Processing spatial layout by perception and sensorimotor interaction

Bruce Bridgeman and Merrit Hoover
University of California, Santa Cruz, CA, USA

Everyone has the feeling that perception is usually accurate—we apprehend the layout of the world without significant error, and therefore we can interact with it effectively. Several lines of experimentation, however, show that perceived layout is seldom accurate enough to account for the success of visually guided behaviour. A visual world that has more texture on one side, for example, induces a shift of the body's straight ahead to that side and a mislocalization of a small target to the opposite side. Motor interaction with the target remains accurate, however, as measured by a jab with the finger. Slopes of hills are overestimated, even while matching the slopes of the same hills with the forearm is more accurate. The discrepancy shrinks as the estimated range is reduced, until the two estimates are hardly discrepant for a segment of a slope within arm's reach. From an evolutionary standpoint, the function of perception is not to provide an accurate physical layout of the world, but to inform the planning of future behaviour. Illusions—inaccuracies in perception—are perceived as such only when they can be verified by objective means, such as measuring the slope of a hill, the range of a landmark, or the location of a target. Normally such illusions are not checked and are accepted as reality without contradiction.

Most people feel that they see the world as it exists in true geometric reality, a concept known as naïve realism (Ramsperger, 1940). Contrary to this belief, however, people often fail to perceive the world as it truly exists before them, but rather perceive it in a manner that combines the geometric layout of the environment and their own potential to interact with it, distorted by a number of less-than-reliable processing algorithms. This has been demonstrated in many perceptual studies (Bridgeman, 2004; Bridgeman, Gemmer, Forsman, & Huemer, 2000; Post, Welch, & Bridgeman, 2003) that examine various functions of the sensorimotor and cognitive systems and the use of visual information in each system. A focus of perception can be situating the body in space, rather than recognizing patterns in the abstract.

Traditional approaches, such as that of traditional Gestalt psychology or the more modern spatial vision, do not consider the body as a participant in perceptual processes. Patterns are considered to have an independent existence, their important relationships being with one another rather than with the observer. Here, two empirical examples will show the importance of embodied cognition, considering the links between perception and action in the grounding of cognition.

Correspondence should be addressed to Bruce Bridgeman, Department of Psychology, University of California, Santa Cruz, CA 95064, USA. E-mail: bruceb@ucsc.edu

http://www.psypress.com/qjep DOI:10.1080/17470210701623712
Normally we think of perception as accurate, so that we consider illusions to be interesting exceptions to the normal process of perception, when something goes awry, and the normal mechanisms of perception fail us. Psychophysicists use illusions as a kind of destructive testing of visual processing; just as an engineer might stretch a cable until it breaks and learn about the strength of the cable in the process, a psychophysicist stretches some aspect of perception until it breaks, producing an illusion where the perception does not match reality. I can make a low-reflectance ‘black’ surface seem white, for instance, by shining a spotlight on it in an otherwise dark room. Under normal illumination, however, the same surface will appear black. The abnormal illumination causes normally reliable perceptual mechanisms to fail.

In this context reality can be defined as something that remains the same upon closer examination, while illusion is something that is revealed to change upon closer examination. The surface that appeared white, for example, when it was the lightest surface in the room, appears black when all surfaces receive illumination. The whiteness was an illusion. What happens, though, if there is no closer examination? If the illusion is never revealed, the perception will be accepted as reality no matter how far it deviates from the actual surface features and geometric layout of the world.

A historical example of an unchecked perception is the appearance of the fixed stars, which all appear to be at the same distance from us as though fixed to a heavenly sphere that slowly rotates around us, with the north star as its axis. Until well into the 19th century this illusion was unchecked, and everyone in the world accepted it as reality. Christians still sing confidently of the heavenly spheres in some of their hymns to the almighty. Now we know, of course, that the heavenly sphere consists of stars only a few light years away, giant galaxies a billion times farther from us, and everything in between. The heavenly sphere is a perspective illusion. Still, it seems spherical when I look up at night; none of the normal depth cues tell me that the stars aren’t all equidistant from me.

The thesis of this paper is that illusions are not exceptions—rather, they are everywhere, and quantitatively accurate perception of reality is often a lucky accident. Since illusory perceptions are seldom checked, however, we accept the products of our perceptual systems as reality and move confidently through a world full of illusion.

Illusions occur not only in perception of space, but in other cognitive processes such as reasoning, where failures of elementary logic are disturbingly common. Even in serious domains such as estimating disease risk in diagnosis problems, given a test’s hit rate, false-positive rate, and prior probability of having a disease, most people come to wildly inaccurate conclusions. Even a majority of physicians fall prey to such cognitive illusions (Gigerenzer, 1998), analogous to geometric illusions because only formal calculation reveals the correct answer.

The examples I use are primarily spatial, a mode in which perceptions are often faulty but seldom checked. Distance estimates are frequently inaccurate by large amounts, for example, with verbal estimates by different people varying widely for the same conditions. Beyond a few dozen metres, we have little quantitative idea how far away things are. The hills we see in the distance might be 5 kilometres away, or 20. Normally, though, we have a good enough qualitative impression to make our way through the world. Perception can still be useful if it points us in the right direction, even when it is wildly inaccurate in a geometric sense.

For controlling motor activity, though, a qualitative impression is not good enough. If we don’t grab the right object, or jump the stream accurately to the opposite bank, we can be in big trouble. Action, even visually guided action, is often accurate even while perceptions of the same conditions and events as well as verbal descriptions suffer from illusions.

The first modern work demonstrating accurate motor activity despite illusory perception concerned induced motion. A target that seemed to be moving because of background motion, but
was actually fixed, elicited pointing to the fixed location, not to two different locations at the extremes of the apparent motion (Bridgeman, Lewis, Heit, & Nagle, 1979). In a complementary condition of the same experiment, a target whose real motion was nulled by an antagonistic induced motion elicited pointing to two different locations at the extremes of the real motion, despite lack of apparent motion (a double dissociation).

Here a more explicit definition of perception becomes necessary. The term is sometimes used to refer to any information pick-up that is subsequently used to control behaviour, whether verbal or instrumental. That definition, however, runs afoul of overgenerality: A thermostat, for instance, picks up environmental information and uses it to make a discrimination (too cold or not too cold), followed by an action (turning a furnace on or off). By the more general definition, the thermostat would have a capability for perception, violating the normal definition of the term. Here, perception is linked with awareness, which in turn can be defined in terms of language. A rough-and-ready definition of perception, then, is that if you can talk about it it’s perception, and if you can’t it’s not.

Milner and Goodale (1995) summarized a wealth of evidence for separate pathways mediating perception and motor action, including data from patients in whom one or the other pathway is damaged. Subsequent work has investigated this two-visual-systems hypothesis in detail. Patients, however, might adapt to a deficit by segregating functions that are normally integrated, isolating their deficit with a mental “firewall”. A contrast between illusory perception and accurate grasping in normal humans was demonstrated for the Ebbinghaus illusion (Haffenden & Goodale, 1998), the illusion that a small circle surrounded by large circles appears smaller than it really is, and vice versa. But there is a confound in the study; the perceptual measure summated the small-among-large circle illusion with the large-among-small illusion, while grasping at a circle involved only one illusion or the other. Franz, Gegenfurtner, Bülthoff, and Fahle (2000) pointed out that the two illusions sum nonlinearly and failed to replicate the contrast between perception and grasp when the two illusions were separated. A recent summary of the dissociation, and the controversy, is given in Goodale and Westwood (2004), with the conclusion that the two visual systems, while distinct, are not as independent as had once been assumed.

In the context of embodied cognition, we examine in detail two examples of such dissociations of perception and visually guided behaviour: estimates of location in biased environments and slopes of hills. Both have practical implications for everyday behaviour. In these examples, it has turned out that body position and action planning are critical both to the perceptual process and to interaction with the sensory world.

Estimates of location

Distorting observers’ perceptions of the position of a visual target is frighteningly easy. It requires only that the target be placed in an environment with more texture on one side of the world than on the other—for example, where the linear perspective of a rectangular environment is not centred on the objective straight ahead. Roelofs (1935) noted that the edge of a projected rectangle that is objectively centred in front of a subject will appear deviated to the side opposite the rest of the rectangle. If the rectangle is on the left, for example, its right edge (centred before the observer) will appear deviated to the right. Bridgeman, Peery, and Anand (1997) extended this phenomenon to introduce an induced Roelofs effect: A target inside the rectangle will appear deviated to the side opposite the rectangle’s deviation. The rectangle need not be deviated as much as the original rectangle of Roelofs—any offset will do. Illusions of location occur whenever a frame is presented asymmetrically in the visual field (Figure 1).

In an extensive series of experiments, my laboratory has investigated this effect with both perceptual and motor measures, with targets and inducing rectangles at various eccentricities and responses at varying delays. For this paper, I
review only the portion of the results necessary to illuminate the current issue, under conditions with immediate response and with a target in an observer’s objective midline.

**Perception**

If both an inducing frame and a target are centred, an observer has no difficulty in accurately localizing the target, given a five-alternative forced-choice task (Figure 2). If the frame is offset to one side, however, the observer verbally selects an X on the opposite side, revealing a deviated perception. The inducing conditions for this illusion are not extreme: There is no motion in the field, the observer has unlimited time, and there is no distracting secondary task. Nevertheless, perception is reliably made inaccurate.

In these experiments the frame could take one of three positions on a given trial, either left of centre, centre, or right of centre. The target could appear in one of five positions. A trial began with simultaneous onset of target and frame in an otherwise featureless field; they remained visible for 1 s before disappearing together. There was no motion, either apparent or real. The five-alternative forced choice was made immediately after stimulus offset.

The explanation for this inaccuracy is also illustrated in Figure 2: The observer’s centreline is deviated in the same direction as the frame. It is captured by the frame, so that the subjective centreline and the position of the objectively centred target no longer coincide. The observer, unaware that his centreline has been shifted (to the left in the illustrated example), picks an X to the right of the original target (Dassonville, Bridgeman, Bala, Thiem, & Sampanes, 2004). Perception is shifted in one direction because the subjective centreline has been shifted in the other. What seemed like a purely perceptual task turns out to be critically dependent on body posture and distortion of bodily coordinates.

**Motor control**

The observer’s task is to jab the location of the target just after stimulus offset, without sight of
the finger. Jab position is detected with a touch screen, so that no apparatus encumbers the limb. Under stimulus conditions identical to those that elicited the cognitive Roelofs mislocalization, motor behaviour remains accurate (Figure 3) regardless of frame position. The explanation for the continued accuracy is that the observer jabs in a shifted coordinate system revealed by the shifted straight ahead, measured in other jab trials where the task is to jab straight ahead in the presence of a frame but no target. If the frame is on the left, for example, the observer jabs to the right of a left-deviated centreline, hitting the target. The perceptual and motor localization systems have evolved in such a way that visually guided motor activity remains accurate despite large deviations in perception, which is allowed to wander without serious consequences.

While revealing a functional difference between perception and visually guided action, these experiments take place in an artificial laboratory setting with stimuli that are impoverished by comparison with the real world, though the visual world is seldom completely symmetrical, and small Roelofs-like illusions can be expected everywhere. The next set of experiments shows analogous phenomena in the natural world.

Slopes of hills

One especially salient example of the dichotomy between objective reality and perceived reality is found when estimating slope. During slope perception experiments, observers estimate the slope of a hill verbally and motorically (Proffitt, Bhalla, Gossweiler, & Midgett, 1995). In these studies, subjects were presented with a hill, either in reality or in virtual reality, and were asked to estimate its slope. In a verbal measure, subjects were asked to provide an estimate in degrees, while in a motor measure, subjects used their palms to adjust a tilt-board linked to a protractor. Estimates using the motor measure were significantly more accurate than estimates made in the verbal modality. The verbal measure greatly overestimated the actual slope of the hill.

With these experiments as a starting point, we asked how observers would respond if they were actually standing on the slope of a hill, rather than at its base, as they estimated slopes at various ranges from themselves. A critical difference between this design and that of Proffitt et al. (1995) is that the observer becomes part of the perceptual environment. There are two reasons why one might expect differences in slope estimation at different ranges: different neural processing of near space and greater effort required to negotiate longer climbs.

To the degree that cognition is grounded in action, one might anticipate that space within the immediate action range of near space (within arm’s reach) is processed in a mode appropriate for immediate action, while more distant space would be evaluated primarily in an action-planning mode. Because interaction with near space does not require locomotion, there is no reason to overestimate slope in that range and every reason to judge it accurately. Findings in neurophysiology suggest that near space is processed differently from far space (Andersen, 1989; Andersen, Essick, & Siegel, 1985). Some parietal neurons respond to visual targets that are within reach, but will not respond to the same targets presented more distantly. Further, some neurons in monkey premotor cortex respond to
visual stimuli in the space near the arm or hand, but not to the same stimuli presented out of arm’s reach (Graziano, Yap, & Gross, 1994). In further examination of these neurons, it was hypothesized that there are visuomotor neurons that respond only to visual stimuli within “peripersonal” (near) space. These findings suggest that the space available for immediate interaction may have different neural correlates than space further away.

The peripersonal space is not fixed—it can be extended by use of tools that allow direct manipulation of objects further from the body than arm’s length (Berti & Frassinetti, 2000). In our experiments, since observers used only their bodies to indicate slopes, the peripersonal space could be estimated to extend about 1 m away from the body.

A clinical study also provides evidence that coding of visual information changes as a function of spatial position relative to the individual’s hand or arm. In a patient with damage to the right primary visual cortex (V1), stimulus detection was considerably better when his arm was placed in the left visual hemifield (the “blind” field; Schendel & Robertson, 2004). This evidence suggests that the interpretation of a visual object depends in part on the individual’s ability to interact with the object.

These findings led us to hypothesize that differences in neural coding for near and far space would influence estimations of slope. If neurons code space differently within arm’s reach, then verbal and/or motor estimates may reflect these differences by showing changes between near and far space. Again, the experiments described here form only part of a larger study that covered five slopes and additional variables not reviewed here.

**Perception**

Our observers stood on a long, uniform 11-deg slope, on a paved surface. Looking up the hill, traffic cones were placed 1, 2, 4, 8, and 16 m from their station point. Observers first judged the slope of the hill between themselves and a given cone. The observers were comfortable with estimates in degrees of angle and were reminded that 0 deg represents a flat surface, 90 deg a vertical cliff, and other slopes fit within this range. We took advantage of a finding of Proffitt et al. (1995) that estimates with a hand-held disc, with an adjustable sector indicating the slope, were indistinguishable from direct verbal estimates in degrees, making the use of the hand-held disc unnecessary.

Results with the verbal measure are shown in Figure 4, where observers again greatly overestimate the slope at all distances with the verbal measure, replicating the Proffitt et al. (1995) result. The overestimate also increases with distance, however, increasing by a constant amount for each doubling of the distance over which the estimate is made, expressing quite precisely a logarithmic relationship between distance and slope estimate ($r = .997$). The data also fit a power function very well ($r = .995$).

**Motor control**

Earlier experiments have used apparatus to interface between the body and a motor measure of slope. To make the connection between body and environment more direct in the current experiments, we eliminated the apparatus and used a direct body posture to measure body-defined slope, shifting attention from a palm board to the body itself.

Rather than using the palm board to measure tilt, we used photography to record the motor
estimate. Observers were asked to hold their elbow against the body and the forearm perpendicular to the body and to raise or lower it to match the slope of the hill. They were instructed to look at the goal cone on the hill, not at their arm. The hand was extended straight, in the orientation of the arm. A photograph was taken of the forearm, using a digital camera mounted on a levelled tripod at approximately waist height. In this way we avoided body contact with any other apparatus and were able to rely on body posture alone.

After making verbal and motor estimates for one cone, observers stepped away, turned around once, and stepped up again to do estimates for another cone, selected in balanced order.

With the motor measure the errors were smaller than the errors in perceived slope (Figure 5) and also fitted well to both a logarithmic function \( r = .982 \) and a power function \( r = .981 \). All of the mean slope matches fit into a relatively narrow range between 20 deg and 30 deg.

Discussion

Unfortunately, the data do not support either of the theories that originally motivated the study. The theory that near space is handled by different neurophysiologic mechanisms than far space predicts a break in the function at about arm’s length (1 m), but the functions we found increase smoothly from the closest ranges to the furthest, fitting logarithmic or power functions very well throughout the range. The effort hypothesis predicts a linear increase of apparent slope with range, because an 8-m climb should take twice as much effort as a 4-m climb, and climbing 16 m should take four times as much effort. An effort hypothesis modified by anticipated fatigue in climbing the longer distance is implausible because the distances involved are not excessive for our young, healthy observers; they routinely climb longer and steeper slopes in their daily travels around the campus. Effort tempered by a foreshortening of apparent distance is also contradicted, because distance estimates are fairly accurate out to about 20 m, becoming foreshortened only at longer distances (Klatzky, Lippa, Loomis, & Golledge, 2003).

For the purposes of this paper, however, the implications of the results are clear—even in a full-cue natural environment, observers make large and systematic mistakes in judging the spatial layout of an important aspect of their world, while motor-oriented estimates of the same environment are more accurate.

GENERAL DISCUSSION

Both of the projects described here point to the centrality of the body in perceptual as well as motor activities. The Roelofs effect results were explained not only by the relative exocentric positions of target and mask, but also by their relation to the straight-ahead orientation of the body. Even for perception, it was not possible to exclude the body from estimates of position, though cognitive and motor systems made use of body information in different ways. The results of the experiments on the slope of a hill showed that estimates of the layout of the world are affected by the position of the body relative to objects in the world, as well as being influenced by the modality of response—cognitive or motor.

The results extend the range of evidence that demonstrates intimate relationships among body
sense, motor activity, and perception, relationships that apply also in other domains. Gentilucci and Dalla Volta (2008 this issue) have demonstrated not only that arm and hand movements can influence intonation, but that semantics can influence movement. The motor control system is also a primitive meaning processor. There are also close relations between actions and action verbs (Nazir et al., 2008 this issue; Taylor & Zwaan, 2008 this issue), showing that informational influence goes in both directions—actions affect verbal semantics, and semantics affect actions.

In general, the combination of these results with those of others supports a new way of looking at brain function. It now seems clear that language evolved at least in part from a gestural communication system, and brain control centers that evolve from one another remain in close contact. Most cognitive capabilities evolved from motor control systems and exploit ancient motor-oriented neurological organizations. The capabilities studied in the experiments described here are no exception.

CONCLUSION

The research on the Roelofs effect and on the slopes of hills illustrates a more general point: that perception frequently suffers from large errors, but that the errors usually go unnoticed and are without consequences. As long as the real-time control of motor activity remains accurate, perception need be only accurate enough to converge on the right answer—quantitative accuracy in all conditions is neither required nor available.

Perceptual illusions, rather than being exceptional, are ubiquitous, but usually inconsequential because our perceptual information is good enough to allow us to navigate the world. Perception need not provide the quantitatively accurate picture that we think we have. As long as the erroneous information is not checked, we accept it as reality, and naive realism can survive in a world of constant illusion. The function of perception then is not only to relay the layout of the world, but also to inform us of relationships and affordances that aid in planning future behaviour, beyond the immediate requirements of motor organization.

The brain can tolerate illusions of space because its job of facilitating action is distinct from providing a picture of the world. In this context the body is central to linking representations of space and of action, using perceptual information to plan movements and to predict their consequences.

REFERENCES

Gentilucci, M., & Dalla Volta, R. (2008). Spoken language and arm gestures are controlled by the same motor control system. The Quarterly Journal of Experimental Psychology, 00, 00–00.


