



Review

Embodied cognition and the perception–action link

Bruce Bridgeman*, Philip Tseng

Department of Psychology, University of California, Santa Cruz, Santa Cruz, CA 95064, USA

Accepted 8 January 2011

Available online 14 January 2011

Communicated by L. Perlovsky

Abstract

Perception is interpreted as a set of capabilities that facilitate two functions necessary for survival; learning about the environment and controlling real-time behavioral interactions with it. Perceptual capabilities evolve in the context of an organism and its environment, adapted to an organism's ecological niche. The relation between embodied perception and action can be studied in the context of the only muscles that serve only to enable perception – the eye muscles. The only eye movements under cognitive control are saccades, the rapid jumps of binocular fixation from one target to another. The world is perceived as stable while the retinal image, and the corresponding projections inside the brain, are displaced with each saccade. This space constancy forms the stable platform for all other visual functions and requires an explanation that involves visual short-term memory. This memory, and the change detection that it makes possible, is enhanced when there is a physical interaction between the observer and the visual stimulus. Perception is something you do, not something that happens to you.

© 2011 Elsevier B.V. All rights reserved.

Contents

1. Functions of perception	74
2. Evolution of perception	74
3. Control of action	75
4. Perception and eye movements – an example	75
5. Spatiotopic fusion and efference copy	76
6. Change blindness and scene representation	76
6.1. Why does change blindness occur?	78
6.2. Encoding	78
6.3. Maintenance	78
6.4. Retrieval and comparison	78
6.5. Report	78
6.6. The role of action in change detection	78
6.7. Two functions of vision	79
7. Embodied cognition and the two visual systems	79

* Corresponding author.

E-mail address: bruceb@ucsc.edu (B. Bridgeman).

8. Application to naturalistic images	81
8.1. Neurological evidence	82
8.2. Conclusion	82
References	82

1. Functions of perception

What is perception for? Animals and humans are exquisitely sensitive to a wide variety of events and conditions in their environments. The human visual system can detect events that deliver only a handful of photons, or events that are twelve orders of magnitude brighter; it can make use of spatial offsets that are smaller than the minimum angular distance between its photoreceptors; and with only 4 types of receptors it can differentiate thousands of colors. Audition can detect movements on the order of magnitude of the diameter of a hydrogen atom, and can localize sounds using time offsets at the two ears of only a few microseconds. Touch, taste and smell have similarly impressive capabilities.

Why do we consider ourselves to have only five senses? Aristotle's list of sight, hearing, touch, taste and smell seems hopelessly out of date when we look into the body and find not five but dozens of different types of sensory receptors. There are sensors for the lengths and tensions of muscles, for position and acceleration of the head, and other physical events. Separate systems of skin receptors pick up light touch, pressure, vibration, heat and cold. Much of what we consider to be taste is really mediated by olfactory receptors that report the presence of chemicals in the outside environment during inhalation, but detect chemicals in food during exhalation. The system reorganizes itself twice with every breath. Without modern neurophysiology we would not know that we have two kinds of auditory receptors, four kinds of visual receptors, perhaps a thousand different kinds of olfactory receptors. The connection between receptors and experience seems tenuous indeed.

Yet the concept of five senses endures. The reason why was suggested by J.J. Gibson [30], who identified the senses not as collections of receptors but as five kinds of activities that humans engage in while investigating their environments, five modes of orienting – seeing, hearing, touching, tasting and smelling. These are verbs, activities that the body engages in to explore the world. The many types of receptors collapse into a few ways of orienting, five modes of sensory attention. Receptors for internal states, such as those sensing the condition of the muscles or the orientation of the head, do not participate in this exploration, and so do not add to the list of senses.

At about the same time Merleau-Ponty [56] was pointing out the importance of the body in organizing perception. Perception is something you do, and you do it to situate the body in a sensory environment of possibilities for action. The processes of perception could no longer be considered as what happens inside a passive organism when something occurs in the surrounding world. Rather, the organism actively seeks out information about the world in its five different ways, and interprets that information in terms of what its own body can do with it.

2. Evolution of perception

Why did perception evolve? Clearly, animals that could detect the layout and events in their surroundings did better than those who could not, or who could detect and interpret them less well. That much is obvious. But the evolution was closely tied at every stage to the capabilities of the organism. The sensory systems did not develop in a vacuum, but in the swirl of daily challenges and opportunities present in the environment.

These challenges and opportunities are defined in terms of possibilities for action. A rabbit is an opportunity for a hawk, but a hawk is a challenge for the rabbit. So the rabbit evolves panoramic vision. It doesn't have to move even its eyes to monitor everything; it has a small binocular field in the front, and another in the back. It sees about equally well everywhere, and not very clearly anywhere. The hawk's visual adaptation is just the opposite, with very high acuity (better than a human's) in a small central field, and excellent binocular resolution. The contrasting strategies of vision serve each animal because of the interactions between their bodies and their environment. The more stress the rabbit feels, the more it freezes, immobile, relying on its camouflage. It can't outrun the hawk. So the rabbit and the hawk are caught in an evolutionary arms race, the rabbit becoming better at hiding and the hawk better at searching, each driven by the contrasting demands of its body and its corresponding ecological niche. The function of perception

is to control the body, and nothing else. Perceptual capabilities evolve to serve an animal in its ecological niche just as surely as limbs and teeth do.

3. Control of action

Action is normally the limiting factor in cognitive function – though the information-gathering capacity of our senses is enormous, we can do only one thing at a time. The process of selective attention serves to narrow the focus of both sensory and cognitive resources to the task at hand. Indeed, preparation for action is accompanied by increased capability or acuity in many senses. The question historically has been whether this is a real increase in sensitivity, or rather a change in criterion for accepting detections and discriminations in the relevant categories. The answer now appears to be that action–effect anticipation affects sensitivity directly [16], as shown by a signal-detection analysis that separates sensitivity from response bias.

4. Perception and eye movements – an example

The relation between perception and action is a complex one, because most motor effectors are also used to collect information. The hands, for example, can explore a surface for tactile information, or they can lift a weight. The tongue moves tactile and taste receptors around even as it participates in chewing food. Separating the two functions of the hands or the tongue adds a complication to untangling the perception–action link and interpreting the mechanisms of embodied perception.

There is an exception, though, to the dual use of muscles as motor effectors and as components of active sensory systems. That is in the eye muscles, effectors that function only to orient the eyes for optimal perception. They do not work on the outside world, yet their actions directly affect perception. There are two types of eye muscles: intraocular (inside the eye) muscles dilate or constrict the pupil and execute the focusing of the lens; extraocular (outside the eye) muscles direct the eye to locations of interest in the outside world.

The action of the intraocular muscles is as automatic as the actions of digestion – the observer has no influence over their functions. The actions of the extraocular muscles, however, promise to inform us about embodied perception. There are twelve muscles, six for each eye, that engage in several types of eye movements. Vergence movements, convergence and divergence, activate pairs of medial or the lateral extraocular muscles respectively, bringing the point of sight (the spot where the two lines of sight converge) closer to or further from the eyes. Pursuit movements track an object that is moving relative to the head, with the two eyes coordinated to move in the same direction at the same time. Quick jumps of the eyes, called saccadic movements, move the point of sight rapidly from one target object to another. Other eye movements, such as the vestibularly controlled movements of compensation for head movements, will not be considered here.

Of the three types of exploratory eye movements – vergence, pursuit and saccades – it is only the saccades that are under voluntary control. One can choose to move the eyes to fixate a new target, though usually even this function is automatic and progresses without the attention or even the awareness of the observer. Pursuit, however, requires a target to pursue. One cannot simply generate these movements in a fixed environment. To demonstrate this, you need only ask someone to imagine an object moving smoothly across their visual field, and observe their eye movements. They will generate a series of saccadic eye movements interspersed with periods of steady fixation. If they try harder the saccades may become smaller and more frequent, but smooth movements cannot be generated in this way.

To generate smooth pursuit, one must saccade to a target that happens to be moving; to generate vergence, one must saccade to a target that is moving toward or away from the observer. This leaves saccadic eye movements as the only movements that can be planned, or can be coordinated with anticipated actions. Thus to investigate eye movements in the service of action, we should start with saccades.

The retinal image seems to be closely tied to perception of the world of actions and possible actions, but it is not [21–23]. The image is sharp only at the fovea, the central area of the retina that has high spatial acuity and good color discrimination. Unperceived blood vessels run in front of the peripheral retina, and the retinal blind spot is only 15 deg from the fovea. The retinal picture bounces around with every saccade; the peripheral part of the image (but never the image on the fovea) slides across the retina during pursuit; and small, active eye movements make the retina drift across the optical image even during what seems like steady fixation. The world is perceived as stable nonetheless; stability establishes the consistent platform on which all other visual perception and sensorimotor

coordination depend. The mismatch between chaotic events at the retina and our perception of a stable, full-color, sharp-focus visual world demands an explanation if information gathered at the retina is to be used in controlling action. The problem will be examined at behavioral and physiological levels.

5. Spatiotopic fusion and efference copy

How does the visual system achieve perception of a stable and continuous world across blinks and saccades? Blinks and saccades not only bring disruptions to visual continuity due to lack of visual input [82]; they also shift the entire world by many degrees in the case of a saccade. Yet blinks and saccades happen three times per second on average and cause no noticeable visual disturbances to the observer.

The traditional explanation for space constancy, the perception of a stable visual world, is that whenever the motor centers of the brain send a command to change the position of the eyes, a copy of that command (the efferent [exiting] nerve signal) branches off to go to perceptual centers in the brain. That signal functions to compensate for the retinal displacement that will result. Because the brain receives a copy of the efference sent to the eyes, the compensating signal is called an efference copy.

Given what we know about efference copy and iconic memory, it is natural to assume that such information (e.g., about the size and direction of a saccade) is used to consolidate the images resulting from many visual fixations into a bigger internal representation of the external world, one that would not shift with every eye movement as the retinal image does. For nearly two centuries, therefore, it has been the dominant explanation for why the world appears stable despite shifts of the retinal image.

However, efference copy has its limitations as a hypothesized mechanism to achieve space constancy [8,9]. It fails at high speed [44,61], and underestimates how far the eyes deviate from their centered position [15]. Notably, a visual target displacement less than 1/3 of the saccade magnitude during a saccade is rarely noticed by people [11]. This size of error suggests that efference copy is inadequate to mediate space constancy. Thus the possibility is ruled out that there is some representation in the brain that puts together the images resulting from successive visual fixations [68,69], calibrated by precise changes in the efference copy signal [10]. The resulting stability, called spatiotopic fusion, must be accomplished by some other mechanism [53].

Behavioral evidence has come to the same conclusion [18,50]. Bridgeman and Mayer [13] used a simple dot matrix to test whether different snapshots are fused together with efference copy information. Within a 5×5 array (25 possible stimuli locations), these authors displayed 12 dots before a saccade and the remaining 12 complimentary dots after saccade completion. If spatiotopic fusion is performed by the visual system across saccades, the combined and fused dot patterns would easily reveal the location of the missing dot [49]. This was not the case according to their data, as subjects performed below chance level.

Similar findings were also reported by O'Regan and Levy-Schoen [63], who displayed half of a word before saccade, and the remaining half afterwards. Again, word fusion should be performed effortlessly if spatiotopic fusion across saccades was performed. Similarly, these authors did not observe evidence for spatiotopic fusion. Together, converging evidence suggests that efference copy is not accurate enough for spatiotopic fusion, and images across saccadic eye movements are rather independent from each other [62].

6. Change blindness and scene representation

The absence of image fusion across eye movements raises the question of how much information is actually left from one fixation to another if nothing is pasted together [36,37,45]. This question was exemplified in the past decade by the phenomenon of “change blindness”, where visual disruptions besides saccades can also dramatically lower people’s accuracy in detecting changes in a visual scene [72,80]. In this section we briefly review the change blindness literature that has shed light on how we internally represent the external world.

Change blindness can be observed in many ways, as long as there is a visual disruption such as an entire blank screen [72] or local occlusions [64]. This idea was first developed to mimic the blackout period between eye movements [7], where there is no visual input. To provide a quantitative measure, Rensink et al. [72] designed a flicker paradigm that alternated two nearly-identical pictures (pictures A and A') repeatedly with a blank screen inserted between each alternation, and measured the number of alternations presented until subjects became aware of the change. Surprisingly, people on average took 10, and up to 30, cycles until change was detected. In a real world scenario,

switching of conversational partners due to occlusions (analogous to the blank screen) was detected only 50% of the time [79].

These changes are hard to detect because the blank screen has masked the transient motion signals that usually accompany physical changes in the world. Thus, the changes no longer “pop out” at the observers, and require a top-down search to be revealed. This was used as evidence against a detailed internal visual representation, presumably because if such a representation existed, then a superimposition of the two images should reveal the change effortlessly. The fact that people had to spend many cycles searching serially through each object seems to suggest that such detailed representation does not exist, and therefore the focus of attention needs to search serially (like a spotlight) in order to identify the change [31].

Supporting evidence for such an attentional account also came from O’Regan et al. [64]. These authors took out the blank screen and placed several smaller checkerboard patches over certain areas of the image, but never covering the change. Change detection performance was equally poor, suggesting that changes do not need to be covered. Rather, motion transients (as produced by the patches) that capture one’s attention away from the change, much like an exogenous attentional cue, were enough to produce change blindness. Lastly, perhaps the most effective and clever demonstration of scene change blindness came from Henderson and Hollingworth [38]. Henderson and Hollingworth [38] displayed pictures that were half occluded by many gray vertical bars as if one were looking at the world through a prison window. When subjects made a saccade across an invisible predefined boundary, the vertical bars would shift slightly to the side so that the un-occluded half became occluded, and the originally-occluded areas became visible. This clever method thus altered every pixel between pictures A and A’, while keeping the higher level representation (e.g., gist) [76] intact. These authors observed poor change detection performance as well. Again, if one possessed a rich representation of the visual scene, a comparison between A and A’ would reveal two completely different pictures.

Not all studies have observed the same change blindness effect, however. To begin, for a long time it has been known that people can form very robust memories of at least 10,000 pictures [67,81], and hold them for weeks. We also know that iconic memory lasting only a fraction of a second is rich in details. Therefore, change blindness cannot be indicative of impoverished information in the iconic store [24] or an absence of visual representations in one’s visual long-term memory (VLTM) [46,47]. Instead, it must be pointing to post-iconic and pre-VLTM processes that somehow prevent information from getting stored or consolidated. Indeed, studies that used a free-viewing paradigm, which allowed participants to view a scene for a few seconds before triggering a change or presenting a questionnaire, have reported accurate visual memory and change detection [42,55]. Becker, Pashler, and Anstis [6] also observed accurate change detection when they increased the inter-stimulus interval, thus allowing iconic information to be consolidated into VSTM. Notably, Hollingworth and Henderson [42] gave their participants an object discrimination task online (VSTM) or after 5 to 30 min (VLTM). Change detection was good in both conditions, suggesting that visual information indeed can accumulate in VLTM, as long as it is transferred without being overwritten or interfered with [6,90].

Note that since VSTM has limited capacity [52], a recency effect has been reported by Zelinsky and Loschky [95] as the number of intervening objects increased. Thus, these results suggest that VSTM capacity places a constraint on the number of objects that can be stored, and information must be transferred to VLTM or medium-term storage [43,54] to prevent its being lost.

It is important to note the role of attention throughout this transfer and consolidation process [73]. Supporting evidence came from Wheeler and Treisman [90], who presented participants with four items that either changed color, changed shape, or swapped color or shape (the binding condition) in a change detection task. They found that performance in the binding condition declined when the test array consisted of all four items rather than one single probe item. These authors concluded that attention was distracted in the four-item condition and consequently was withdrawn from the items in VSTM, which suggested a critical role for attention in maintaining coherent representation.

Subsequent studies questioned whether maintaining binding information demanded more attentional resources than single-feature objects [2,48]. Nevertheless, despite different observed magnitudes in the effect of attention, these studies all reported worse performance when attention is distracted, thus suggesting that attention is critical in the processes of change detection.

6.1. *Why does change blindness occur?*

If internal representations do exist and accumulate to some extent, how can change blindness occur in the first place? The answer is not a simple one, as change detection requires several successful stages of processing, including encoding, maintenance, retrieval, comparison, and report of visual information. Failure at any particular step would render change detection unsuccessful. Here we describe each stage and how it contributes to change blindness.

6.2. *Encoding*

The encoding process is the first necessary step to ensure change detection [17]. Obviously, a changing object must be fixated before and after the change in order for the difference to be encoded. This explains why Hollingworth and Henderson observed higher detection accuracy when they controlled for eye movements. Another important aspect is the fragility of a visual representation in VSTM before it gets consolidated into VLTM [41]. Therefore, presenting the next stimulus too early can disrupt the formation processes of the current representation, causing improper encoding due to information overwrite [5]. This perhaps explains why it is particularly difficult to detect changes in the flicker paradigm, because the alternating rate may be too fast in comparison to other paradigms that allow free viewing.

6.3. *Maintenance*

Given the limited resolution and capacity of VSTM [29,52], it is clear that not all visual information can be maintained. Indeed, Zelinsky and Loschky [95] found that although [33–35] visual representation does accumulate, it hits a ceiling at three objects, which perfectly reflects the capacity of VSTM in the literature.

As mentioned before, bound representations also require attention during the maintenance period [90]. As a result, only a fraction of the encoded objects are maintained over time against decay, interference, and overwriting.

6.4. *Retrieval and comparison*

When information is accurately encoded and successfully maintained, it needs to be retrieved correctly in order to be useful for conscious detection. Angelone et al. [4] found that when people failed to detect changes, they could recognize the changed object above chance, suggesting that change blindness can occur when information is stored but not retrieved. Hollingworth [40] also observed improved change detection when retrieval and comparison load was reduced by clearly indicating object to compare at the end of each trial.

6.5. *Report*

Report and judgment of response is often overlooked by vision scientists. Specifically, a change blindness task is usually done with explicit verbal reports, which require one to be well aware of the change, thus creating a high threshold for reporting changes. Researchers have reported that people can guess the location or identity of a change reliably better than chance when responding with measures other than explicit verbal report [27,74]. Therefore, the amount of information in the visual memory can be underestimated by explicit verbal report. Indeed, studies on implicit change detection suggest that alternative measures that do not require explicit report, such as guesses [27] and direct motor gestures [87], can often reveal information that would not have been picked up by explicit verbal reports.

6.6. *The role of action in change detection*

So far we have provided a brief overview of scene perception and representation in the context of change detection, but where does the perception–action link come into the picture? If the goal of perception is to plan action [32], it is plausible that one may perceive an object or scene differently if such stimulus is of significance in terms of action. Therefore, if one can signal the visual system that certain stimuli are action-relevant, an additional recruitment of the sensorimotor system may possibly change the way we perceive and represent those stimuli.

Indeed, Witt, Proffitt, and Epstein [92,93] conducted a series of studies documenting how an intention to act may change the way distance or weight is perceived. Notably, distance is perceived as further when the perceiver intends to throw, and closer when the perceiver intends to use a tool. Thus, perception seems to be modulated by the possibility of asserting action, or even one's intention to act. Perception changes when the job of vision changes from describing the world to acting on it.

6.7. *Two functions of vision*

Here we first summarize findings supporting the idea of two qualitatively different visual systems that focus on visual recognition and action planning respectively, then describe how this distinction may have an impact on how we perceive and represent scene information.

Vision has two kinds of jobs to do. The first, and the one most studied to date, concerns perception of properties of the visual world and events in it. This type of vision can be called cognitive, because it is divorced from action. A passive observer can pick up enormous amounts of information from the world without interacting physically with it, except for eye movements that have no influence on the environment. This information is used to inform about the environment; the information can then guide the planning of future actions.

The second job of vision is to control visually guided behavior in real time. Threading a needle, throwing a stone, or picking a cherry are jobs that require close coordination between vision and action, a close sensorimotor coordination. Much of the processing takes place quickly and unconsciously.

The qualitative differences between the cognitive and sensorimotor visual systems have been well documented. The need for different systems that process representational and spatial information separately was first raised by Trevarthen [85] in monkeys, and later demonstrated by Bridgeman, Kirsh and Sperling [12] in humans, where participants could point accurately to a visual target whose change in position they could not consciously detect due to saccadic suppression. Mishkin and Ungerleider [59] further proposed two anatomically distinct ventral and dorsal pathways for processing vision. Since then, Milner and Goodale [58] labeled the two pathways as the “what” and “how” to differentiate the two by the functions that they serve. Bridgeman, Peery and Anand [14] used Paillard's [65] nomenclature of “cognitive” and “sensorimotor” systems to emphasize the idea of two visual *systems* and the qualitative differences between their functions, distinct from the neurological underpinnings of the two systems.

Since Bridgeman et al.'s first demonstration (1979), the psychophysical differences between the cognitive and sensorimotor systems have been studied extensively, especially over the last decade. A notable demonstration was the differential magnitude of perceived size each system exhibits toward the Ebbinghaus illusion [3], the perception of a circle as smaller if it is surrounded by large circles than if it is surrounded by small circles. Aglioti et al. proposed that the cognitive system is easily influenced by the contextual information (critical to the Ebbinghaus illusion), whereas the sensorimotor system is not. Although this topic is still subject to debate [28], it is generally accepted that the two systems do differ in their reactions to some illusions [14,88], although to a lesser degree than originally reported [89]. Some patients with visual agnosia can also utilize the sensorimotor system and perform certain goal-directed behaviors above the chance level despite their inability to “see” these objects [19,66]. Bridgeman and colleagues also tested this difference between the two visual systems with a shifted frame that biased participants' spatial reference, which resulted in static mislocalizations when measured with a verbal probe but not with a pointing action [14]. These differences originate from the different purposes that the two systems serve: the cognitive system supports visual analysis and identification whereas the sensorimotor system supports real-time motor action.

7. Embodied cognition and the two visual systems

Several studies have extended the two visual systems idea into embodiment, to suggest that visual analysis of objects can be altered when one's own hands are near the display, perhaps by creating a stronger attentional engagement [1,25,70,71]. This is distinct from previous studies using visual illusions because there is no difference between the way the cognitive and sensorimotor systems should respond (i.e., button press in all the studies reviewed above, no verbal vs. pointing dissociation in responses). Using a covert attention paradigm, Reed et al. [71] found faster reaction time to target onset detection by asking participants to simply place one hand near the computer display. These authors concluded that a spatial prioritization for attention was created when one's hand (or hands) was in proximity.

Using visual search, inhibition of return, and attentional blink paradigms, Abrams et al. [1] also reported slower shifts of attention when one's hands were near the display (see [83], for similar effects in saccadic eye movements). Covering the hands did not attenuate the slowing effect of nearby hands – proprioception, or knowledge that the hands were near the display, were enough. Abrams and colleagues interpreted their findings by suggesting a mechanism that could facilitate the detailed evaluation of objects for potential manipulation [1].

This interpretation, however, is not entirely supported by the data because a slower search rate could arise from a number of negative factors, and does not necessarily imply a positively 'altered' visual processing. Therefore, without accuracy data on one's sensitivity to external stimuli, the claim of a "detailed evaluation" is premature. There has been only one study that reported an increased accuracy in spatial discrimination when hands are placed in proximity [26]. The task involved an on-line discrimination of small lights from different near-hand and far-hand positions; greater accuracy was observed only when the hands were near the lights. Dufour and Touzalin's on-line discrimination task can be somewhat difficult, and to an extent requires cognitive effort such as inhibitory control.

How can the sensorimotor system play a role in change detection? Several studies have demonstrated an improved visual acuity and attention when one's hands are placed next to the stimulus display, suggesting a unique neural population that is recruited when the hands are involved [1,26,51,71]. Tseng et al. [87] combined this hand manipulation with a change detection paradigm and indeed found an improved change detection performance whenever the observer's hands were near the display. Interestingly, these authors found that the facilitative effect was mainly driven by the right hand, but adding the left hand produced a superadditive overall facilitation, at least in right-handed individuals.

These results suggest that the brain is particularly interested in where the hands are, presumably because that is where actions happen [32]. In addition, since this nearby-hand facilitation has been observed in other paradigms and is not exclusive to change detection, it is likely that the facilitation takes place very early at the perceptual and encoding stage, rather than at the maintenance, retrieval, comparison, and report stages.

Based on these findings, a more stringent test is required of the nearby-hand facilitation with a change detection paradigm. Unlike previous tasks investigating the effects of nearby-hand positions, change detection is a complex task that requires the use of visual short-term memory (VSTM) [91]. Successful change detection begins with attentional allocation, encoding and maintenance of object details, and ends with retrieval and comparison processes. Failure in any of these steps may render change detection unsuccessful.

Furthermore, difficulty of the change detection task can be easily manipulated by changing the number of objects available for comparison [39]. Therefore, the nature of the nearby-hand effect, whether positive or negative, can be observed from its influence on one's attention and VSTM via change detection performance. To this end, we designed an experiment with varying levels of difficulty to test whether findings on nearby-hand facilitations in reaction time could also be observed with an accuracy measure. Difficulty was manipulated by changing either one of eight or one of twelve objects in an array.

More important, can facilitatory effects be observed in a task that requires extensive involvement of high level processes such as memory, comparison, and decision-making? In subsequent work we probed the nature of the facilitatory effect and investigated whether the magnitude of the effect was localized, graded, or uniformly distributed, depending on the distance between the stimuli and the hands.

We used a classic change detection paradigm from Luck and Vogel [52], which had often been used to determine the capacity of VSTM. The prediction here is straightforward: if nearby-hand positions really bring forth a more detailed analysis of visual stimuli, then participants should be more sensitive to changes when they place their hands near the stimuli than when their hands are in their laps. We observed a significant enhancement of change detection for hand position and set size, but no interaction between them. The novel finding here is the facilitatory effect of hand position regardless of task difficulty, indicating that participants were discriminating the items better when their hands were on the display, which led to better change detection. We also concluded that participants encoded more information in VSTM when their hands were near the display.

To investigate the question of the distribution of the hand effect across the display, we made a few modifications to our experimental setup. First, the visual display stretched to cover the entire monitor so that all stimuli could be positioned anywhere from the center of the monitor to where the hands were placed. Second, in addition to the left and right hands-on (hands-horizontal) position, we added a top-and-bottom (hands-vertical) condition as well as one-hand conditions. Therefore, if the facilitation was of a graded nature, we should observe an improvement in performance

that, from the center of the display, spreads outward horizontally in the hands-horizontal condition and vertically in the hands-vertical condition.

Most of the regional gains were quite stable and evenly distributed, though. None of the columns and rows in the display were significantly different from the others in any condition, suggesting a uniform improvement in change detection performance near the hands. Therefore, the visual system seems to perform an equally extensive analysis for all items on the display. This suggests that it is not the hands *per se* that attract attention, which would predict a localized facilitation near the hands. Rather, the nearby hands create a stronger attentional engagement [1,71] with the display as a whole.

However, these findings do not entirely rule out the graded distribution perspective. Objects near the hands presumably represent items that are more likely to provide possibilities for bodily engagement in visually guided action. Therefore it is not surprising that these objects should receive prioritized allocation of attention.

This logic, however, hits a stall point when the prioritized region gets too big, or if the number of stimuli within one area gets too large. For example, when the size of the prioritized region exceeds the coverage of a single saccadic eye movement, is there a prioritized location among all the prioritized regions? Similarly, if too many stimuli are competing for attentional resources within one given area, it is plausible that certain areas are prioritized first.

Facilitation was stronger when hands were placed vertically. This effect was robust and remained significantly better than the other conditions when task was difficult (set size 12). This is a surprising finding as we did not expect the vertical effect to be any stronger than the horizontal.

One possible explanation is that the vertical hand positions required much more effort to maintain. As a result, the greater physical effort created a greater salience of the motor posture. However, Davoli and Abrams [20] have reported a replication of the Abrams et al. [1] study with imagined nearby-hand postures (relative to imagined hands placed behind participants' back). Therefore the physical effort associated with the hand placement is unlikely to be the only key here. Perhaps a simpler explanation is that by giving our participants a self-paced pausing period between the trials, the chances of missing an image from the following trial (picture A or A') due to blinks became smaller.

When we tested the effect of each hand separately, only the right hand boosted participants' performance, and the left hand was not effective. The right hand was facilitatory even if it was placed on the left side of the display, showing that the effect followed the hand and not the corresponding side of the body. This does not mean, however, that the left hand is not facilitatory, for with both hands on the display the participants performed better than with only the right hand.

Thus the left hand, while offering no facilitation on its own, enhanced the effect of the right hand when both hands were engaged, a superadditive nonlinearity. This may correspond to a difference between a one-hand grasp (almost always with the right hand in right-handed subjects) and a more powerful two-hand grasp. A one-hand left grasp, in contrast, would occur only rarely in these subjects.

The key difference that sets the present study apart from previous investigations on the facilitatory effect of nearby-hand positions is the accuracy measures and the role of working memory in the change detection paradigm. As these results have indicated, nearby hands indeed enhance visual working memory and thereby improve change detection. We believe this facilitatory effect stems from the same underlying mechanism that previous studies have observed [1, 26,70,71,83]. Therefore, the benefits of nearby hands must have originated from an earlier stage of perception rather than acting directly on working memory.

8. Application to naturalistic images

These experiments were performed with stimulus arrays consisting of color patches superimposed on a uniform background. While such displays allowed us to quantify the information content of the images and of the responses to them, they lack the ecological validity of natural environments, the conditions under which embodied cognition evolved to perform. In another set of studies we investigated responses to natural images, photographs of outdoor scenes.

In terms of detecting changes in natural scene pictures, Tseng and colleagues [87] used an implicit change detection paradigm and asked participants to guess the location of a change either with a number (indicating one of six regions of the image) or a pointing motion. To allow for improvement under one condition or another, we allowed only one presentation of an original and a slightly altered photograph, so that the participants hardly ever succeeded in detecting the location of the change.

Participants indeed performed reliably better than chance with their verbal responses, even when they only guessed at the correct location, supporting previous reports of implicit change detection [27]. Most important, participants who responded with a pointing motion outperformed those with verbal response, suggesting an advantage for those who used action.

This study was carried out with a screen 60 cm from the subjects, so that they could directly jab the screen containing the images. The motor advantage disappeared, however, when we moved the display out to 3 m from the participants, and they pointed with a laser pointer. The motor advantage was also absent when participants had to use the laser pointer at the nearer distance, indicating that actual contact with the screen, and not verbal versus action response, was the key to better performance. Again, these results suggest that it is not the distance to the target material, but rather the proximity of hands, that allows people to detect changes more frequently.

8.1. Neurological evidence

The neural mechanism of the nearby-hand effect is still unclear at this point. Previous investigations have relied on the idea of bimodal neurons, which react to both visual and tactile signals, to explain the nearby-hand advantage. The strongest evidence comes from the study of Schendel and Robertson [77], who suggested that arm positions near or within one's blind field can attenuate visual loss in patients with homonymous field deficits.

The findings from this study, however, are still under vigorous debate. It has recently been shown that when patients' blind fields and other methodological details are properly controlled, the attenuation of visual loss due to nearby hands can no longer be observed [78]. Thus, it remains unclear whether bimodal visual–tactile neurons are involved in the enhanced visual analysis we have reported here.

Besides the bimodal neurons, one likely region to mediate the nearby-hand effect is the posterior parietal cortex, which is a part of the frontoparietal attentional network and marks the beginning of the dorsal “action” pathway [59]. The posterior parietal cortex has been shown to update spatial mapping [57,60], to direct attention [75], and most important, to encode and maintain information in VSTM [84,86,94]. All of these functions are necessary components of successful change detection. Therefore the posterior parietal cortex's projection to the dorsal pathway and the attentional network makes it a likely region that mediates the nearby-hand effect.

These findings are best explained with those bimodal neurons in the parietal region whose receptive fields move with the hand [71]. Specifically, the posterior parietal cortex within the parietal lobe marks the end of the dorsal pathway, which makes it a likely candidate behind the sensorimotor system. Interestingly, neuroimaging studies have reported higher posterior parietal cortex activation as VSTM load increases [84,94]. Disrupting right posterior parietal cortex activity with transcranial magnetic stimulation indeed impaired people's ability to detect changes, and doing so at different specific timing revealed that posterior parietal cortex is especially critical to early encoding processes [86] rather than the later retrieval and response stages, which is consistent with the behavioral findings reviewed above.

8.2. Conclusion

We have reviewed some of the findings from the change blindness literature to show how action and body representation are involved in scene perception and coding. We started with evidence against spatiotopic fusion because efference copy is too slow and too low in gain to mediate space constancy. Thus, space constancy does not depend on a collage of images, but rather each fixation is independent from the previous one. This, however, does not mean that no visual information is stored from each fixation. When given enough time and no disruption, information can be stored in visual short-term memory and later transferred to visual long-term memory, and this process requires attention. Lastly, action seems to benefit these processes, especially at the perceptual and encoding stage, by creating a stronger attentional engagement with the perceptual world. Action and perception are inextricably linked in interactions of the organism with the world.

References

- [1] Abrams RA, Davoli CC, Du F, Knapp WK, Paull D. Altered vision near the hands. *Cognition* 2008;107:1035–47.
- [2] Allen RJ, Baddeley AD, Hitch GJ. Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General* 2006;135:298–313.
- [3] Aglioti S, DeSouza JF, Goodale MA. Size-contrast illusions deceive the eye but not the hand. *Current Biology* 1995;5:679–85.

- [4] Angelone BL, Levin DT, Simons DJ. The relationship between change detection and recognition of centrally attended objects in motion pictures. *Perception* 2003;32:947–62.
- [5] Becker M, Pashler H. Volatile visual representations: failing to detect changes in recently processed information. *Psychonomic Bulletin and Review* 2002;9:744–50.
- [6] Becker MW, Pashler H, Anstis SM. The role of iconic memory in change-detection tasks. *Perception* 2000;29:273–86.
- [7] Blackmore SJ, Brelstaff G, Nelson K, Troscianko T. Is the richness of our visual world an illusion? Trans-saccadic memory for complex scenes. *Perception* 1995;24:1075–81.
- [8] Bridgeman B, Lathrop B. Interactions between cognitive space and motor activity. In: Mast F, Jänke L, editors. *Spatial processing in navigation, imagery and perception*. Berlin: Springer; 2007.
- [9] Bridgeman B. In: Gilchrist I, editor. *Space constancy*. Oxford: Oxford University Press; 2010.
- [10] Bridgeman B, Hejden LVd, Velichkovsky B. A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences* 1994;17:247–92.
- [11] Bridgeman B, Hendry D, Stark L. Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research* 1975;15:719–22.
- [12] Bridgeman B, Kirch M, Sperling A. Segregation of cognitive and motor aspects of visual function using induced motion. *Perception & Psychophysics* 1981;29:336–42.
- [13] Bridgeman B, Mayer M. Failure to integrate visual information from successive fixations. *Bulletin of the Psychonomic Society* 1983;21:285–6.
- [14] Bridgeman B, Peery S, Anand S. Interaction of cognitive and sensorimotor maps of visual space. *Perception and Psychophysics* 1997;59:456–69.
- [15] Bridgeman B, Stark L. Ocular proprioception and efference copy in registering visual direction. *Vision Research* 1991;31:1903–13.
- [16] Cardoso-Leite P, Mamassian P, Scuitz-Bosback S, Waszak F. A new look at sensory attenuation: action–effect anticipation affects sensitivity, not response bias. *Psychological Science* 2010;21:1740–5.
- [17] Castelano MS, Henderson JM. Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance* 2007;33:753–63.
- [18] Currie CB, McConkie GW, Carlson-Radvansky LA, Irwin DE. The role of the saccade target object in the perception of a visually stable world. *Perception and Psychophysics* 2000;62:673–83.
- [19] Danckert J, Rossetti Y. Blindsight in action: what can the different sub-types of blindsight tell us about the control of visually guided actions? *Neuroscience Biobehavioral Research* 2005;29:1035–46.
- [20] Davoli CC, Abrams RA. Reaching out with the imagination. *Psychological Science* 2009;20:293–5.
- [21] Deubel H, Schneider WX. Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision Research* 1996;36:1827–37.
- [22] Deubel H, Schneider WX, Bridgeman B. Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research* 1996;36:985–96.
- [23] Deubel H, Schneider WX, Bridgeman B. Transsaccadic memory of position and form. *Progress in Brain Research* 2002;140:165–80.
- [24] Di Lollo V. Temporal characteristics of iconic memory. *Nature* 1977;267:241–3.
- [25] di Pellegrino G, Frassinetti F. Direct evidence from parietal extinction of enhancement of visual attention near a visible hand. *Current Biology* 2000;10:1475–7.
- [26] Dufour A, Touzalin P. Improved visual sensitivity in the perihand space. *Experimental Brain Research* 2008;190:91–8.
- [27] Fernandez-Duque D, Thornton IM. Change detection without awareness: do explicit reports underestimate the representation of change in the visual system? *Visual Cognition* 2000;7:324–44.
- [28] Franz VH, Gegenfurtner KR, Bühlhoff HH, Fahle M. Grasping visual illusions: no evidence for a dissociation between perception and action. *Psychological Science* 2000;11:20–5.
- [29] Gibson BS, Li L, Skow E, Brown K, Cooke L. Searching for one versus two identical targets: when visual search has a memory. *Psychological Science* 2000;11:324–7.
- [30] Gibson JJ. *The senses considered as perceptual systems*. Boston: Houghton–Mifflin; 1966.
- [31] Gilchrist ID, Harvey M. Refixation frequency and memory mechanisms in visual search. *Current Biology* 2000;10:1209–12.
- [32] Hayhoe M, Aivar P, Shrivastava A, Mruczek R. Visual short-term memory and motor planning. *Progress in Brain Research* 2002;140:349–63.
- [33] Henderson JM. Identifying objects across saccades: effects of extrafoveal preview and flanker object context. *Journal of Experimental Psychology: Learning Memory and Cognition* 1992;18:521–30.
- [34] Henderson JM. Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General* 1994;123:410–26.
- [35] Henderson JM. Human gaze control during real-world scene perception. *Trends in Cognitive Science* 2003;7:498–504.
- [36] Henderson JM, Hollingworth A. High-level scene perception. *Annual Review of Psychology* 1999;50:243–71.
- [37] Henderson JM, Hollingworth A. Eye movements and visual memory: detecting changes to saccade targets in scenes. *Perception and Psychophysics* 2003;65:58–71.
- [38] Henderson JM, Hollingworth A. Global transsaccadic change blindness during scene perception. *Psychological Science* 2003;14:493–7.
- [39] Henderson JM, Siefert AB. Types and tokens in transsaccadic object identification: effects of spatial position and left–right orientation. *Psychonomic Bulletin and Review* 2001;8:753–60.
- [40] Hollingworth A. Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance* 2003;29:388–403.
- [41] Hollingworth A. Constructing visual representations of natural scenes: the roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance* 2004;30:519–37.

- [42] Hollingworth A, Henderson JM. Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance* 2002;28:113–36.
- [43] Hollingworth A, Schrock G, Henderson JM. Change detection in the flicker paradigm: the role of fixation position within the scene. *Memory and Cognition* 2001;29(2):296–304.
- [44] Ilg U, Bridgeman B, Hoffmann K-P. Influence of mechanical disturbance on oculomotor behavior. *Vision Research* 1989;29:545–51.
- [45] Irwin DE, Yantis S, Jonides J. Evidence against visual integration across saccadic eye movements. *Perception and Psychophysics* 1983;34:49–57.
- [46] Irwin DE, Zacks JL, Brown JS. Visual memory and the perception of a stable visual environment. *Perception and Psychophysics* 1990;47:35–46.
- [47] Irwin DE, Zelinsky GJ. Eye movements and scene perception: memory for things observed. *Perception and Psychophysics* 2002;64:882–95.
- [48] Johnson JS, Hollingworth A, Luck SJ. The role of attention in the maintenance of feature bindings in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance* 2008;34:41–55.
- [49] Jonides J, Irwin DE, Yantis S. Integrating visual information from successive fixations. *Science* 1982;215(4529):192–4.
- [50] Jonides J, Irwin DE, Yantis S. Failure to integrate information from successive fixations. *Science* 1983;222:188.
- [51] Lloyd DM, Azanon E, Poliakoff E. Right hand presence modulates shifts of exogenous visuospatial attention in near perihand space. *Brain and Cognition* 2010;73:102–9.
- [52] Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. *Nature* 1997;390:279–81.
- [53] McConkie GW, Currie CB. Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance* 1996;22:563–81.
- [54] Melcher D. Persistence of visual memory for scenes. *Nature* 2001;412:401.
- [55] Melcher D. Accumulation and persistence of memory for natural scenes. *Journal of Vision* 2006;6:8–17.
- [56] Merleau-Ponty M. *The visible and the invisible*. Evanston, IL: Northwestern University Press; 1968.
- [57] Merriam EP, Genovese CR, Colby CL. Spatial updating in human parietal cortex. *Neuron* 2003;39:361–73.
- [58] Milner AD, Goodale MA. *The visual brain in action*. New York: Oxford University Press; 1995.
- [59] Mishkin M, Ungerleider LG, Macko KA. Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences* 1983;6:414–7.
- [60] Morris AP, Chambers CD, Mattingley JB. Parietal stimulation destabilizes spatial updating across saccadic eye movements. *Proceedings of the National Academy of Sciences USA* 2007;104:9069–74.
- [61] Nagle M, Bridgeman B, Stark L. Voluntary nystagmus, saccadic suppression, and stabilization of the visual world. *Vision Research* 1980;20:1195–8.
- [62] Oliva A, Wolfe JM, Arsenio HC. Panoramic search: the interaction of memory and vision in search through a familiar scene. *Journal of Experimental Psychology: Human Perception and Performance* 2004;30:1132–46.
- [63] O'Regan JK, Levy-Schoen A. Integrating visual information from successive fixations: does trans-saccadic fusion exist? *Vision Research* 1983;23:765–8.
- [64] O'Regan JK, Rensink RA, Clark JJ. Change-blindness as a result of 'mudsplashes'. *Nature* 1999;398:34.
- [65] Paillard J. Cognitive versus sensorimotor encoding of spatial information. In: Ellen P, Thinus-Blanc C, editors. *Cognitive processes and spatial orientation in animal and man*. Dordrecht, The Netherlands: Martinus Nijhoff; 1987. p. 43–77.
- [66] Perenin MT, Rossetti Y. Grasping without form discrimination in a hemianopic field. *Neuroreport* 1996;7:793–7.
- [67] Potter MC. Very short-term conceptual memory. *Memory and Cognition* 1993;21:156–61.
- [68] Rayner K, McConkie GW, Ehrlich S. Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance* 1978;4:529–44.
- [69] Rayner K, McConkie GW, Zola D. Integrating information across eye movements. *Cognitive Psychology* 1980;12:206–26.
- [70] Reed CL, Betz R, Garza JP, Roberts RJ. Grab it! Biased attention in functional hand and tool space. *Attention Perception & Psychophysics* 2010;72:236–45.
- [71] Reed CL, Grubb JD, Steele C. Hands up: attentional prioritization of space near the hand. *Journal of Experimental Psychology: Human Perception and Performance* 2006;32:166–77.
- [72] Rensink RA, O'Regan JK, Clark JJ. To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science* 1997;8:368–73.
- [73] Rensink RA. Change detection. *Annual Review of Psychology* 2002;53:245–77.
- [74] Rensink RA. Visual sensing without seeing. *Psychological Science* 2004;15:27–32.
- [75] Rushworth MF, Taylor PC. TMS in the parietal cortex: updating representations for attention and action. *Neuropsychologia* 2006;44:2700–16.
- [76] Sampanes AC, Tseng P, Bridgeman B. The role of gist in scene recognition. *Vision Research* 2008;48:2275–83.
- [77] Schendel K, Robertson LC. Reaching out to see: arm position can attenuate human visual loss. *Journal of Cognitive Neuroscience* 2004;16:935–43.
- [78] Smith DT, Lane AR, Schenk T. Arm position does not attenuate visual loss in patients with homonymous field deficits. *Neuropsychologia* 2008;46:2320–5.
- [79] Simons DJ, Levin DT. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review* 1998;5:644–9.
- [80] Simons DJ, Rensink RA. Change blindness: past, present, and future. *Trends in Cognitive Sciences* 2005;9:16–20.
- [81] Standing L. Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology* 1973;25:207–22.
- [82] Thiele A, Henning P, Kubischik M, Hoffmann KP. Neural mechanisms of saccadic suppression. *Science* 2002;295:2460–2.
- [83] Thura D, Boussaoud D, Meunier M. Hand position affects saccadic reaction times in monkeys and humans. *Journal of Neurophysiology* 2008;99:2194–202.
- [84] Todd JJ, Marois R. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 2004;428:751–4.

- [85] Trevarthen CB. Two mechanisms of vision in primates. *Psychological Research* 1968;31:299–337.
- [86] Tseng P, Hsu TY, Muggleton NG, Tzeng OJL, Hung DL, Juan CH. Posterior parietal cortex mediates encoding and maintenance processes in change blindness. *Neuropsychologia* 2010;48:1063–70.
- [87] Tseng P, Tuennermann J, Roker-Knight N, Winter D, Scharlau I, Bridgeman B. Enhancing implicit change detection through action. *Perception* 2010;39:1311–21.
- [88] Vishton PM, Rea JG, Cutting JE, Nuñez L. Comparing effects of the horizontal-vertical illusion on grip scaling and judgment: relative vs. absolute, not perception vs. action. *Journal of Experimental Psychology: Human Perception and Performance* 1999;25:1659–72.
- [89] Vishton PM, Sepsens NJ, Nelson LA, Morra SE, Brunick KL, Stevens JA. Planning to reach for an object changes how the reacher perceives it. *Psychological Science* 2007;18:713–9.
- [90] Wheeler ME, Treisman AM. Binding in short-term visual memory. *Journal of Experimental Psychology: General* 2002;131:48–64.
- [91] Williams CC, Henderson JM, Zacks RT. Incidental visual memory for targets and distractors in visual search. *Perception and Psychophysics* 2005;67:816–27.
- [92] Witt JK, Proffitt DR, Epstein W. Perceiving distance: a role of effort and intent. *Perception* 2004;33:570–90.
- [93] Witt JK, Proffitt DR, Epstein W. Tool use affects perceived distance but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance* 2005;31:880–8.
- [94] Xu Y, Chun MM. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 2006;440:91–5.
- [95] Zelinsky GJ, Loschky LC. Eye movements serialize memory for objects in scenes. *Perception and Psychophysics* 2005;67:676–90.