

When does action resist visual illusions?

Nicola Bruno

Over the past decade, many studies of non-pathological individuals have reported functional dissociations between perceptual judgments and motor responses. These results suggested an interpretation of the ventral and dorsal streams in the primate visual system as independent modules for visual awareness and the visual guidance of actions. However, recent comparisons of perception and grasping responses in size-contrast displays have been widely reported to show that apparent dissociations are actually experimental artifacts. An overview of the literature suggests that the issue of visuomotor dissociations in healthy individuals is far from settled. Some results suggest that immunity from visual illusions might be found when task requirements emphasize observer-relative reference frames. These results suggest that the functional specialization of the two visual subsystems might be less rigid than originally posited.

The primate visual system shows two broad streams of projections from the primary visual area: a ventral projection to the inferotemporal cortex and a dorsal projection to the posterior parietal cortex. Although the existence of separate pathways for higher level visual processing has been known for some time, their functional interpretation is still debated. An earlier view¹ hypothesized that the ventral projection subserved object identification (the 'what' function) and the dorsal projection was concerned with spatial processing ('where'). More recently, this view has been questioned by the idea that the ventral projection might serve visual awareness, whereas the dorsal projection might be primarily concerned with the nonconscious control of actions². Originally motivated by the study of perception-action dissociations in neuropsychological patients³, this characterization of visual functions has proved attractive to a broad audience of vision scientists, in part, presumably, because it promises to reconcile long-standing theoretical differences concerning the role of explicit representations in vision⁴ (Box 1).

'The crucial factor might be the selection of a specific frame of reference...'

In the context of broader theoretical implications, widespread interest has been given to functional dissociations between perceptual and motor tasks in neurologically healthy individuals. Such dissociations are directly predicted by the perception-action interpretation of ventral and dorsal functions⁵. To support action, spatial maps must emphasize relationships between

an observer's body and external objects rather than relationships between external objects alone, they must be available rapidly, but they need not be maintained in memory as the unfolding of the action renders them quickly obsolete. Conversely, to support conscious perception, spatial maps must emphasize relational invariants that mediate object constancy under changes of the viewpoint. These include relationships of an object to its surroundings, rather than the observer's body, and temporal relationships underlying phenomenal object constancy rather than short-lived representations. Thus, tasks that elicit different spatial or temporal frames of reference for perception and action should result in different measurable outputs from exactly the same visual information. Several studies of more or less comparable perceptual and action tasks have reported results that are generally consistent with this prediction. For example, observers show systematic biases in perceptually evaluating distance but, when asked to walk the same distances wearing a blindfold, they are essentially accurate⁶. Observers placed at the bottom of a hill overestimate hill slope, but their motoric adjustments of an unseen hand paddle to match the same slope are much less prone to the illusion⁷. Delaying an action can shift size-scaling responses from absolute to relative metrics⁸.

Contradictory grasps?

Although many of these findings remain relatively uncontroversial, contradictory results have been recently reported concerning the existence of functional dissociations between visually controlled grasping and perceptual estimates of size in the Ebbinghaus-Titchener size-contrast display (Fig. 1a). In the Ebbinghaus-Titchener display, the same central disk appears slightly larger when surrounded by smaller circles and slightly smaller when surrounded by larger circles. A much-cited study by Aglioti *et al.*⁹ reported that when actors actually picked up the central disks using their thumb and forefinger, the maximum grip aperture during the action was highly correlated with the physical size and did not show the size-contrast effect. Along with later similar observations¹⁰, the finding of Aglioti *et al.* has been widely cited as evidence supporting the perception-action interpretation of ventral and dorsal functions. However, other studies have recently suggested that the seeming dissociation might, in fact, be an artifact resulting from inadequate matching of experimental conditions.

At the core of the issue is the nature of the perceptual and grasping tasks. Aglioti *et al.* asked their observers to compare the two disks directly to perform the perceptual size estimates. However, corresponding motoric instructions directed their actors' attention to one of the disks at a time. Thus, their grasping task was not fully comparable to the perceptual task in that it did not require a simultaneous appreciation of two surface sizes relative to their contrasting surrounds. When the perceptual task was modified to remove such

Nicola Bruno
Dipartimento di
Psicologia, Università di
Trieste, via S. Anastasio 12,
34143 Trieste, Italy.
e-mail: nicola.bruno@
univ.trieste.it

Box 1. Koffka's question

According to Gestaltist Kurt Koffka^a, the fundamental question of vision science was the following: 'Why do things look as they do?' As Koffka himself recognized, this question can be understood in two different ways. The first is phenomenological: why is it that experience possesses certain qualities in our conscious awareness? The second is functional and adaptive: 'how is it that percepts veridically map onto objects in the external environment?' For Koffka, and for many other theorists^b, the phenomenological question came first and led to constructivism, that is, the idea that conscious awareness is 'built' by internal processes of representation. Constructivism was forcefully challenged by James J. Gibson's 'direct perception' approach^c. For direct theorists, vision is to support action in the environment by simply picking up optical information. The two positions have been hotly debated^d. However, constructivist and direct theories might both find their scope if visual function consists of a subsystem concerned with building conscious experience and a subsystem to control action independently of visual phenomenology.

References

- a Koffka, K. (1935) *Principles of Gestalt Psychology*, Harcourt Brace Jovanovich
- b Rock, I. (1983) *The Logic of Perception*, MIT Press
- c Michaels, C.F. and Carello, C. (1981) *Direct Perception*, Prentice-Hall
- d Fodor, J.A. and Pylyshyn, Z.W. (1981) How direct is visual perception? Some reflection on Gibson's 'ecological approach'. *Cognition* 9, 139–196

simultaneous appreciation, then the perceptual size-contrast effect reduced and became similar to the effect on grip scaling shown by the maximum pre-shape aperture^{11,12}. These findings have been widely reported as evidence against dissociations between visual judgements and grasping responses in healthy individuals. For instance, a *TICS* Monitor piece¹³ suggested recently that size-contrast illusions do not provide evidence for the perception–action interpretation of ventral and dorsal functions.

Explanation still up for grasps

However, a balanced review of the literature suggests that the issue is far from settled. Two studies^{14,15} showed that the scaling of grip aperture during a prehension movement is unaffected by the converging lines in a 'railroad tracks' pattern (the Ponzo illusion; Fig. 1b), whereas grip force in picking up the object shows the illusory bias. Although the reason for this particular dissociation is not clear, it seems fair to say that the inadequate matching argument does not work here (the measures are different aspects of the same action), as was indeed noted in a recent *TICS* Comment¹⁶. Another study has proposed that seeming illusory effects on grip scaling in the Ebbinghaus illusion result from non-illusory visuomotor mechanisms, such as treating flanker elements as obstacles to be avoided¹⁷. However, why drawn flankers would be treated as actual obstacles is not clear¹⁸. Several studies^{19,20} have suggested that 'actual' actions are immune to illusions, whereas 'mimed' (no real target present) or 'motor estimate' (motor reports of conscious percepts) actions are not. But *a priori* criteria for defining these different types

of motor responses are not obvious. For example, one would classify adjustments of an unseen paddle as a motor estimate response to hill slope. However, such motor measure is largely immune to the overestimation typical of verbal or visual matches⁷. Finally, it has been suggested that movement times in pointing are affected by illusory size contrast in the Ebbinghaus pattern²¹. This result seems to imply that visuomotor dissociations do not extend to the transport component of a grasp. However, a recent reanalysis of pointing movement times suggested that a dissociation can be observed in these conditions provided that the action is not delayed after the display is turned off²².

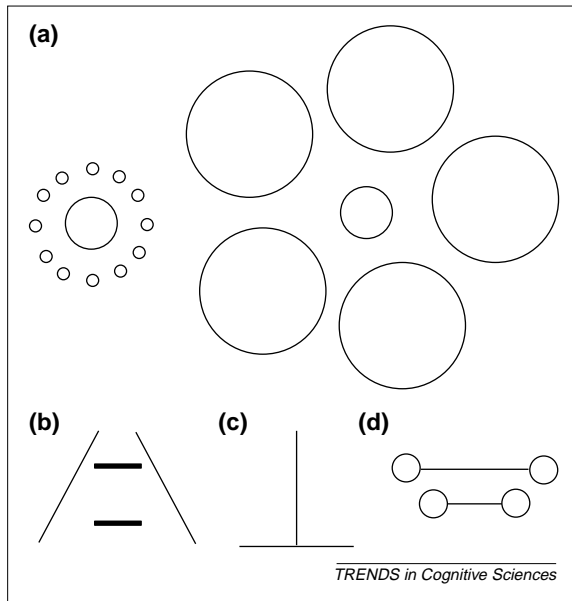
Thus, there is a great deal of disagreement in the literature concerning what is dissociable between perception and action and when dissociations are found. A *TICS* Opinion article¹⁸ recently suggested that most (but not all) grasping studies can be reconciled with the two-visual-systems model as long as one recognizes that some aspects of action might be prone to illusory effects, whereas others remain relatively immune. This review also voiced concerns about using two-dimensional patterns drawn on paper as stimuli for motor responses. Another recent article²³ published in *Current Biology* concluded that it remains unclear whether observed visuomotor dissociations occur because of the operation of two separate visual modules or simply result from different information processing operations within an integrated system.

A frames-of-reference hypothesis

Surprisingly, the current debate has paid little attention to two findings that I consider illuminating for current interpretations of visuomotor dissociations. In a study on the well-known horizontal–vertical illusion (Fig. 1c; the horizontal segment of a 'T' pattern is underestimated relative to the vertical segment), Vishton *et al.*²⁴ performed four experiments. In the first experiment, grip scaling showed essentially no underestimation of the vertical, whereas perceptual size estimates did. In the second and third experiments, however, perceptual estimates also failed to show underestimation, provided that the observer's attention was directed to a single element of the display. Finally, in a fourth experiment, grip scaling showed a marked underestimation of the horizontal when reaching with three fingers to the vertices of a triangular figure. Thus, the horizontal–vertical illusion deceived both the eye and the hand, depending on whether the task emphasized object- or observer-relative ('absolute') metrics.

In another study, Wraga *et al.*²⁵ examined perceptual and blindwalking responses to a walkable, 'dumbbell' version of the Müller–Lyer pattern (Fig. 1d). In one condition, the observer responded while standing on one end of the dumbbell, thereby emphasizing an observer-relative coding of distance. In another condition, they stood away from the dumbbell pattern, emphasizing object-relative coding. Wraga *et al.* found that both numeric estimates of segment length and the extents walked by blindfolded observers

Fig. 1. Size-contrast illusions. (a) Ebbinghaus-Titchener size-contrast illusion, (b) Ponzo illusion, (c) horizontal-vertical illusion and (d) 'dumbbell' version of the Müller-Lyer illusion.



were affected by the position of the hoops at the end of the dumbbells, provided that conditions emphasized object-relative coding. Conversely, emphasizing viewer-relative coding rendered both dependent variables essentially immune to the illusion.

Although this second study looked at walking instead of grip scaling, its results are important in relation to the interpretation of Vishton's results. In Vishton's first three experiments, participants performed grasps to a two-dimensional pattern drawn on paper. In his fourth experiment, they performed a reach with three fingers that requires a somewhat precarious balancing of finger trajectories. It could have been, therefore, that Vishton's results in the first three experiments were because of the nature of the display, whereas his results in the fourth experiment were a consequence of the peculiar kinematics of the action. None of these criticisms applies to Wraga's walking response.

Taken together, therefore, these two findings strongly suggest that current interpretations of ventral and dorsal function might have placed too much emphasis on response mode (perceptual versus motor). Instead, the important distinction could lie in the frame of reference used to perform the task. It remains to be seen whether such frame-of-reference interpretation will survive further analysis. One cannot help noticing, however, that the hypothesis is general enough to accommodate several seemingly contradictory results. For instance, walking from a starting position to a target (an 'actual' action if ever there was one) is a natural candidate for using an observer-relative coding of distance. Hence one would expect no illusory compression and this is what is usually found^{6,25}. Adjusting an unseen paddle to the slope of the hill also encourages evaluating the slope relative to the observer. Therefore, again one would not expect slope overestimation and this is what is found⁷, even though one would classify this as a sort of 'motor estimation' action. Judging the size of a target disk in a

Questions for future research

- Visual perception and visually planned action are dissociable in a variety of tasks. Is this because of modular streams specialized for perception and action, or merely a consequence of multiple spatial maps using different frames of reference?
- Visual perception and visually planned action are not always dissociable. How are tasks yielding perception-action links different from those yielding dissociations?
- Even if dissociable at some stage, visual perception and visually planned action must coordinate at some other stage. Which one and where?

full Ebbinghaus illusion pattern emphasizes size relationships and should produce a size-contrast illusion. Grasping the target disk focuses an actor's attention on that disk only and should produce the same amount of size contrast that is found when only half the Ebbinghaus pattern is presented. This is exactly what Franz *et al.*¹² reported and this account seems to me to offer more promise than the 'flankers as obstacles' hypothesis¹⁷ for accommodating the contradiction with the earlier grasping results of Aglioti *et al.*⁹ There seems to be no compelling rationale for predicting object-relative frames of reference for grip force as opposed to observer-relative frames for grip scaling. However, it seems reasonable to assume that grip force might be more dependent on object characteristics than are observer-relative extents. If this assumption holds true, then the frame-of-reference hypothesis could accommodate even these results^{14,15}.

Conclusion

If proved, visuomotor dissociations in healthy observers would have sweeping implications for perceptual theory²⁶. Although some results have suggested that such dissociations are experimental artifacts, the present review indicates that the issue is, in fact, still not solved. Other results indicate that the crucial factor in visuomotor dissociations might be the selection of a specific frame of reference and not whether observers perform a visual match rather than a visually guided action. If both illusory and veridical responses can be obtained in both perception and action contingent on which frame of reference is emphasized, then the functional interpretation of the ventral and dorsal stream as independent modules is probably too rigid. It would seem more parsimonious to propose a partly interconnected system that uses different frames of reference in different tasks. Even if partly incorrect, the two-visual-system hypothesis proposed by Milner and Goodale² has demonstrated great heuristic value. It will be an exciting task for future research to establish the nature of the interconnections between visual subsystems and to make resulting amendments to our understanding of their functional architecture of the visual system.

Acknowledgements

Supported by Murst grant 9911333852. The author is indebted to the editor and the anonymous reviewers for comments on an earlier draft of this work.

References

- 1 Ungerleider, L.G. and Mishkin, M. (1982) Two cortical visual systems. In *Analysis of Visual Behavior* (Ingle, D.J. *et al.*, eds), pp. 549–586, MIT Press
- 2 Milner, A.D. and Goodale, M.A. (1995) *The Visual Brain in Action*, Oxford University Press
- 3 Milner, A.D. *et al.* (1991) Perception and action in 'visual form agnosia'. *Brain* 114, 405–428
- 4 Bruce, V. *et al.* (1996) *Visual Perception: Physiology, Psychology and Ecology*, Psychology Press
- 5 Goodale, M.A. and Haffenden, A. (1998) Frames of reference of perception and action in the human visual system. *Neurosci. Biobehav. Rev.* 2, 161–172
- 6 Loomis, J.M. *et al.* (1992) Visual space perception and visually directed action. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 906–921
- 7 Proffitt, D.R.P. *et al.* (1995) Perceiving geographical slant. *Psychonomic Bull. Rev.* 2, 409–428
- 8 Goodale, M.A. (2000) Grasping after a delay shifts size-scaling from absolute to relative metrics. *J. Cogn. Neurosci.* 12, 856–868
- 9 Aglioti, S. *et al.* (1995) Size-contrast illusions deceive the eye but not the hand. *Curr. Biol.* 5, 679–685
- 10 Ellis, R.R. *et al.* (1999) The influence of visual illusion on grasp position. *Exp. Brain Res.* 125, 109–114
- 11 Pavani, F. *et al.* (1999) Are perception and action affected differently by the Titchener circles illusion? *Exp. Brain Res.* 127, 95–101
- 12 Franz, V.H. *et al.* (2000) Grasping visual illusions: no evidence for a dissociation between perception and action. *Psychol. Sci.* 11, 20–25
- 13 Snowden, R. (2000) The last grasp? *Trends Cogn. Sci.* 4, 213
- 14 Brenner, E. and Smets, J.B. (1996) Size illusion influences how we lift but not how we grasp an object. *Exp. Brain Res.* 111, 473–476
- 15 Jackson, S.R. and Shaw, A. (2000) The Ponzo illusion affects grip force but not grip aperture scaling during prehension movements. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 418–423
- 16 McIntosh, R.D. (2000) Seeing size and weight. *Trends Cogn. Sci.* 4, 442–444
- 17 Haffenden, A.M. and Goodale, M.A. (2000) Independent effects of pictorial displays on perception and action. *Vis. Res.* 40, 1527–1607
- 18 Carey, D.P. (2001) Do action systems resist visual illusions? *Trends Cogn. Sci.* 5, 109–113
- 19 Otto-de Haart, E.G. *et al.* (1999) More thoughts on perceiving and grasping the Müller–Lyer illusion. *Neuropsychobiology* 37, 1437–1444
- 20 Haffenden, A.M. and Goodale, M.A. (1996) The effect of pictorial illusion on prehension and perception. *J. Cogn. Neurosci.* 10, 122–136
- 21 van Donkelaar, P. (1999) Pointing movements are affected by size-contrast illusions. *Exp. Brain Res.* 125, 517–520
- 22 Fisher, M.H. (2001) How sensitive is hand transport to illusory context effects? *Exp. Brain Res.* 136, 224–230
- 23 Plodowsky, A. and Jackson, S.R. (2001) Getting to grips with the Ebbinghaus illusion. *Curr. Biol.* 11, 304–306
- 24 Vishton, P.M. *et al.* (1999) Comparing effects of the horizontal-vertical illusion on grip scaling and judgment: relative vs. absolute, not perception vs. action. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 1659–1672
- 25 Wraga, M. *et al.* (2000) Perception–action dissociations of a walkable Müller–Lyer configuration. *Psychol. Sci.* 11, 239–243
- 26 Georgeson, M. (1997) Vision and action: you ain't seen nothin' yet. *Perception* 26, 1–6

Language and perceptual categorisation

Jules Davidoff

In a pioneering set of experiments, Rosch investigated the colour processing of a remote traditional culture. It was concluded that colours form universally natural and salient categories. However, our own cross-cultural research, backed up by neuropsychological data and interference studies, indicates that perceptual categories are derived from the words in the speaker's language. The new data support a rather strong version of the Whorfian view that perceptual categories are organized by the linguistic systems of our mind.

Why do category members belong together? Or, put another way, why are category members seen as similar and different from members of other categories? For most categories, it can be concluded that the answers to these questions are determined by theories about the world, rather than perceptual similarity between category members¹. However, for perceptual categories (e.g. colours, facial expressions) the role of perceptual similarity in establishing categories seems more plausible. I will argue that though plausible, it is not perceptual similarity, but rather linguistic similarity that is the critical factor in perceptual categorisation. It has also been argued, in the case of colour, that there are underlying, universal, neurophysiological mechanisms determining

categorisation^{2,3}. I will argue against that view. The arguments in favour of language will draw on neuropsychological and cross-cultural research; these will be reinforced by results from interference studies.

Colour categories are not innate

The proposal for universal colour categories² is held to gain strength from the known properties of wavelength-sensitive neurones⁴. Based on the opponent-process mechanism of neurones in the lateral geniculate nucleus and in V1, it was argued that there are two elemental achromatic categories (black, white) and four elemental colour categories (red, green, yellow and blue)^{3,5,6}. The four colour categories are held to form around natural foci that produce uniquely red, green, yellow and blue sensations. The argument is based on the finding that there are two wavelengths for which opponent-process neurons termed R–G give no output⁴. Similarly, there is a wavelength that corresponds to no output from the other type of opponent-process neurones, termed Y–B. However, the respective wavelengths chosen to correspond to the typical or unique colours of blue, yellow and green do not consistently match the predictions from neurophysiology⁷. In fact, it ought to go without saying that no firm conclusion concerning neurones could really be drawn by asking a person who already has the concept of blue, yellow or green to indicate a colour that is uniquely blue (or yellow or green). Furthermore, the unique colours produced by colour-blind observers do not tally with the predictions made from their altered retinal output⁸. In fact, the neurophysiological data show that neurones simply respond selectively to particular wavelengths⁹, or to combinations of wavelength and brightness¹⁰. Such selectivity is insufficient to allow that the neurones act

Jules Davidoff
Dept of Psychology,
Goldsmiths College,
University of London,
Lewisham Way, London,
UK SE14 6NW.
e-mail:
J.Davidoff@gold.ac.uk