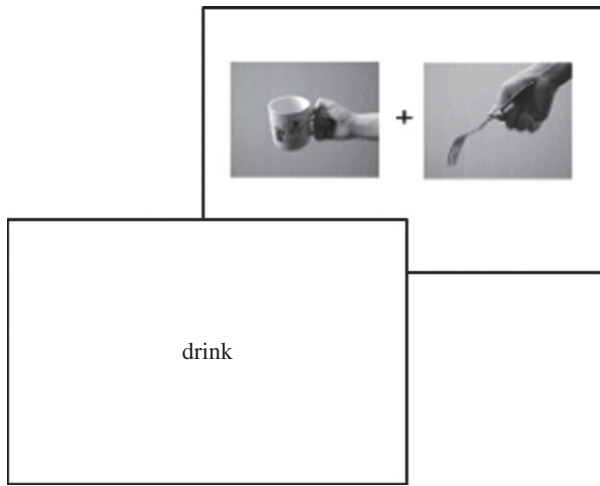


Figure 5. Stimuli used by Kumar *et al.* [41]. Participants saw a verb and then had to verify if an object associated with the verb was present. In our example, the target was present and gripped correctly, whereas the distractor was gripped incorrectly (TC-DIC).

a correct or an incorrect grip. A 2×2 design was employed in which targets were gripped correctly or incorrectly and paired with a distractor which itself was gripped correctly or incorrectly (figure 5). We measured the N2pc as our index of the ease of selecting and orienting attention to a target (see above; also [42]).

The magnitude of the N2pc over the PPC is shown in figure 6. The results show that the N2pc was reduced in amplitude and took longer to reach its peak when the target was correctly gripped and the distractor incorrectly gripped (low red line, figure 6, condition TC-DIC), compared with the other conditions (both target and distractor correctly gripped (TC-DC); both target and distractor incorrectly gripped (TIC-DIC); target gripped incorrectly and distractor gripped correctly (TIC-DC)). This result illustrates at least three points. One is that, matching our prior results [29,32], hand grip is selected along with an object that is cued for the task (otherwise effects of hand grip would not have been apparent). Second, a novel result is that attentional orienting to a target, indexed through the N2pc, is itself affected by how the object is grasped—orienting is easier when the target has a congruent hand grip. Third, orienting was also influenced by the congruency of the grip applied to the distractor (even when the target was assigned a congruent grip, the N2pc decreased when the distractor had an incongruent grip). This last result again provides evidence that action-related properties of objects are processed even when the objects are not attended.

We can think of the effect of grip congruence in either of two ways. Mirroring the arguments we made about the effects of associative distractors on visual search [4,5,7], we can propose effects based on bottom-up stimulus coding or top-down priming. On the bottom-up view, congruently gripped targets generate a strong motor and then a visual response, which drives attention to their location. By contrast, incongruently gripped distractors generated only a weak competing response. The net effect is that target selection is facilitated. Alternatively, it may be that in the experiments on orienting [41] the cue word (e.g. drink) primes associated visual and motor responses, and that these primed responses then enhance stimulus processing and help direct attention to the cued target. One reason to favour a bottom-up account

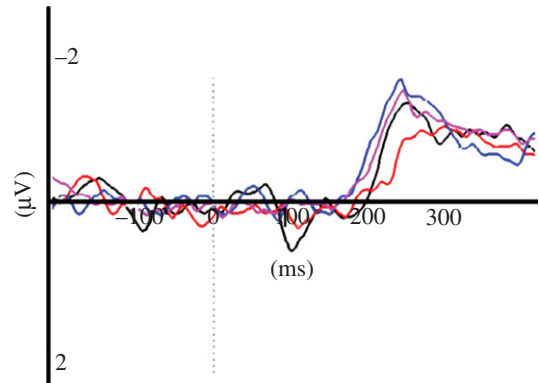


Figure 6. The N2pc responses recorded by Kumar *et al.* [41]. Red, target object correctly gripped and distractor incorrectly gripped (TC-DIC); black, target object correctly gripped and distractor gripped correctly (TC-DC); purple, target object incorrectly gripped and distractor gripped incorrectly (TIC-DIC); blue, target object gripped incorrectly and distractor gripped correctly (TIC-DC).

here is that there were effects of the grip applied to the distractor, yet the distractor was never cued. However, further work is required to tease these possibilities apart.

(a) Motor preparation

Although the work on attentional orienting does not fully distinguish bottom-up and top-down accounts of the effects of grasp-action on performance, there is other work indicating that a pre-prepared motor response can affect visual attention in a top-down manner. We again start with a neuropsychological example. Humphreys & Riddoch [42] reported evidence from unilateral neglect, where patients show an attentional bias against contralesional space and may be unaware of stimuli falling there. They examined cases where the patient was cued to find an object by its name ('find the cup') and when the cue was action-related ('find the object to drink from'). Neglect was less apparent when the object was cued by its associated action compared with when it was cued by its name. The result suggests that neglect was reduced when the patient could respond to the action-related properties of the object compared with when the semantic properties of the object were more strongly 'weighted' (from the object's name). This enhanced response to action properties could also have been facilitated by the patient preparing to make the appropriate motor action from the cue ('drink').

More specific evidence for effects of motor preparation has been reported by Bekkering & Neggers [43] and Forti & Humphreys [44] with normal viewers. Bekkering & Neggers tracked eye movements when participants searched for targets defined by a conjunction of orientation and colour features (green horizontal target among red horizontal and green vertical distractors). Participants either made pointing or grasp responses to targets. When a pointing response was made, the first saccade often went to a distractor carrying the same colour as the cued target (green vertical). However, when the task was to grasp the target, then first saccades went more frequently to a distractor matching the orientation of the target (red horizontal). Bekkering & Neggers proposed that preparing a particular action increased the attentional weighting given to action-related properties of the stimulus (e.g. preparing a grasp as opposed to a pointing response increases the weight given to stimulus orientation rather

than colour), so that these properties played a stronger role in attentional selection (guiding the first saccade to a distractor with the cued orientation rather than colour).

Forti & Humphreys [44] contrasted search to an object name (cup) and search to the action that would be performed with the object (drink). They found that cueing by action facilitated the detection of targets in the lower visual field. The lower visual field has a greater representation than the upper visual field within the parietal cortex [45], perhaps due to visually directed actions being more frequent in the lower than the upper visual field. Forti & Humphreys argued that cueing the action for an object enhanced perceptual selection by priming regions of parietal cortex sensitive to visuo-motor response mapping.

Vainio *et al.* [46] reported an even more dramatic finding. They had participants prepare a motor response (precision or power grip) that had to be carried out when participants detected a change to an object (using a procedure in which the objects flashed on and off, and one (target) object was substituted by another in alternative displays). Participants were much more likely to detect a change, and any changes were detected much more rapidly, when the object was congruent with the prepared grip (e.g. preparing a power grip enhanced the detection of change to a large object). If an inappropriate grip was prepared (e.g. a precision grip when a change was made to a large object), then often the change would go unnoticed. These data indicate that the preparation of a particular motor action can even determine whether people 'see' a particular visual event taking place. The result fits with the data on neglect [42], where patients may not 'see' objects unless the objects are congruent with a prepared motor response.

Here, we can think that preparation of a motor action prior to the occurrence of a stimulus provides a particularly important input into the visual selection process, perhaps exerting a greater effect even than any motor response evoked by the stimulus itself as it is processed.

5. Theoretical implications

Most theories of attention have stressed that visual selection is based on low-level properties of stimuli—their orientation, colour and so forth (see [47–49] for examples). However, the data on the effects of action information on selection indicate that selection can be determined by relatively high-level representations, sensitive to the spatial inter-relations between interacting objects, and also to motor simulation and motor preparation effects and their relations to objects. The effects of action preparation can be thought of in terms of priming of the motor system and top-down feedback to bias visual processes in terms of stimuli consistent with the prepared action (cf. [50]). The effects of action relations in the image (between separate objects and between a hand grasp and an object), however, suggest that visual elements may be coded in a relatively elaborate fashion even without full attention being allocated to the stimuli and, further, this information may be relayed rapidly to motor and pre-motor regions concerned with motor simulation (see [30,32]). Theories of visual selection need to be adapted to enable such higher order representations to modulate attention. This is feasible in frameworks such as the Selective Attention for Identification Model (SAIM) put forward by Heinke &

Humphreys [48]. In this model, visual stimuli compete for selection into a 'FOA'. The FOA is a bottleneck through which stimuli of different spatial extents are mapped, and representations within the FOA are then matched against 'attentional templates' for known objects. The activated templates then feed back excitation to favour matching stimuli. Within SAIM, the nature of the templates can vary, and the results we have summarized here are consistent with templates based on learned action relations between objects. Partial activation of such templates could be passed on rapidly to motor cortex to drive fast motor preparation responses, and the templates could also feed back to earlier visual processes, to enable objects within a common action-template to be selected together. The notion of rapid activation of pre-frontal brain regions which create a 'hypothesis' about a stimulus, which then modulates earlier processing, is similar to proposals put forward by Bar and co-workers [51,52]. However in our case, we suggest that this form of re-entrant processing is based on early activation of a motor response to the stimulus, rather than other types of 'perceptual hypotheses' (e.g. about the semantic properties of the stimulus).

A final caveat is worth noting, however. If there is evidence for affordances being coded pre-attentively (e.g. in patients showing visual extinction) and then affecting the allocation of attention, then does this mean that we are constantly registering all the affordances that may be offered in complex environments? Here, we must remember that the experiments we have presented have used restricted, sparsely populated displays, rather than the more complex environments we encounter in everyday life. Patients may be presented with two rather than one object. Search is tracked across perhaps a dozen objects in a sparse display (see [44]), but not across a complex and cluttered scene such as in a 'where's Wally?' puzzle. It may be that, with sparse displays, we have sufficient perceptual capacity to code action relations without attention. It remains a moot point, though, as to whether multiple affordances are computed in more complex environments or whether there is a limitation on the number of affordance cues we can compute without attention. The data show that pre-attentive coding of affordances is possible, but not that it is characteristic of visual perception and attention in all contexts (see [53] for further discussion of this point).

6. Conclusions: attention and affordance

The results discussed here indicate that action relations between objects, and between a hand depicted in interaction with an object, affect visual attention. In addition, visual attention is also modulated by the preparation of a particular motor response. When stimuli are shown interacting, there is evidence that attention is drawn to both items, facilitating the selection of both stimuli in patients with visual extinction. The corollary of this is that a member of an interacting pair (e.g. the hand grasping an object) tends to be selected even when it is irrelevant for the task. The response to interacting objects seems to be mediated at both visual and motor levels of processing, and the motor 'resonance' to the interaction may feed back to modulate the visual response and to affect the ease of orienting attention to an object in the first place. This effect of an evoked motor response is also apparent when people prepare a particular action to make for a response, with the evidence suggesting that preparation of a particular motor

response feeds back to enhance the processing of visual features compatible with the prepared response.

Gibson [54] coined the term 'affordance' to capture the idea that visual objects may trigger a direct visuo-motor response according to the behavioural context the observer is in. For example, a log may 'afford' a grasp response when we are thinking to make a fire, while a sitting response may be afforded when we are tired and needing to rest. We suggest that visual affordances are psychologically real and moderate both visual selection of attended stimuli and also how attractive stimuli are for attention. Furthermore, the context in which selection takes place is also determined by the preparation of a particular

action (e.g. preparing to sit will favour the affordance of the log as a seat). We speculate that visual affordances are learned through a process such as Hebbian learning based on the statistical regularities present when action events occur. These statistical regularities reflect the possibility for action based on the physical properties of the objects [9]. We contend that affordances become part of the learned repertoire of attentional responses to our environment.

Funding statement. This work was supported by grants from the European Research Council (PePe – 323883), the NIHR (Oxford Cognitive Health Clinical Research Facility) and the Stroke Association.

References

- Theeuwes J. 1992 Perceptual selectivity for color and form. *Percept. Psychophys.* **13**, 27–36.
- Downing PE. 2000 Interactions between visual working memory and selective attention. *Psychol. Sci.* **11**, 467–473. (doi:10.1111/1467-9280.00290)
- Folk CL, Remington RW, Johnston JC. 1992 Involuntary covert orienting is contingent on attentional control settings. *J. Exp. Psychol.: Hum. Percept. Perform.* **18**, 1030–1044. (doi:10.1037/0096-1523.18.4.1030)
- Moore E, Laiti L, Chelazzi L. 2003 Associative knowledge controls deployment of visual attention. *Nat. Neurosci.* **6**, 182–189. (doi:10.1038/nn996)
- Telling A, Kumar S, Meyer AS, Humphreys GW. 2010 Electrophysiological evidence of semantic interference in visual search. *J. Cogn. Neurosci.* **10**, 2212–2225. (doi:10.1162/jocn.2009.21348)
- Luck SJ, Hillyard SA. 1990 Electrophysiological evidence for parallel and serial processing during visual-search. *Percept. Psychophys.* **48**, 603–617. (doi:10.3758/BF03211606)
- Belke E, Humphreys GW, Watson DG, Meyer AS, Telling AL. 2008 Top-down effects of semantic knowledge in visual search are modulated by cognitive but not perceptual load. *Percept. Psychophys.* **70**, 1444–1458. (doi:10.3758/PP.70.8.1444)
- Rappaport SJ, Humphreys GW, Riddoch MJ. 2013 The attraction of yellow corn: reduced attentional constraints on coding learned conjunctive relations. *J. Exp. Psychol. Hum. Percept. Perform.* **39**, 1016–1031. (doi:10.1037/a0032506)
- Humphreys GW, Riddoch MJ. 2007 How to define an object? *Mind Lang.* **22**, 534–547. (doi:10.1111/j.1468-0017.2007.00319.x)
- Hebb DO. 1949 *The organization of behaviour*. New York, NY: John Wiley & Sons.
- Riddoch MJ, Humphreys GW, Edwards S, Baker T, Willson K. 2003 Seeing the action: neuropsychological evidence for action-based effects on object selection. *Nat. Neurosci.* **6**, 82–89. (doi:10.1038/nn984)
- Riddoch MJ, Chechlacz M, Mevorach C, Mavritsaki E, Allen HA, Humphreys GW. 2010 The neural mechanisms of visual selection: the view from neuropsychology. *Ann. NY Acad. Sci.* **1191**, 156–181. (doi:10.1111/j.1749-6632.2010.05448.x)
- Riddoch MJ, Pippard P, Booth L, Rickell J, Summer J, Brownson A, Humphreys GW. 2011 Effects of action relations on the configural coding between objects. *J. Exp. Psychol.: Hum. Percept. Perform.* **37**, 580–587. (doi:10.1037/a0020745)
- Rossion B. 2013 The composite face illusion: a whole window into our understanding of holistic face perception. *Vis. Cogn.* **29**, 1–115.
- Green C, Hummel JE. 2006 Familiar interacting object pairs are perceptually grouped. *J. Exp. Psychol.: Human Percept. Perform.* **32**, 1107–1119. (doi:10.1037/0096-1523.32.5.1107)
- Yoon EY, Riddoch MJ, Humphreys GW. 2010 The paired object affordance effect. *J. Exp. Psychol.: Hum. Percept. Perform.* **36**, 812–824. (doi:10.1037/a0017175)
- Wulff M, Humphreys GW. 2013 Visual responses to action between unfamiliar object pairs modulate extinction. *Neuropsychologia* **51**, 622–632. (doi:10.1016/j.neuropsychologia.2013.01.004)
- Roberts KL, Humphreys GW. 2010 The one that does leads: action relations alter the perceived temporal order of graspable objects. *J. Exp. Psychol.: Hum. Percept. Perform.* **36**, 776–780. (doi:10.1037/a0018739)
- Roberts KL, Humphreys GW. 2010 Action relationships concatenate representations of separate objects in the ventral visual system. *Neuroimage* **42**, 1541–1548. (doi:10.1016/j.neuroimage.2010.05.044)
- Kim JG, Biederman I. 2011 Where do objects become scenes? *Cereb. Cortex* **21**, 1738–1746. (doi:10.1093/cercor/bhq240)
- Kim JG, Biederman I, Juan JG. 2011 The benefit of object interactions arises in the lateral occipital cortex independent of attentional modulation from the intraparietal sulcus: a transcranial magnetic stimulation study. *J. Neurosci.* **33**, 7846–7855.
- Corbetta M, Shulman GL. 2002 Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215. (doi:10.1038/nrn755)
- Chechlacz M, Rotshstein P, Hansen PC, Deb S, Riddoch MJ, Humphreys GW. 2013 The central role of the temporo-parietal junction and the superior longitudinal fasciculus in supporting multi-item competition: evidence from lesion-symptom mapping of extinction. *Cortex* **49**, 487–506. (doi:10.1016/j.cortex.2011.11.008)
- Vuilleumier P, Schwartz S, Verdon V, Maravita A, Hutton C, Husain M, Driver J. 2008 Abnormal attentional modulation of retinotopic cortex in parietal patients with spatial neglect. *Curr. Biol.* **18**, 1525–1529. (doi:10.1016/j.cub.2008.08.072)
- Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F, Rizzolatti G. 2005 Parietal lobe: from action organization to intention understanding. *Science* **308**, 662–667. (doi:10.1126/science.1106138)
- Johnson-Frey SH, Newman-Norlund R, Grafton ST. 2005 distributed left hemisphere network active during planning of everyday tool use skills. *Cereb. Cortex* **15**, 681–695. (doi:10.1093/cercor/bhh169)
- Humphreys GW, Wulff M, Yoon EY, Riddoch MJ. 2010 Effects of action relation and affordance that do and do not depend on a self-reference frame: neuropsychological evidence. *J. Exp. Psychol.: Learn. Mem. Cogn.* **36**, 659–670. (doi:10.1037/a0019317)
- Borghini AM, Bonfiglioli C, Ricciarelli P, Rubichi S, Nicoletti R. 2005 Visual hand primes and manipulable objects. In *COGSCI2005. Proc. 27th Annual Conf. of the Cognitive Science Society, Stresa, Italy, 21–23 July* (eds B Bara, L Barsalou, B Bucciarelli), pp. 322–327. Mahwah, NJ: Lawrence Erlbaum Associates.
- Yoon EY, Humphreys GW. 2005 Direct and indirect effects of action on object classification. *Mem. Cogn.* **33**, 1131–1146. (doi:10.3758/BF03193218)
- Kumar SJ, Yoon EY, Humphreys GW. 2012 Perceptual and motor-based responses to hand actions on objects: evidence from ERPs. *Exp. Brain Res.* **220**, 153–164. (doi:10.1007/s00221-012-3126-4)
- Reid VM, Hoehl S, Grigutsch M, Groendahl A, Parise E, Striano T. 2009 The neural correlates of infant and adult goal prediction: evidence for semantic processing systems. *Dev. Psychol.* **45**, 620–629. (doi:10.1037/a0015209)
- Kumar SJ, Riddoch MJ, Humphreys GW. 2013 Mu rhythm desynchronization reveals motoric influences of hand action on object recognition. *Front. Hum. Neurosci.* **7**, 1–9. (doi:10.3389/fnhum.2013.00066)
- Pfurtscheller G, Neuper C, Krausz G. 2000 Functional dissociation of lower and upper frequency mu rhythms in relation to voluntary limb movement.

- Clin. Neurophysiol.* **111**, 1873–1879. (doi:10.1016/S1388-2457(00)00428-4)
34. Pfurtscheller G, Stancak Jr A, Neuper C. 1997 Event-related synchronization (ERS) in the alpha band—an electrophysiological correlate of cortical idling: a review. *Int. J. Psychophysiol.* **24**, 39–46. (doi:10.1016/S0167-8760(96)00066-9)
 35. Derambure P, Defebvre L, Bourriez JL, Cassim F, Giueu JD. 1999 Event-related desynchronization and synchronization. Reactivity of electrocortical rhythms in relation to the planning and execution of voluntary movement. *Neurophysiol. Clin.* **29**, 53–70. (doi:10.1016/S0987-7053(99)80041-0)
 36. Pineda J. 2005 The functional significance of mu rhythms: translating ‘seeing’ and ‘hearing’ into ‘doing’. *Brain Res. Rev.* **50**, 57–68. (doi:10.1016/j.brainresrev.2005.04.005)
 37. Muthukumaraswamy SD, Johnson BW. 2004 Changes in rolandic mu rhythm during observation of a precision grip. *Psychophysiology* **41**, 52–156. (doi:10.1046/j.1469-8986.2003.00129.x)
 38. Muthukumaraswamy SD, Johnson BW, McNair NA. 2004 Mu rhythm modulation during observation of an object-directed grasp. *Brain Res. Cogn. Brain Res.* **19**, 195–201. (doi:10.1016/j.cogbrainres.2003.12.001)
 39. Hoenig K, Sim EJ, Bochev V, Herrnberger B, Kiefer M. 2000 Conceptual flexibility in the human brain: dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *J. Cogn. Neurosci.* **20**, 1799–1814. (doi:10.1162/jocn.2008.20123)
 40. Kiefer M, Sim EJ, Helbig H, Graf M. 2011 Tracking the time course of action priming on object recognition: evidence for fast and slow influences of action on perception. *J. Cogn. Neurosci.* **23**, 1864–1874. (doi:10.1162/jocn.2010.21543)
 41. Kumar SJ, Riddoch MJ, Humphreys GW. Submitted. Action information modulates attentional orienting: an ERP study. *Exp. Brain Res.*
 42. Humphreys GW, Riddoch MJ. 2001 Detection by action: evidence for affordances in search in neglect. *Nat. Neurosci.* **4**, 84–88. (doi:10.1038/82940)
 43. Bekkering H, Nèggers SFW. 2002 Visual search is modulated by action intentions. *Psychol. Sci.* **13**, 370–374. (doi:10.1111/j.0956-7976.2002.00466.x)
 44. Forti S, Humphreys GW. 2008 Lower visual-field sensitivity to object viewpoint and action instruction in search. *Psychol. Sci.* **19**, 42–48. (doi:10.1111/j.1467-9280.2008.02044.x)
 45. Previc FH. 1990 Functional specialization in the lower and upper visual fields in humans: its ecological origins and neurophysiological implications. *Behav. Brain Sci.* **13**, 519–542. (doi:10.1017/S0140525X00080018)
 46. Vainio L, Symes E, Ellis R, Tucker M, Ottoboni G. 2008 On the relations between action planning, object identification, and motor representations of observed actions and objects. *Cognition* **108**, 444–465. (doi:10.1016/j.cognition.2008.03.007)
 47. Bundesen C. 1990 A theory of visual attention. *Psychol. Rev.* **97**, 523–547. (doi:10.1037/0033-295X.97.4.523)
 48. Heinke D, Humphreys GW. 2003 Attention, spatial representation and visual neglect: simulating emergent attentional processes in the Selective Attention for Identification Model (SAIM). *Psychol. Rev.* **110**, 29–87. (doi:10.1037/0033-295X.110.1.29)
 49. Wolfe JM. 1994 Guided search 2.0. A revised model of visual search. *Psychon. Bull. Rev.* **1**, 202–238. (doi:10.3758/BF03200774)
 50. Desimone R, Duncan J. 1995 Neural mechanisms of selective attention. *Annu. Rev. Neurosci.* **18**, 193–222. (doi:10.1146/annurev.ne.18.030195.001205)
 51. Bar M. 2003 A cortical mechanism for triggering top-down facilitation in visual object recognition. *J. Cogn. Neurosci.* **15**, 600–609. (doi:10.1162/089892903321662976)
 52. Kestutis Kveraga K, Boshyan J, Bar M. 2007 Magnocellular projections as the trigger of top-down facilitation in recognition. *J. Neurosci.* **27**, 13 232–13 240. (doi:10.1523/JNEUROSCI.3481-07.2007)
 53. Gillebert C, Humphreys GW. In press. Mutual interplay between perceptual organization and attention: a neuropsychological perspective. In *Handbook of perceptual organization* (ed. J Wagemans). Oxford, UK: Oxford University Press.
 54. Gibson JJ. 1979 *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.