

---

## Guest editorial

---

### Vision and action: you ain't seen nothin' yet ....

#### 1. Paradigms and puzzles: blindsight in normal observers

Most scientists, most of the time, are engaged in 'normal science'. They advance knowledge within a fairly well-defined, and widely shared, framework of understanding---a paradigm (Kuhn 1962). Occasionally, however, a result so astonishing is obtained that it stretches the prevailing paradigm to breaking-point, and one senses that a scientific revolution is under way. In visual science, the recent experiments of Kolb and Braun (1995) on "blindsight in normal observers", described in some detail below, are certainly astonishing and---in concert with other recent findings on the relations between perception, action, and awareness---could well be paradigm-breaking. Of course, like a comfortable but leaky pair of shoes, an old paradigm is not abandoned simply because it fails. There will always be unexplained results that simply sit outside the paradigm as unresolved puzzles. Rather, old paradigms are displaced by better ones that can both accommodate existing results and incorporate the 'puzzles' in a principled fashion. It is too early to be sure, but the shape of a new paradigm for vision could well look something like that described by Milner and Goodale (1995) in their remarkable book *The Visual Brain in Action*. Readers of this journal will want to take notice, if they haven't already, because the gist of Milner and Goodale's thesis is that vision and perception are *not* the same thing.

Let's begin with Kolb and Braun. On the face of it, in their experiments they used a standard form of texture segmentation task, presenting observers with an array of oriented elements (the background) containing a smaller sub-array (the target) whose elements were oriented at right angles to the background elements. In experiment 1 the elements were dots moving obliquely; in experiment 2 each element of the array was a small stationary patch of oblique stripes (a Gabor function; see [figure 1a](#)). On each trial, observers had to identify which of four quadrants contained the target area, distinguishable by the orientation of its elements. In addition, they had to rate on a scale from 1 to 10 their confidence that the choice of target location was correct. In the standard conditions (as just described) the results were quite straightforward; there was a high correlation between rated confidence and actual performance. In other words, in this standard texture segmentation task *subjects know what they are doing*. The idea that subjects 'know what they are doing' (my phrase not theirs) was formalised by a novel application of ROC analysis to the confidence ratings. The details needn't detain us here, but the essence is that confidence was high when responses were correct, and low when incorrect. Thus in a strict signal-detection sense, observers knew about their own performance trial-by-trial, even though the stimulus characteristics were fixed. However, in two modifications of this task, subjects performed equally well (around 70%--75% correct, where 25% is chance) but *did not know what they were doing*.

For experiment 2, the critical modification was to rotate all the elements in one eye through 90° ([figure 1b](#)). Thus at every array location both orientations were present, one in each eye. Rivalry and suppression generally occur with such displays, but at the brief duration used here 'anomalous fusion' is often reported (Wolfe 1983). Physically summing the two eyes' images would eliminate all trace of the target (as shown in [figure 1c](#)), and so it is fairly easy to see that if binocular summation occurred before texture segregation, the target would disappear. Indeed, according to Kolb and Braun subjects reported no awareness of the target area and claimed to be 'guessing', even though they were not ( $d'=1.7$ ). Importantly, we do not have to rely only on these subjective reports, since the ROC analysis shows objectively that observers exhibited no knowledge of their own performance ( $d'$  was close to 0). They had sight, but no insight. Kolb and Braun conclude that "binocular fusion of orthogonally oriented elements conceals a target from visual awareness but does not prevent its localization" (page 337).

Strictly speaking, we should be very careful about moving from a clear experimental result---showing no knowledge of performance---to the conclusion that observers had no visual awareness of the target. There are

two tasks involved here, not one. Using vision to judge the target location is one task; judging the likelihood of response validity is another. But there is undeniably a close association between performance on these two tasks in the normal condition that is lacking in the critical condition. In the absence of any other operational definition, Kolb and Braun have effectively *defined* awareness as 'knowing what you are doing', and if we allow that then their conclusion follows directly. Note that this is much more than the familiar finding that difficult forced-choice tasks may yield fairly high performance coupled with low confidence. That is no great mystery, and claims for 'subliminal perception' on that basis do not need to be taken very seriously. The important novelty here is the comparison between two closely related conditions: when performance levels, and hence difficulty, were about equal, one condition showed insight and knowledge of performance while the other did not.

It is also important that the dissociation between performance and awareness was not restricted to conditions of dichoptic presentation. In experiment 1, with moving dots, no binocular rivalry was involved, but results were similar to those just described. In the critical condition, dots moving in opposite directions (for 50 ms, at the same oblique angle) were paired quite closely in space, for both the target and background regions, while in the standard condition dots moving in opposite directions were more widely separated. This apparently innocuous change has a strong impact on motion perception (Qian et al 1994), and in the critical, paired condition Kolb and Braun reported that flicker, not motion, was seen. What links the critical conditions here and in experiment 2? My guess is that both involve a breakdown in figure-ground segmentation between target and background regions, yielding no *visible* target object. The mystery is that, in spite of this, performance on the task was not impaired.

## 2. Paradigm-breaking: new views of vision and action

What, then, is the paradigm that such puzzling results might break? It is no less than the 'sensory metatheory of mind': 'Since the time of Aristotle the mind has been regarded as intrinsically sensory in nature, as a passive black box or window [...] Vision, conceived as the passive reception of information that both exists and possesses an intrinsic psychological character independently of the organism, became the paradigm exemplar of mental processing. At the same time, behaviour, or motor activity, became divorced from mental activity and was seen as a consequence [...] of [...] mental events' (Weimer 1977, page 268). Thus Marr (1982), who did more than most to shape our current conceptions of vision, agrees with 'the plain man's answer and Aristotle's too' that to see is 'to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is' (page 3). And again: '... the quintessential fact of human vision [is] that it tells about shape and space and spatial arrangement' (page 36).

Of course, when pressed further, all workers in the field of perception will concede that the purpose of vision is to support cognition (recognition, comparison, thought, memory, language ...) and action. But few of us feel it necessary to involve real actions in our experiments, and the topic of motor control is regarded as a separate discipline, of which we are largely ignorant. Perceptual representations of the spatial world are the basis for action, but can be studied in isolation from action. Thus to study (say) perceived depth of objects, or even perceived direction of heading in self-motion, the observer can sit passively in front of a computer graphics simulation. S/he does not actually have to reach out for the object, or move through the world, to reveal the visual processes and representations on which those actions would be based. This separation of effort is justifiable only as long as the sensory metatheory is sustainable. The thrust of Milner and Goodale's (1995) argument is that it is not. Vision is more than the representation of spatial information.

They argue instead that vision is based on *two* broadly separate systems of processing whose representations and purposes are different, only one of which should be regarded as 'perceptual'. The other is concerned with visuomotor guidance of actions. Evidence for this thesis is carefully marshalled from neuroanatomy, neurophysiology, neuropsychology, eye movement studies, and psychophysics, with special emphasis on the dissociations between perception and the visual control of action that may be observed, in both monkeys and humans, after damage to different visual areas of the brain.

The idea of 'two visual systems' is of course well-known and has been around for at least three decades. Schneider (1969) argued from the behavioural effects of selective lesions that (for rodents) the cortex and colliculus constituted separate systems for 'what' and 'where', ie for pattern discrimination or object recognition and spatial localisation respectively. This functional distinction has been largely preserved in more recent thinking about primate vision, albeit translated into two cortical systems or 'streams' by the formulation of Ungerleider and Mishkin (1982). The ventral stream, passing from striate cortex (V1) through V2, V3, V4 into

the inferotemporal cortex, is associated with pattern perception and object recognition, while the dorsal stream, passing from V1 through V2, V3, V3A, MT, MST etc into the parietal cortex, was associated with localisation of objects in 3-D space around the animal, and more recently with motion perception and the analysis of optic flow. Thus the two streams together could indeed "know what is where by looking".

Milner and Goodale's (1995) theory amounts to a radical reworking of the two-systems idea, because "... the division of labour in cortical visual systems is based on a distinction between the requirements for action and perception and ... this distinction cuts right across any distinction between spatial and object vision" (pages 42--43). "One stream [ventral] is concerned only with the world 'out there' independent of the observer, while the other [dorsal] is concerned only with the observer's actions within that visual world" (page 63). In summary, the theory holds that "both cortical streams process information about the intrinsic properties of objects and their spatial locations, but the transformations they carry out reflect the different purposes for which the two streams have evolved. The transformations carried out in the ventral stream permit the formation of perceptual and cognitive representations which embody the enduring characteristics of objects and their significance; those carried out in the dorsal stream, which need to capture instead the instantaneous and egocentric features of objects, mediate the control of goal-directed actions" (pages 65--66).

The evidence used to construct this new model of vision is too extensive to summarise here, but involves detailed consideration of the physiological properties of neurons in different brain areas along the two streams, as well as neuropsychological evidence from "blindsight" (which Milner and Goodale interpret as 'action without perception') and disturbances of visual control, such as optic ataxia. Here patients with parietal lesions are unable to reach out for seen objects appropriately, even though visual localisation in the affected visual field, and reaching in the intact visual field, are both normal. It is visual guidance of action which is disturbed, and Milner and Goodale build the argument that most of the evidence linking parietal cortex with spatial localisation can be better understood in terms of visuomotor function. There was a striking double dissociation between the patient RV (optic ataxia; parietal lesions) and patient DF (visual form agnosia; presumed ventral stream lesions). Both observers were presented with the same set of smoothly contoured, but oddly shaped 'poker chips'. Patient RV could discriminate visually between the different shapes, but could not use that information to form an appropriate grip between finger and thumb to pick the shapes up. Patient DF was just the reverse; she could not discriminate the shapes in a visual task, but could use detailed shape information to grasp them appropriately, like normal observers. Thus for DF the shape information denied to perception is available for action, and vice-versa for RV. This double dissociation is one of the key findings that support the new two-systems theory.

Dissociations between the visual information available to perception and to action are also evident in normal subjects. Along with a host of studies on saccadic eye movements (too detailed to review here), Milner and Goodale report a striking result, quite central to the interests of this journal, on an illusion of perceived size. In the 'Titchener circles', surrounding a test circle with larger circles makes the test circle appear smaller, and vice-versa. However, even as observers were in the act of judging one test disk to be larger than another (by picking up the apparently larger one), the aperture between finger and thumb used to pick up the disk revealed that they also had information about the true size of the object. Thus, in the course of a single act, perception and action were based on different codes for object size.

Milner and Goodale argue more generally that the ventral stream develops object-centred representations suitable for long-term storage, recognition, constancy, and generalisation; the dorsal stream develops viewer-centred representations that are needed, here and now, for on-line interaction with objects at particular distances and orientations. Finally, Milner and Goodale cautiously but specifically attribute visual awareness to the activity of the ventral (perceptual) stream, while denying it to the dorsal (visuomotor) stream. Activation of the dorsal stream leads to action, not perception.

This has been of necessity a very condensed account of their argument, and in fact Milner and Goodale blur the sharp two-system dichotomy with some interesting ideas about crosstalk between the two streams, for example in the idea that area MT, in the dorsal stream, may relay motion information to V4 in order to make it available to perception. I've concentrated in this brief review on the Big Ideas that the new theory offers, and suggested that it may constitute a new paradigm for vision research. Like any new paradigm it provides an incomplete framework that needs to be filled-in by painstaking 'normal science'.

### **3. Implications for perception and Perception**

The scope for new work within this framework is clearly enormous, and work with normal subjects will be

particularly important in establishing the extent to which the two visual systems, dissociable by lesions, can cooperate and inform each other during the course of actions by the intact observer. But what are the implications for our approaches to perception and psychophysics? No doubt they are many, but one concerns the choice of appropriate methods for investigating perceptual and action-directed visual codes. Forced-choice methods, long favoured for their rigour and objectivity, push the observer to use any cue that enables correct performance on a task. We can now see that this cue may or may not be perceptual. Kolb and Braun's experiment, which I began with, appears to be just such a case. The experiment is psychophysical, and involves no lesions, and no reaching or grasping actions. Yet, as we have seen, the strong implication is that performance in the standard conditions was perceptual, while in the critical conditions it was not. In the framework of Milner and Goodale, my hunch is that the critical conditions defeated binocular, second-order, figure-ground segmentation in the ventral pathway, hence eliminating visual awareness of any target shape, while allowing some effective segmentation in the dorsal path, based upon different cues (for example, monocular cues in experiment 2, or non-opponent motion signals in experiment 1). These could be sufficient to program a saccade to the target location. Subjects would then find themselves looking at the target quadrant but not seeing anything; hence they could respond correctly (on the basis of gaze direction) but would not know why they were doing so. This is 'action without perception', and it would account for the dissociation between confidence and performance in the critical conditions. Further work, perhaps monitoring eye movements, is needed.

A more general implication from all this is that if we wish to study perception (and publish in *Perception*?) then we must use tasks in which observers actually see something. If not, they may be using visual, but nonperceptual, cues. It follows that we can either use explicit perceptual instructions, criterion-dependence and all, or employ forced-choice methods backed by confidence ROCs in the manner of Kolb and Braun. Crudely, when  $d'$  values for performance and confidence agree, that's perception. Conversely, it may be possible to study visual processing in the nonperceptual path specifically by showing that  $d'$  (confidence) is much less than  $d'$  (performance). In any event, these are exciting new developments, and if they really are paradigm-breaking then I (not very confidently) predict that in ten years' time this journal will be called *Perception and Action*.

Mark Georgeson

School of Psychology, University of Birmingham, Birmingham B15 2TT;  
e-mail: [m.a.georgeson@bham.ac.uk](mailto:m.a.georgeson@bham.ac.uk)

**Note added in proof:** After this was written, Morgan et al (1977) described several unsuccessful attempts to replicate Kolb and Braun's (1995) findings with dichoptic textures. In the experiment that was closest to Kolb and Braun's conditions, performance was much worse in the critical condition than the standard condition, and confidence was correlated with performance. This could be interpreted as the signature of ventral, rather than dorsal, stream performance. The reasons for this empirical discrepancy are unclear, but further work looking at the effects of instructions, level of practice, and monitoring eye movements could be revealing.

## References

- Kolb F C, Braun J, 1995 "Blindsight in normal observers" *Nature (London)* **377** 336--338
- Kuhn T S, 1962 *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press)
- Marr D, 1982 *Vision: a Computational Investigation into the Human Representation and Processing of Visual Information* (San Francisco: W H Freeman)
- Milner A D, Goodale M A, 1995 *The Visual Brain in Action* (Oxford: Oxford University Press)
- Morgan M J, Mason A J S, Solomon J A, 1997 "Blindsight in normal subjects?" *Nature (London)* **385** 401--402
- Qian N, Andersen R A, Adelson E H, 1994 "Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics" *Journal of Neuroscience* **14** 7357--7366
- Schneider G E, 1969 "Two visual systems: brain mechanisms for localization and discrimination are dissociated by tectal and cortical lesions" *Science* **163** 895--902
- Ungerleider L G, Mishkin M, 1982 "Two cortical visual systems", in *Analysis of Visual Behaviour*

Eds D J Ingle, M A Goodale, R J W Mansfield (Cambridge, MA: MIT Press)

Weimer W B, 1977 "A conceptual framework for cognitive psychology: motor theories of mind" Chapter 10, in *Perceiving, Acting & Knowing: Toward an Ecological Psychology* Eds R Shaw, J Bransford (Hillsdale, NJ: Erlbaum Associates)

Wolfe J M, 1983 "Influence of spatial frequency, luminance, and duration on binocular rivalry and abnormal fusion of briefly presented dichoptic stimuli" *Perception* **12** 447--456

[Return to main contents](#)

© 1997 Pion Ltd