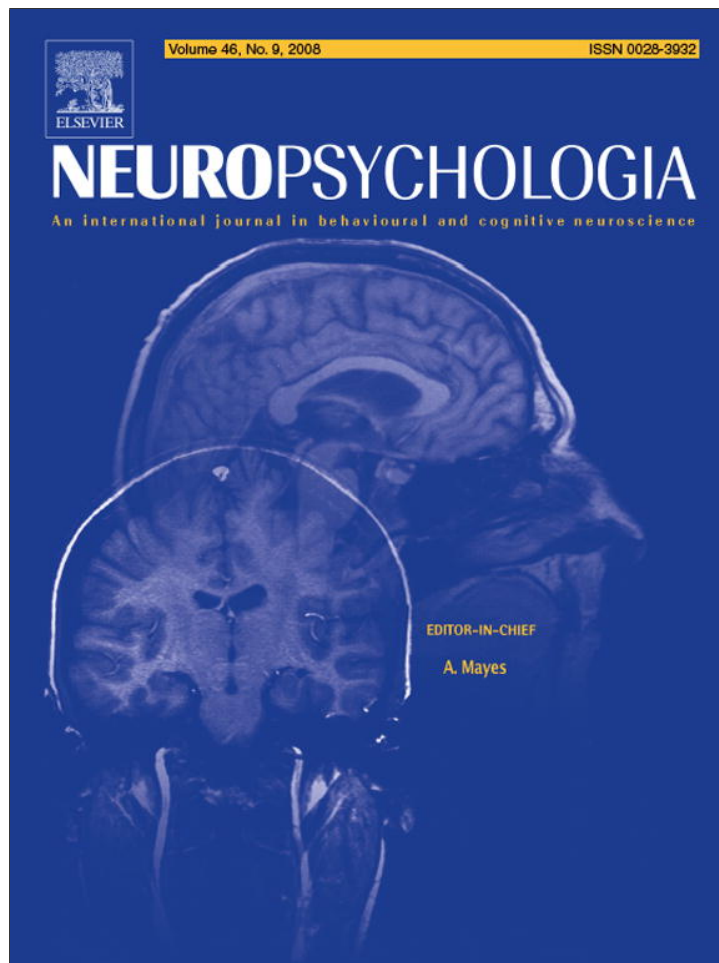


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

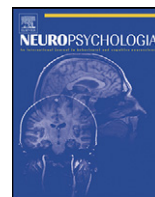
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Note

Matching boxes: Familiar size influences action programming

Robert D. McIntosh*, Gavin Lashley

Human Cognitive Neuroscience, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK

ARTICLE INFO

Article history:

Received 4 December 2007
 Received in revised form 4 March 2008
 Accepted 8 March 2008
 Available online 15 March 2008

Keywords:

Ventral stream
 Dorsal stream
 Perception
 Action
 Depth cues

ABSTRACT

The perception/action model is the dominant account of the primary division of labour in the human visual pathway. Integral to this model is the idea that goal-directed actions are guided spatially by bottom-up vision, independent of perceptual recognition and top-down object knowledge. We question this idea by showing that the expected size of familiar objects (matchboxes) affects the amplitude of reaches made to grasp them, and the pre-shaping of the hand, even when binocular cues are available. This suggests that perceptual recognition routinely influences action programming.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Sixteen years ago, the well-known anatomical split between the ventral and dorsal streams of human vision was functionally interpreted in terms of a distinction between vision-for-perception and vision-for-action (Goodale & Milner, 1992). A large body of data now supports this broad pattern of specialization: it is clear that the dorsal (occipito-parietal) stream is necessary for normal visuomotor guidance, whilst object and scene recognition depend critically upon the ventral (occipito-temporal) stream (Milner & Goodale, 2007). What is less clear at present is the nature and extent of interactions between these two visual subsystems in normal behaviour.

Some specific contributions of perceptual recognition to action guidance are uncontroversial within the perception/action model. For instance, semantic processing dependent upon object recognition must necessarily influence action selection (Carey, Harvey, & Milner, 1996). Perceptual recognition may also be needed for the anticipatory programming of fingertip forces when picking objects up, since weight prediction requires access to stored knowledge of objects and materials (McIntosh, 2000). On the other hand, the visuospatial guidance of action is held to be encapsulated from object recognition, depending exclusively upon direct, bottom-up information (Milner & Goodale, 2007). The most-cited evidence for this idea is the remarkable preservation of visuomotor abilities in patient DF, a woman with profoundly impaired object recognition following bilateral lesions to the lateral occipital area, a crucial

node in the ventral pathway (Milner & Goodale, 2007; Milner et al., 1991). More recent investigations, however, have suggested that DFs success in visuomotor tasks may owe partly to functional reorganisation, and not only to dorsal stream integrity. Although she reaches to visual targets as accurately as do healthy individuals, DF is much more reliant on binocular vision (Mon-Williams, Tresilian, McIntosh, & Milner, 2001; Wann, Mon-Williams, McIntosh, Smyth, & Milner, 2001). This may imply that the healthy visuomotor system uses additional distance cues, unavailable to DF, presumably mediated by the ventral, perceptual pathway.

A second line of evidence used to support a strong perception–action division is the apparent insensitivity of visually guided actions to illusions of size (e.g. Aglioti, DeSouza, & Goodale, 1995), position (e.g. Bridgeman, Peery, & Anand, 1997) or orientation (e.g. Dyde & Milner, 2002) created by manipulating visual context. On the assumption that such illusions arise within the ventral stream, their failure to affect action can be taken to demonstrate the independence of action from perception, even in healthy individuals (Milner & Goodale, 2007). However, several studies have reported positive effects of contextual illusions on action, casting doubt on the basic dissociation (see Bruno, 2001; Franz, 2001; Smeets & Brenner, 2006, for reviews). Unfortunately, even these latter results do not afford simple interpretation, since the neural bases of the pictorial illusions employed are still uncertain. An effect of contextual illusions on action might imply a functional contribution of the perceptual pathway to action guidance; alternatively, the illusory effects might arise within the visuomotor system separately, or within the early visual system, prior to the split between processing streams (Milner & Dyde, 2003). Notwithstanding the vigour with which ‘illusions-in-action’ phenomena are debated,

* Corresponding author. Tel.: +44 131 6503444; fax: +44 131 6503461.
 E-mail address: r.d.mcintosh@ed.ac.uk (R.D. McIntosh).

contextual illusions seem ill-suited to address the functional relationship between vision-for-perception and vision-for-action.

A more powerful, direct test of whether perceptual recognition contributes to action guidance is afforded by the pictorial depth cue of familiar size. Since the visual angle subtended by an object is inversely proportional to viewing distance (Emmert's law), an object's absolute distance can be recovered from its retinal image size, provided that the viewer recognises the object and can access a stored representation of its true (or typical) size. This cue thus depends, by definition, on object recognition. Familiar size has long been known to influence distance estimation, especially when other cues to distance are sparse (Holway & Boring, 1941). However, the conditions under which it influences action guidance in humans are presently unclear.

Marotta and Goodale (2001) asked participants to grasp featureless spheres under monocular or binocular viewing conditions. In their familiar-size condition, the same sphere was presented on 17 consecutive trials, and the effects of this repeated exposure were assessed over the next 17 trials, in which the standard sphere was substituted for a larger or smaller 'probe' on six occasions. The main outcome was that the speed profile of the reaches showed more secondary peaks, indicative of programming errors and corrective sub-movements, for probe trials in the monocular condition. The authors concluded that familiar size does not normally inform reach-to-grasp movements, but that observers may use it when binocular cues are denied. However, we suggest that featureless spheres may not be good stimuli for encouraging the use of familiar size, and that the high rate of probe trials (47% of trials in the perturbation phase) could have attenuated any familiarity effects rapidly. We now re-assess whether familiar size affects action, using more meaningful objects and a lower rate of perturbation. We show that familiar size influences both the distance and size estimates used for programming reach-to-grasp movements, even when veridical bottom-up distance information is available binocularly.

2. Methods

2.1. Participants

Twenty-four participants were tested. All were right-handed, as assessed by the Edinburgh Handedness Inventory (EHI; Oldfield, 1971), had normal stereoscopic vision, as assessed by the screening plates of the TNO stereotest, and had normal or corrected-to-normal visual acuity. Twelve participants were assigned to a binocular viewing condition, and 12 to a (right-eye) monocular viewing condition. Ocular dominance was assessed by the Porta Test, and right-eye dominant participants only were assigned to the monocular condition; six participants in the binocular condition showed left- or mixed-ocular dominance. Participants in the monocular condition (9♀, 3♂) had a median age of 24.5 years (range: 21–52) and a median EHI laterality quotient of +91 (range: 54–100). Participants in the binocular condition (10♀, 2♂) had a median age of 21.5 years (range: 20–33) and a median laterality quotient of +100 (range: 82–100). Mann–Whitney *U*-tests found no reliable differences between the groups in age ($U = 52.5$, $p = 0.27$) or laterality quotient ($U = 52.5$, $p = 0.09$). This experiment was conducted in accordance with the 1964 Declaration of Helsinki and with the approval of the Ethics Committee of the School of Philosophy, Psychology and Language Sciences at the University of Edinburgh. All participants gave their informed verbal consent prior to testing.

2.2. Procedure

Participants sat at a table under diffuse illumination, with their head immobilised in a chinrest, wearing LCD shutter glasses (PLATO, Translucent Technologies) and holding a start-point on the table between their right index finger and thumb. Each trial began with the LCD glasses turning from opaque to clear, revealing a matchbox presented at one of five distances directly in front of the eye(s) (270, 315, 360, 405, 450 mm). The box was supported from behind by a thin wooden dowel, to which it was attached by a small magnet, and which was concealed from the participant by the box itself. The box was viewed against a white canvas backdrop, 725 mm from the eyes, which otherwise filled the entire field of view afforded by the glasses. Half-a-second after viewing onset, a tone cued the participant to reach out and grasp the box, top-to-bottom, between finger and thumb. The LCD glasses remained clear for 2 s after the tone, so that the movement was always completed with visual feedback available. Movements were recorded by sampling, at 200 Hz, the 3D positions



Fig. 1. The four matchboxes. The upper row shows the standard *Swan Vestas* (79 mm × 45 mm × 13 mm) and *Scottish Bluebell* boxes (53 mm × 36 mm × 14 mm) presented in the baseline trials. The lower row shows the 0.8-scale replica *Swan* box and 1.25-scale replica *Bluebell* box presented in the perturbation trials. Boxes were presented in the orientation shown and grasped, top-to-bottom, between finger and thumb.

of infrared emitting markers attached to the wrist, and to the distal phalynxes of the thumb and index finger (Optotrak, Northern Digital Incorporated). Movement data were analysed offline; raw data were filtered at 20 Hz using a second-order Butterworth dual-pass filter and analysed using custom software. Movement onset and offset were defined from the tangential speed of the wrist marker, using a threshold of 50 mm/s.

Standard matchboxes of the (locally-common) *Swan Vestas* and *Scottish Bluebell* brands were used for the first 42 (and six preceding practice) trials (see Fig. 1). Baseline trials occurred when the *Bluebell* box was presented at 360 mm ('near' baseline trials) or the *Swan* box at 450 mm ('far' baseline trials). These baseline trials were repeated nine times each, interspersed randomly with three filler trials for each box at each of the four other distances. Perturbations of the standard box sizes were then applied on two trials. On trial 43, a 0.8-scale replica of the *Swan* box was presented at 360 mm; this presented the same box-height for grasping (36 mm) at the same distance as did the standard *Bluebell* box in the near baseline trials, but projected a retinal image pictorially consistent with the standard *Swan* box seen at 450 mm (i.e. far baseline trials). On trial 44, a 1.25-scale replica of the *Bluebell* box was presented at 450 mm; this gave the same box-height for grasping (45 mm) at the same distance as did the standard *Swan* box in the far baseline trials, but projected a retinal image consistent with the standard *Bluebell* box seen at 360 mm (i.e. near baseline trials). If familiar size of the standard boxes contributes to the distance estimates used for reaching, then participants should over-reach, relative to baseline, for the near perturbation and under-reach for the far. However, if bottom-up information only is used, then grasping should be impervious to the perturbations of box size.

3. Results

Average movements in the baseline and corresponding perturbation trials showed a consistent influence of familiar size (see Fig. 2). Overall, participants over-reached for the small *Swan* box at the near distance, and under-reached for the large *Bluebell* box at the far distance. These errors were observed informally during the experiment, in that participants (especially in the monocular condition) would occasionally collide with the small *Swan* box and/or grasp ineffectually at the air in front of the large *Bluebell* box. Inferential analyses focused on the depth displacement of the index finger at the end of the movement in the baseline and perturbation trials (equivalent results were obtained for the thumb and wrist). A mixed-model ANOVA was conducted with the between-subjects factor of viewing condition (monocular, binocular). The within-subjects manipulations were coded as orthogonal factors of physical distance (near, far) and familiar-size-specified distance (near, far), which were thus congruent in the baseline trials, but incongruent in the perturbation trials. There were reliable main effects for both physical distance ($F_{(1,22)} = 325.04$, $p < 0.0005$) and familiar-size-specified distance ($F_{(1,22)} = 21.78$, $p < 0.0005$), and

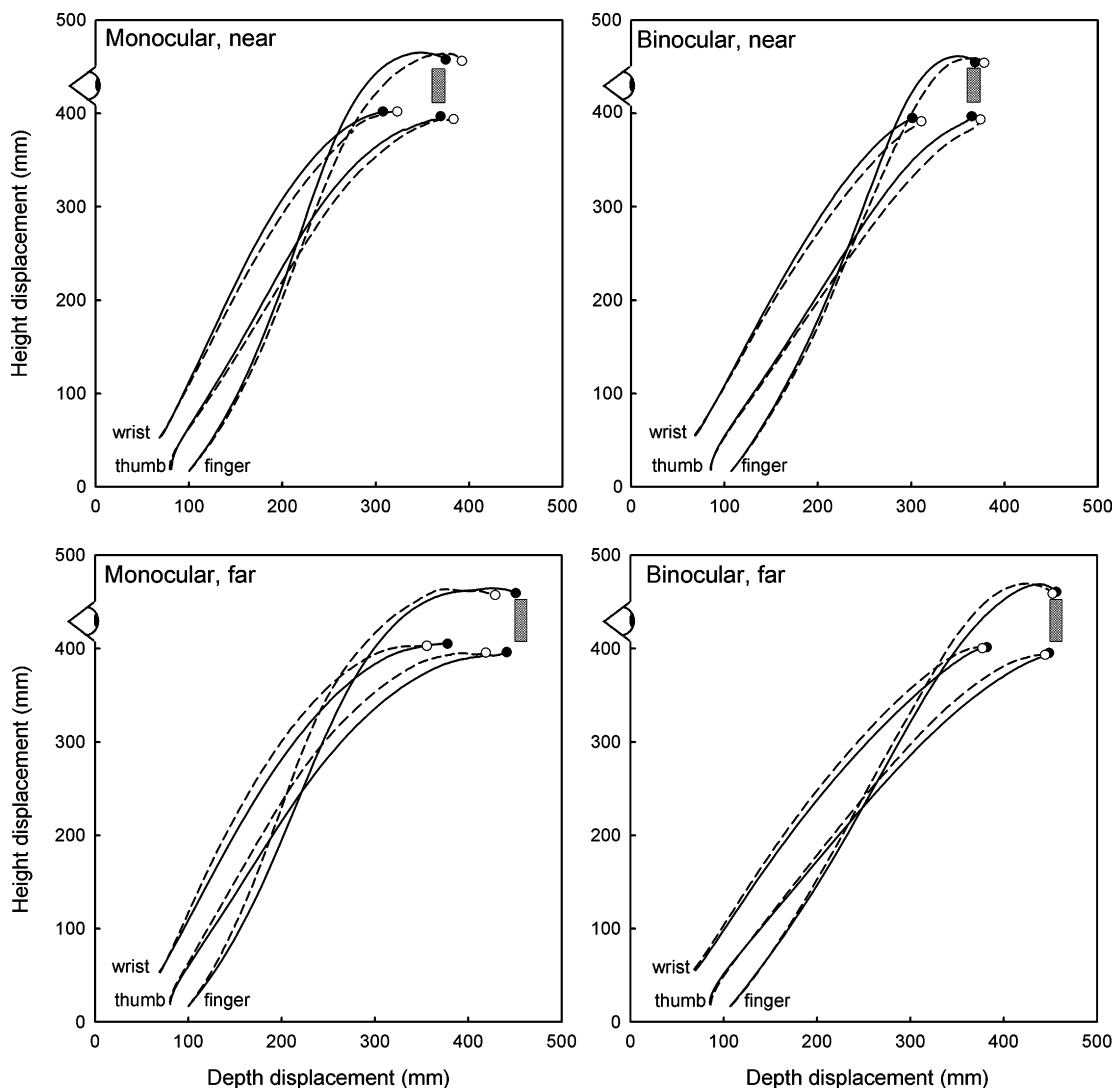


Fig. 2. Average movements and endpoints in the sagittal plane for the baseline (solid lines, filled circles) and perturbation (dashed lines, unfilled circles) trials. Eye position and target boxes are indicated. Movement onset and offset were defined by comparing the tangential speed of the wrist marker against a threshold of 50 mm s^{-1} . Prior to spatial averaging, each movement was normalized to 100 frames.

an interaction between familiar-size-specified distance and viewing condition ($F_{(1,22)} = 5.16$, $p < 0.0005$), such that the effect of familiar size was amplified with monocular viewing. Crucially, repeated measures ANOVAs performed separately for each viewing condition confirmed a reliable influence of familiar-size-specified distance in the binocular ($F_{(1,11)} = 30.28$, $p < 0.0005$) as well as the monocular condition ($F_{(1,11)} = 12.63$, $p < 0.01$). Indeed, although the biasing effect of familiar size was metrically smaller in the binocular condition (6.82 mm vs. 19.75 mm), the effect size was actually greater (partial $\eta^2 = 0.73$ vs. 0.54), due to the higher precision of movement with binocular viewing.

To assess the influence of familiar size upon grasp preparation, mean maximum grip apertures (MGA: the maximum separation between finger and thumb measured during the movement) in the baseline and perturbation trials were entered into a mixed-model ANOVA by physical size (small, large), familiar size (small, large) and viewing condition (monocular, binocular). Reliable main effects were found for both physical size [$F_{(1,22)} = 9.78$, $p < 0.01$] and familiar size [$F_{(1,22)} = 11.33$, $p < 0.005$], with familiar size having a marginally stronger influence both statistically (partial $\eta^2 = 0.34$ vs. 0.31) and metrically (2.58 mm vs. 1.91 mm). Although the biasing effect of

familiar size was metrically greater in the monocular than in the binocular condition (3.50 mm vs. 1.66 mm), the interaction with viewing condition was not reliable [$F_{(1,22)} = 1.44$, $p = 0.24$]. Overall, participants' grasps were over-scaled for the small *Swan* box, and under-scaled for the large *Bluebell* box, with familiar size having at least as much influence as physical size.

Finally, we studied the incidence of overt corrective sub-movements, as indexed by the number of peaks in the speed profile of the wrist marker, in the baseline and perturbation trials. For this analysis, the within-subjects factors were coded in terms of physical distance (near, far) and perturbation condition (baseline, perturbation). The main effect of viewing condition ($F_{(1,22)} = 12.23$, $p < 0.005$) reflected reliably more speed peaks in the monocular group in all conditions (monocular mean = 1.40, S.D. 0.38; binocular mean = 1.01, S.D. 0.02). The main effect of perturbation condition ($F_{(1,22)} = 5.61$, $p < 0.05$) was modified critically by an interaction with viewing condition ($F_{(1,22)} = 6.84$, $p < 0.05$), such that size perturbations increased the number of overt corrective sub-movements in the monocular condition (mean increase 0.37, S.D. 0.52), but not in the binocular condition (mean increase -0.02 , S.D. 0.04), consistent with the results of [Marotta and Goodale \(2001\)](#).

4. Discussion

These results indicate that familiar size contributes to the programming of reach-to-grasp movements, modifying both reach amplitude and hand pre-shaping. It would seem that the vision-for-perception system, whilst specialised for object and scene recognition, contributes fundamentally to the spatial programming of action by mediating access to stored object knowledge. This conclusion is bolstered by the fact that familiar size has a robust influence even when binocular vision provides rich, bottom-up distance information to the action system.

Our conclusion is thus rather different from that of a prior study on this issue, which reported no effect of familiar size on grasping movements under binocular conditions (Marotta & Goodale, 2001). We believe that the more visible influence of familiar size in the present study can be attributed to our use of *meaningful* graspable objects, and rare perturbations of familiar size. Notably, as well as measuring clear effects of familiar size perturbation on reach distance and grasp preparation, we also replicated Marotta and Goodale's key finding of an increased incidence of corrective sub-movements during perturbation trials in the monocular condition. Rather than demonstrating that familiar size perturbations do not induce programming errors under binocular conditions, this may simply reflect the fact that binocular vision allows on-line corrections to be implemented more smoothly and efficiently (e.g. Servos & Goodale, 1994).

Our data show that familiar size influences action, but we also observed that this influence was amplified with the removal of binocular cues. This result is in accord with the predictions of the modified weak fusion model of depth cue combination (Landy, Maloney, Johnston, & Young, 1995). In such a weighted averaging scheme, the weightings assigned to the available cues must sum to one; therefore, if binocular cues are removed, the weightings accorded to the remaining cues (including familiar size) should increase proportionately. An alternative, non-mutually exclusive possibility is that the smaller perturbation effects in the binocular condition reflect the more efficient use of visual feedback to correct for induced programming errors. Further experiments manipulating feedback availability would be required to test this idea.

One question raised by our results is the precise nature of the contribution that familiar size makes to spatial representations for action. This cue evidently modifies the distance estimate used in programming reach amplitude. We also observed that the pre-shaping of the hand was biased towards the familiar size of the target box. It is unclear, on present evidence, whether this grasp effect arises because the retinal size of the target is interpreted using a distance estimate that is modified by familiar size, or whether the familiar size of an object contributes in a more direct manner to the programming of the grasp formation. Similarly, our methods do not allow us to determine whether the memory representations that mediate familiar size in this experiment are enduring semantic representations of the target boxes, unstable temporary traces formed by associative learning during the baseline trials, or reflect some combination of long- and short-term processes. In any case, unless the perceptual and memory capabilities of the dorsal stream

are vastly richer than previously thought, we would infer a necessary role for the ventral stream in matching current retinal input against these stored object representations.

Overall, the present data question the notion that the visuomotor system accesses top-down knowledge only as a last resort, when binocular vision is denied. Instead, distance information derived via the ventral pathway may be available continuously to action systems, with the weightings adjusted dynamically depending upon the availability and reliability of other cues (Landy et al., 1995). Familiar size is an ideal test case with which to settle this debate, since its influence cannot be explained without a top-down influence from stored object knowledge. When we consider the range of objects that we interact with in a typical day, it is hard to generate many instances in which the object is familiar neither individually nor as a member of a known class. This consideration implies that the involvement of perceptual recognition in action programming is the rule and not the exception.

References

- Aglioti, S., DeSouza, J. F. X., & Goodale, M. A. (1995). Size contrast illusions deceive the eye but not the hand. *Current Biology*, 5, 679–685.
- Bridgeman, B., Peery, S., & Anand, S. (1997). Interaction of cognitive and sensorimotor maps of space. *Perception & Psychophysics*, 59, 456–469.
- Bruno, N. (2001). When does action resist visual illusions? *Trends in Cognitive Sciences*, 5, 379–382.
- Carey, D. P., Harvey, M., & Milner, A. D. (1996). Visuomotor sensitivity for shape and orientation in a patient with visual form agnosia. *Neuropsychologia*, 34, 329–338.
- Dyde, R. T., & Milner, A. D. (2002). Two illusions of perceived orientation: one fools all of the people some of the time; the other fools all of the people all of the time. *Experimental Brain Research*, 144, 518–527.
- Franz, V. H. (2001). Action does not resist visual illusions. *Trends in Cognitive Sciences*, 5, 457–459.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, 15, 20–25.
- Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology*, 54, 21–37.
- Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, 35, 389–412.
- Marotta, J. J., & Goodale, M. A. (2001). The role of familiar size in the control of grasping. *Journal of Cognitive Neuroscience*, 13, 8–17.
- McIntosh, R. D. (2000). Seeing size and weight. *Trends in Cognitive Sciences*, 4, 442–444.
- Milner, A. D., & Dyde, R. T. (2003). Why do some perceptual illusions affect visually guided action, when others don't? *Trends in Cognitive Sciences*, 7, 10–11.
- Milner, A. D., & Goodale, M. A. (2007). *The visual brain in action* (2nd ed.). Oxford, UK: Oxford University Press.
- Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., et al. (1991). Perception and action in 'visual form agnosia'. *Brain*, 114, 405–428.
- Mon-Williams, M., Tresilian, J. R., McIntosh, R. D., & Milner, A. D. (2001). Monocular and binocular distance cues: insights from visual form agnosia I (of III). *Experimental Brain Research*, 139, 127–136.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Servos, P., & Goodale, M. A. (1994). Binocular vision and the on-line control of human prehension. *Experimental Brain Research*, 54, 121–132.
- Smeets, J. B. J., & Brenner, E. (2006). 10 years of illusions. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1501–1504.
- Wann, J. P., Mon-Williams, M., McIntosh, R. D., Smyth, M., & Milner, A. D. (2001). The role of size and binocular information in guiding reaching: insights from virtual reality and visual form agnosia III (of III). *Experimental Brain Research*, 139, 143–150.