

Of Two Minds About Two Visual Systems

Oliver H. Turnbull

School of Psychology
University of Wales, Bangor
Gwynedd, LL57 2DG
U.K.

o.turnbull@bangor.ac.uk

Copyright (c) Oliver H. Turnbull 1999

PSYCHE, 5(8), May 1999

<http://psyche.cs.monash.edu.au/v5/psyche-5-08-turnbull.html>

KEYWORDS: two visual systems, object recognition, visuo-spatial, agnosia, neuropsychology.

COMMENTARY ON: A. David Milner & Melvyn A. Goodale. (1995) *The Visual Brain in Action*. (Oxford Psychology Series, No. 27). Oxford: Oxford University Press. xvii + 248pp. ISBN: 0198524080. Price: \$35 pbk.

ABSTRACT: Milner and Goodale's work stands squarely at the centre of the debate on 'two cortical visual systems', and is often seen as offering a reinterpretation of the classic Ungerleider and Mishkin (1982) account of visual specialisation. Several 'problems' with this original account are presented, including reviewing the work in the context of the history of human neuropsychology. However, it is argued that Milner and Goodale's reinterpretation offers much that is advantageous to our understanding of extra-striate cortical specialisation. In particular, the possible role of the '3rd stream' is reviewed, including discussion of work which casts doubt on the claim that object recognition is restricted exclusively to the ventral stream. Rather, the 3rd stream (the inferior parietal lobule) is presented as offering an optional visuo-spatial resource - which would be used for object recognition under non-optimal circumstances. Such an interpretation might offer a compromise between the two competing 'two visual systems' accounts.

Milner and Goodale's (1995) book stands squarely at the centre of the debate on 'two cortical visual systems' - the term used in the original Ungerleider and Mishkin (1982) article to carve up much of the neural basis of vision. The core of their argument was, of course, that visual information arrives in the primary visual areas, and is then parcelled out to two extra-striate regions: an occipito-temporal system concerned with object recognition (the ventral, or 'what' stream) and an occipito-parietal system concerned with spatial abilities (the dorsal, or 'where' stream). This 'two visual systems' model has become *the* catch-phrase in visual neuroscience for the last dozen-or-so years - and has its staunch adherents and vociferous opponents. I must confess that, strangely, I am both an

adherent *and* an opponent - and my reasons for holding these views reflect intimately on the position taken by Milner and Goodale.

There are many reasons for people's opposition to the 'two visual systems' concept - for example the demonstration that the systems are not *entirely* anatomically independent, or that the complexities of the visual system are undersold when we use simple dichotomous labels (see Zeki, 1993, Chapter 20 for a readable critique). I too have an objection to the concept, which I feel I share with others, but which never seems to appear in print. As someone working in traditional human neuropsychology, I have always felt that the original version of the 'two visual systems' account (i.e. Ungerleider and Mishkin, 1982) seemed to offer little that was genuinely *new* to my understanding of neuropsychology.

I must confess that I was surprised by the impact that the Ungerleider and Mishkin 'revelation' had upon visual neuroscience - I wondered (and still wonder) if perhaps I'd missed something? Certainly, that classic paper describes excellent lesion studies on animals. However, in *functional* terms, and in human neuropsychology, the essence of these findings was described many years ago (for example, see De Renzi, 1982; Brown, 1988; Grusser & Landis, 1991; Code et al, 1996, or any other text with a sense of history). Knowledge of the location of the primary visual areas dates back well over a century, as do descriptions of a ventral lesion site causing visual agnosia (the prototypical disorder of the 'what' system). Links between a range of spatial abilities and the parietal lobes (the 'where' system) were made at various times in the first half of this century - names such as Kleist, Poppelreuter and Zangwill spring to mind. Indeed, in the same year that the original Ungerleider and Mishkin paper was published, De Renzi (1982) was able to write:

by the early 1950's... well informed neurologists were aware that a brain lesion could produce a selective disruption of spatial skills, quite independently of the impairment of elementary visual functions and of other forms of visual agnosia (p.3).

Thus, from the perspective of traditional human neuropsychology, the Ungerleider and Mishkin account does not appear especially novel. However, Milner and Goodale's work offers a fine example of potential *new* interpretations that this 'two systems' argument might bring to human neuropsychology.

In part, this is because Milner and Goodale suggest that a better description of the function of the dorsal visual system is the control of visuo-motor behaviour - i.e. action. Their first, and most elegant, demonstration of this was that DF (a visual agnostic patient) was unable to use information on a visuo-*perceptual* task: being unable to match the orientation of a card to a slot of variable orientation. However, she was remarkably accurate in her abilities using the same visual information on a visuo-*motor* task: being able to post the card into the same slot. The widely touted outcome of this finding has been to re-cast the dichotomy between object recognition and spatial abilities (i.e. 'what' versus 'where') as a dichotomy between object recognition and visuo-motor abilities (i.e. between 'what' and '*how*' - if you prefer this user-friendly nomenclature). In essence, the

Milner and Goodale argument has been seen as overthrowing, or at least competing with, the traditional Ungerleider and Mishkin (1982) position. However, my main reason for admiring the Milner and Goodale account has more to do with my earlier complaints about the original theory - that it doesn't say anything new.

Is it then that Milner and Goodale have offered a better localisation of the neurological basis of visually-guided action? Certainly not. Their patient DF offers (by their own admission) rather inadequate anatomical data, and besides, the likely importance of the occipito-parietal region for visually-guided action was clearly outlined by Balint almost 100 years ago (De Renzi, 1982, or De Renzi's chapter in Code et al., 1996). Rather, the novelty of the Milner and Goodale account is that it offers a radically different interpretation of the unity of perceptual experience. The obvious implication of the DF data (and of the obverse dissociation, seen in patient RV) is that the object recognition and action systems are not mere areas of 'specialisation', but independently operating *channels*. Thus, the observer-object relationship is not simply that which all intuitively believed: a single perceptual system, with separate regions for specialist processing. Rather, we have (at least two) apparently independent visual 'worlds', running in parallel - visual worlds that commonly coincide, but may often offer different versions of reality.

This line of argument is bolstered by the fascinating findings on the Titchner illusion in neurologically *normal* subjects (Milner & Goodale, 1995, pp.167-170), showing that even though our perceptual system is 'fooled' by the illusion, our motor system is not. We might certainly *think* that we see the world in a unified way - the way I do now, as I look around me. Incredibly, however, it seems that I see the world in *different* ways, depending on which *task* I am engaged in - depending upon whether I am recognising my mug or reaching for it. To remind us that such a stark division of labour is not unprecedented, Milner and Goodale ask us to recall that this kind of specialised visual 'channel' approach has clear antecedents in the animal literature. To take an example of channels within the domain of action, they review Ingle's work with frogs (pp.6-11) showing separation of the neural substrate of visually-guided feeding behaviour, versus visually-guided locomotion.

The Milner and Goodale reinterpretation of the 'two visual systems' account comes primarily from their larger-scale re-appraisal of the importance of action as the critical output for vision - indeed, action driven in an almost 'Gibsonian' manner. Thus, in their terms, a hundred years of psychology have been misguided because of the "phenomenological intuitions that regard vision as a purely perceptual phenomenon" (p.13). This is the (widely held) idea that the goal of vision is to form some kind of idealised 'representation' of the world, which we may (or may not) choose to act upon. It is what I once heard Irving Biederman call the 'couch potato' model of vision: a visual system requiring no more direct interaction with the world than that needed to press the remote control button. Evolution, as Milner and Goodale are always keen to remind us, operates only at the level of direct action with the world: death follows for the primate whose 'representation' of a branch may be beautiful, but is geometrically inaccurate. In this regard, their gripe is not merely with faculty psychology, or with the modern cognitivists, but also with behaviourists - who minded not whether responses were

measured by direct interaction with the relevant object, or by responses which were no longer yoked intimately to the reinforcer, such as bar-presses or pecks.

Because they are so vexed by the historical link between vision and *perception*, their work stresses the importance of action to an almost messianic degree. In a discussion with Mel Goodale a few years ago (on an Italian ski-lift, strangely enough) we spoke of an ex-student of his who had taken religious orders. Did this mean, I asked Mel, that he - like Sperry and Eccles - might succumb to that ill-defined mysticism that often overtakes neuroscientists in late career? And what form might this take? With a certain inevitability, we cast an argument in terms of two visual systems. It seemed to us both that, if pushed, God might well wish to exert his influence on the world through the *dorsal* stream, where his manipulations would go unnoticed by consciousness. What then of the ventral stream, I asked? With mock seriousness, Mel chose to display his contempt for those who have viewed vision solely as a vehicle for passive perception, responding: "The ventral stream is the Devil's playground"!

If the role of the dorsal stream is that of visually-guided action, where does that leave the spatial abilities that Ungerleider and Mishkin have attributed to the dorsal stream? The answer of Milner and Goodale has been some tentative suggestions about a possible 'third' stream (seen depicted below in Figure 1). The substrate of this system would be the *inferior* parietal lobule (leaving the superior parietal lobule to visually-guided action), making the functions of the 'third' stream a postulate that is broadly consistent with the list of 'spatial' abilities (in the widest sense of the term) that are impaired after lesion to the inferior parietal lobule of humans.

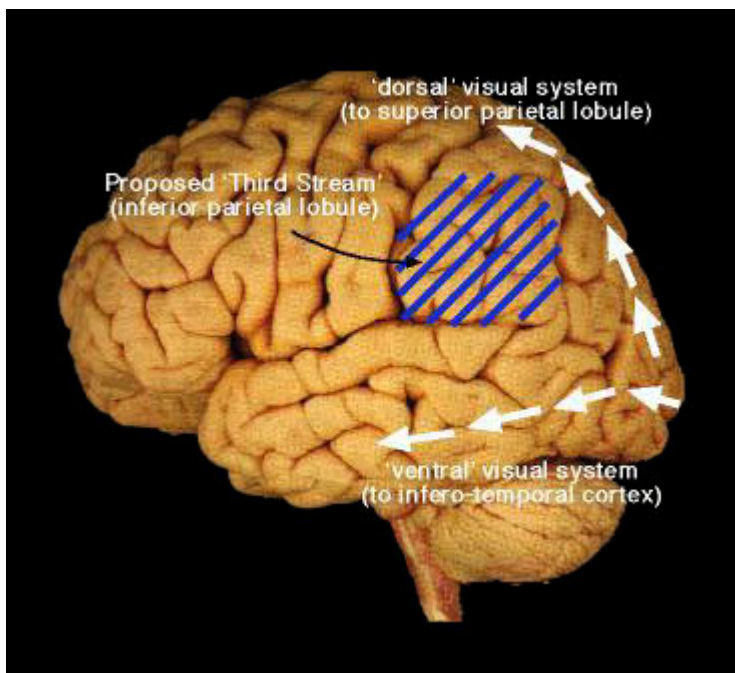


Figure 1

A tentative diagram of the various 'streams' in the human brain. The specifics of this diagram are highly speculative - partly because the original 'two visual systems' model was developed from work in non-human primates, where the tools for localisation studies (in neurophysiology and lesion work) are far more precise than those currently available in humans. This is likely to change as the quality of research in functional imaging develops.

However, Milner and Goodale have been rather reticent about detailing the possible role of this novel system - and seem content with somewhat ill-defined notions about it operating in the 'interaction' between the classic dorsal and ventral streams. Recently, we (Turnbull, Carey, & McCarthy, 1997) have tried to flesh out some of these ideas, with respect to the task of object recognition. Traditionally, object recognition (i.e. the 'what' system) has been seen as an occipito-temporal, or ventral system, function. I have no quarrel with the narrow sense of this interpretation - and I feel comfortable with the idea that the rapid, effortless, and largely automatised recognising of objects, that we accomplish dozens of times every minute of our day, is run by this system. We have also accumulated some evidence (see Solms et al., 1998; Turnbull & McCarthy, 1996; Turnbull, Beschin, & Della Sala, 1997) that this system may well find it convenient to down-play, perhaps even ignore, certain kinds of visual information - such as orientation, distance and size - that are largely irrelevant for a recognition system that strives to achieve object constancy. In other words, in *normal* object recognition, it may be an advantage for the ventral system to ignore many kinds of spatial information. Why should an object recognition system care about the orientation and distance of my mug? This information is more-or-less unimportant for recognising the object - because the identity of my mug does not change when it is rotated, or when I move closer or further from it. Note, however, that such spatial information would be *crucial* if I was trying to act upon the object.

Thus far, this sounds more-or-less like the classic 'two visual systems' account - albeit embellished by additional detail on the precise function of the ventral system. However, it is also clear from human neuropsychology that areas *outside* of the classic ventral stream seem to be involved in object recognition - although only under what we have referred to as 'non-optimal' circumstances. These would include objects seen from 'unusual views' (where the principal axis is fore-shortened or features are occluded), objects seen under unusual lighting conditions, objects seen only in silhouette, objects which appear as fragmented images (i.e. being partially occluded), and perhaps even objects seen in circumstances requiring mental rotation (Warrington & James, 1986; Warrington & Taylor, 1973). The critical lesion sites in patients with such deficits appears to be the inferior parietal lobule - that is, arguably, in the spatial ('where') system (see Turnbull, Carey, & McCarthy, 1997, for review).

How might the disorders seen in these patients map onto the classic Ungerleider and Mishkin (1982) view of the organisation of spatial abilities? Two points are of clear interest. Firstly, my use of the term 'spatial' to describe the disorders of these patients refers to a far more sophisticated form of spatial ability than that which Ungerleider and Mishkin (1982) had in mind. Their 'where' system was presumed capable of no more than knowledge of simple spatial location, while it seems that the human inferior parietal lobule is involved in tasks of complex spatial transformation and manipulation. The

second point is perhaps even more important. The performance of these patients suggests that the inferior parietal lobule is involved in recognising objects - at least under certain circumstances. This makes these cases inexplicable in the Ungerleider and Mishkin (1982) 'two visual systems' account - we are speaking of disturbances of object recognition (the 'what' stream) after a lesion in the 'spatial' system (the 'where' stream).

Might the Milner and Goodale 'third' stream proposal solve this problem? I suspect that it might. Our explanation (Turnbull, Carey, & McCarthy, 1997) has been to remind ourselves that, although object recognition might *normally* be best served by ignoring spatial information such as orientation and size, spatial information is useful for object recognition under non-optimal circumstances. Thus, we have suggested that the ventral stream is the primary route to recognition, but that (contrary to the classic 'two visual system' account) there is a substantial contribution to the process of object recognition from *outside* the ventral stream - though only under non-optimal circumstances. This contribution would be made by the *inferior* parietal lobule - Milner and Goodale's 'third' stream - leaving the superior parietal lobule (i.e. the dorsal stream *proper*) free for the purposes of action, as specified by Milner and Goodale.

What then would be the properties of this 'third' stream ? If it can be described as a 'spatial' system, it clearly differs dramatically from the sense in which a visually-guided action system might be said to be 'spatial'. For example, visually-guided action requires a set of descriptions of objects coded exclusively in egocentric terms - whether the basis for such coding is the eyes, the trunk, or the arm and hand. It does not code for allocentric space - so that it can tell *you* how you should reach for an object, but not how your friend should reach for the object. Also, if we may speak in information processing terms, this information cannot be 'stored' or 'manipulated', but changes on-line. The spatial abilities of the 'third' stream (if we base our interpretation on the deficits of neurological patients with inferior parietal lesions) seem entirely different. Disorders such as constructional apraxia, disturbances of mental rotation ability, and the failure of such patients on recognition tasks where objects are fragmented, poorly lit, or appear in silhouette, suggest that in this system spatial information can be stored and manipulated, and that it is possible to code for location in allocentric space. In terms of object recognition, this system might perhaps be serving to re-organise and normalise an unusual or 'noisy' visual image, in order for another attempt to be made at object recognition - presumably again by the ventral system. (However, such a 'visuo-spatial' system would clearly be useful in a wide range of other situations - including assisting a visually-guided action system under some circumstances). Hopefully this account of the object recognition process offers some additional ideas about the function of the ill-defined 'third' stream - especially as regards the kinds of benefits that might be derived for object recognition from spatially sophisticated systems *outside* the classic ventral stream.

Thus, I feel that our version of the cortical organisation of visual function might well offer a comfortable compromise between the original Ungerleider and Mishkin and the revised Milner and Goodale accounts. In this version, part of the parietal lobe (its superior component) *is*, as Milner and Goodale suggest, dedicated to visually-guided action. Another parietal region (its inferior component) *is*, as Ungerleider and Mishkin

seem to suggest, dedicated to spatial abilities. This seems a happy compromise though, as mentioned above, the Ungerleider and Mishkin interpretation of a 'spatial' ability is a good deal simpler than that which I have suggested. Certainly, this interpretation might explain why the two competing two visual systems models seem to be talking past each other, each thinking that the other has failed to grasp some basic point.

However, this does leave a few problems with the rather glib nomenclature of 'what' and 'where'. As far as I am concerned, the occipito-temporal regions might well be argued to retain the 'what' label, while the superior-parietal system might seem best described by Milner and Goodale's 'how'. But what of the region in between - the 'third' stream? A suitable label eludes me. The label 'where' (Ungerleider and Mishkin's choice) *does* describe some of the spatial characteristics of the 'third' stream, though it fails to capture the active role of the system - its capacity for visuo-spatial manipulation. However, and most importantly for me, it ignores the fact that the inferior parietal lobule *also* makes a contribution to the 'what' process - a term that has already been allocated to the ventral stream. It's a system that can do both 'where' and 'what', or at least make a contribution to each. Now I've tried as many other adverbs as I can think of for this system - but calling it the 'why', 'who' or 'when' system is simply never going to be appropriate... Perhaps we should just let these little labels go - sometimes the brain is a touch too complicated for the simplifications of our language.

References

- Brown, J.W. (1988). *Agnosia and apraxia: Selected papers of Liepmann, Lange, & Potzl*. Hillsdale, New Jersey: Lawrence Earlbaum Associates.
- Code, C., Wallesch, C-W., Joannette, Y. & Lecours, A.R. (1996). *Classic cases in neuropsychology*. Hove: Psychology Press.
- De Renzi, E. (1982). *Disorders of space exploration and cognition*. Chichester: Wiley.
- Grusser O-J. & Landis, T. (1991) *Visual agnosias and other disturbances of visual perception and cognition*. London: Macmillan Press.
- Milner, A.D. & Goodale, M.A. (1995) *The visual brain in action*. Oxford: Oxford University Press.
- Solms, M., Turnbull, O.H., Kaplan-Solms, K. & Miller, P. (1998). Rotated drawing: The range of performance, and anatomical correlates, in a series of 16 patients. *Brain and Cognition*, 38, 358-368.
- Turnbull, O.H., Beschin, N. & Della Sala, S. (1997). Agnosia for object orientation: Implications for theories of object recognition. *Neuropsychologia*, 35, 153-163.

Turnbull, O.H., Carey, D.P. & McCarthy, R.A. (1997). The neuropsychology of object constancy. *Journal of the International Neuropsychology Society*, 3, 288-298.

Turnbull, O.H. & McCarthy, R.A. (1996). Failure to discriminate between mirror-image objects: A case of viewpoint-independent object recognition? *Neurocase*, 2, 63-72.

Ungerleider, L.G. & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale & R.J.W. Mansfield (Eds). *Analysis of Visual Behavior*. Cambridge, MA: MIT Press. pp.549-586.

Warrington, E.K. & James, M. (1986). Visual object recognition in patients with right-hemisphere lesions: Axes or features? *Perception*, 15, 355-366.

Warrington, E.K. & Taylor, A.M. (1973). The contribution of the right parietal lobe to object recognition. *Cortex*, 9, 152-164.

Zeki, S. (1993). *A vision of the brain*. Oxford: Blackwell.