

Visual space perception and action: Introductory remarks

Jochen Müsseler

*Max Planck Institute for Human Cognitive and Brain Sciences, Munich,
Germany*

A. H. C. van der Heijden

Leiden University, The Netherlands

Dirk Kerzel

Giessen University, Germany

Vision evolved from the vital necessity to act in a dynamic environment. Following this view it is clear that perceptual processes and action planning are much more interlocked than is evident at first sight. This is especially evident in visual space perception; actions are performed in space and are guided and controlled by objects in spatial positions. Here we shortly introduce the three research camps dealing with the relationship between space perception and action: the ecological camp, the two-visual-systems camp, and the constructivist camp. We show that these camps emphasize and open different theoretical and empirical perspectives, but that they can be seen to complement each other. We end with an overview of the papers in this special issue.

Vision had no end in itself. The visual system, just as all other “perceptual systems have evolved in all species of animals solely as a means of guiding and controlling action” (Allport, 1987, p. 395). Given this point of view it is clear that empirical and theoretical work concerned with actions that humans and animals have in common—with moving and jumping, with grasping, picking and catching, with approaching and avoiding—has to care about perception, about action, and about their interaction. And it is clear that in such empirical and theoretical work the elaboration of the perception of space and of position is of vital importance; all such actions are performed in space and are guided and controlled by objects on positions.

Please address correspondence to: Jochen Müsseler, Max-Planck-Institut für Kognitions- und Neurowissenschaften, Amalienstr. 33, D-80799 München, Germany. Email: muesseler@psy.mpg.de

CURRENT APPROACHES ON VISUAL SPACE PERCEPTION AND ACTION

Nowadays, the vital importance of space perception in the relationship between perception and action is recognized and discussed in three largely independent research camps: (1) the ecological camp, (2) the two-visual-systems camp, and (3) the constructivist camp. The theoretical points of view emphasized in these camps are not mutually exclusive. In fact, by emphasizing (1) light, (2) brain, and (3) behaviour, they neatly complement each other (for a more elaborated integrating account see Norman, 2002).

The ecological camp emphasizes light. According to the Gibsonian ecological approach (e.g. Gibson, 1979; cf. also Reed, 1996), perception and action are linked by affordances, that is, by the action possibilities that the world offers and that are specified in the structure of the light surrounding the perceiver. Affordances are aspects of—or possibilities in—the environment with reference to the animal's body and its action capabilities, like, for instance, the “climbability” of a rock face or the “sittability” of a stump.

Ecologists stress the active perceiver exploring his/her environment. Because the vital structure of the light is not simply given but has to be extracted, eye, head, and body movements are seen as parts of the perceptual process. Thus, perception means to perceive events, which change over time and space through body and object movements. Space perception comprises the many surfaces that make up the environment. The perceptual performance of an observer consists of the pickup of the (invariant) information inherent in the structured light coming from this environment in a direct manner. Gibson refrained to refer to the processes underlying perception, thus this camp is basically silent about the brain. The only concession was the resonance principle according to which the perceptual system resonates with or is attuned to the invariant structures. Consequently, ecologists analyse the dynamics of perception and action (e.g., Thelen & Smith, 1994) and decline the reductionist experimental paradigms. Experiments with stimulus presentations of a few milliseconds and simple key presses as a behavioural measure are simply seen as inadequate for the analysis of the perception–action interplay. Of course, this was and is always a matter of dispute particularly with the representatives of the constructivist account (see below; for discussions see, e.g., Gordon, 1997, chap. 7; Nakayama, 1994).

The two-visual-systems camp accepts these views on light but emphasizes brain. They provide neuropsychological, neuroanatomical, and behavioural evidence for two channels in the visual system, one channel for perception/cognition, the so-called vision-for-perception pathway, and one channel for action, the so-called vision-for-action pathway. In modern Experimental Psychology, evidence for two visual systems can be found in some studies of the 1960s and 1970s (e.g., Bridgeman, Hendry, & Stark, 1975; Fehrer & Raab, 1962). Somewhat later, Ungerleider and Mishkin (1982) provided evidence for a

ventral pathway leading from the occipital cortex to the inferior temporal lobe, assumed to deal with object identification, and a dorsal pathway leading from the occipital cortex to the posterior parietal cortex, assumed to deal with object location. Later still Goodale and Milner (1992; see also Milner & Goodale, 1995) attributed the dorsal stream the function of the visual control and guidance of motor behaviour. In the recent past, the two-visual system account inspired numerous studies.

While ecologists stress the unity of perception and action, the representatives of the two-visual system camp emphasize diversity and dissociation. Given the assumed modularity of spatial information processing in both functional and structural terms, their studies aimed at and found evidence for dissociations between perception and action also in behavioural studies (e.g., Aglioti, DeSouza, & Goodale, 1995; Haffenden & Goodale, 1998, 2000; Milner & Dyde, 2003). This research strongly suggests that “what one sees” is basically different from “what one needs to act”.

The constructivist camp accepts these views on light and views on brain but emphasizes the importance of the interaction between perception/cognition and action. Constructivists are not convinced that perception and action are largely separated and unconnected cognitive domains and emphasize that, despite the rare examples of diversity and dissociation, the normal case is unity and association. In their view, while the anatomical and functional architecture and the transformational computations involved might be highly complex, there can be no doubt that the system uses spatially coordinated maps for perception and action.

Besides the role of perception for action, this camp emphasizes the importance of action for perception. There is indeed increasing evidence that the functional unity of perception and action works not only from perception to action but from action to perception as well (e.g., Hommel, Müsseler, Ascherleben, & Prinz, 2001; Müsseler & Wühr, 2002). As animals act in their environment, perceptual information cannot be interpreted in an unambiguous way without reference to action-related information. This, in turn, requires that these two pieces of information interact with each other. With regard to visual space perception, the interesting question is: What is the influence of action on the experienced visual space with objects on positions?

In the constructivist camp the central question is the question how the visual system figures out what environmental situation gives rise to the optical image registered on the retina (cf. Rock, 1983). Different kinds of cues, and especially cues resulting from actions in the world, assist in this construction process. By analysing the motion parallax phenomenon, von Helmholtz (1866) could already claim that one’s own movements deliver important depth cues. In contrast to the ecologists, who focus almost exclusively on the external environment as specified in light, constructivists take also into account the internal cognitive mechanisms. In this sense they are representatives of the information processing account. In contrast to the representatives of the two-visual system account, who

might also feel constrained with the information processing account, the constructivists are more interested in the aspects of unity and association between perception and action.

THE CONTRIBUTIONS TO THE SPECIAL ISSUE

This special issue on Visual Space Perception and Action brings together 12 contributions written by various experts in this field, ranging from experimental psychologists and neurophysiologists to computational modellers and philosophers. Each contribution introduces new concepts and ideas that explain how visual space is being established and represented.

The first two papers present theoretical discussions about how position and information about one's actions may be represented in the brain. Wolff contrasts two hypotheses about codes for the position of objects. It is often assumed that the retinotopic organization of cortical and subcortical structures codes the position of objects in space. Wolff however shows that this assumption cannot be maintained in the face of a constantly moving observer. Instead, spatial coding is learned from the sensorimotor contingencies. That is, the relation between observer motion and its sensory consequences may establish a code for position. Whereas Wolff is concerned with an abstract theoretical framework of spatial coding, Bremmer, Schlack, Graf, and Duhamel elaborate on how self-motion may be represented in the parietal cortex. One important visual cue to self-motion is optic flow. Forward motion produces an expansion of image elements, and backward motion a contraction. The precise direction of motion may be derived from a singularity, the focus of expansion. Bremmer et al. show that neurons in the ventral intraparietal area (VIP) of the macaque cortex are sensitive to variations of the focus of expansion and may therefore code the direction of egomotion. These neurons not only respond to visually simulated self-motion, but also to real physical displacement. This shows that VIP may be a multimodal area in which visual and vestibular as well as auditory and somatosensory information is integrated.

The next two contributions deal with information processing around the time of a saccade. Both contributions focus on the detection of position changes that are presented when observers move their eyes from a source object to a target object. In both studies, additional elements are displayed. Deubel asks how the detection of target displacements during the saccade is affected by surrounding objects, referred to as landmarks. He reports that there is a bias to localize the target toward the irrelevant landmarks when these landmarks are close to the target and horizontally aligned with it. Germeys, de Graef, Panis, van Eccelpoel, and Verfaillie pursue the opposite goal of clarifying how memory of surrounding objects, referred to as bystanders, is organized. They find that bystander location is better remembered if the complete scene is presented during recall suggesting that the location of a single bystander is encoded with

respect to the remaining bystanders. As a challenge for current theorizing, Germeys et al. show that the saccade source may be more important for encoding of bystander location than the saccade target. Taken together, the two studies converge on the general conclusion that transsaccadic memory relies strongly on relational information.

The following two papers contribute to an ongoing discussion about whether the perception of space differs from the representation of space used for motor action. Müsseler and van der Heijden examined the hypothesis that two sources may be used to calculate position, a sensory map, and a motor map. The sensory map provides vision while the motor map contains the information for saccadic eye movements. The model predicts that errors in relative location judgements will be observed when the motor map has to provide the information for the visual judgements. The authors provide evidence for this model by showing that the perceived position of differently sized targets follows the same pattern as saccadic eye movements to these targets: Eye movements to a small target undershoot less than eye movements to a spatially extended target. A similar trend is found when observers make perceptual judgements that require relative position judgements: The centre of a small object appears further from the fixation point than the centre of a spatially extended object. Thus, this paper shows association and not dissociation between perception and action.

A similar conclusion is reached in the contribution of Smeets and Brenner. The authors investigate why observers who are asked to connect two points with a straight line fail to do so and draw a curved line between the two positions instead. This may be so because of a spatial distortion, or because the direction of the motion is initially wrong and requires continuous adjustment. The perception of straightness was influenced by using the Hering illusion. In the Hering illusion, a straight line appears curved because a pattern of radiating lines is superimposed. Observers had to judge the straightness of a dot moving across the Hering illusion, as well as draw a straight line across the illusion. Smeets and Brenner found that the curvature that the background induced in the hand's movement path was correlated with the curvature that the background induced in a moving dot's path. Thus again, perception and action are more associated than dissociated.

The next block of four papers is concerned with the perceived position of moving objects. In earlier work, Nijhawan proposed that processing latencies are compensated by motion extrapolation in the visual system such that a flashed stationary object is seen to lag a moving object. A conflicting explanation for the flash-lag effect is that latencies of moving objects are reduced. Both accounts would explain why responses to moving objects are typically accurate. However, Nijhawan, Watanabe, Khurana, and Shimojo show that reaction times to moving stimuli are not reduced compared to stationary objects, and temporal order judgements do not indicate that moving objects are perceived earlier than stationary ones. This contradicts the claim of latency reduction with moving stimuli and favours motion extrapolation.

Stork and Müsseler as well as Thornton and Hayes investigate factors that affect localization of the endpoint of a motion trajectory. Stork and Müsseler show that the endpoint of a moving stimulus is mislocalized in the direction of motion when it is pursued with the eyes. In contrast, judgements of the final target position are accurate when eye fixation on a stationary object is maintained. When observers had control over the target's vanishing point because target disappearance was coupled to a key press, both eye movements and position judgements beyond the vanishing point were reduced, suggesting that intentions affect eye movements and position judgements in a similar manner.

Effects of eye movements and further cognitive factors on endpoint localization were modelled by Erlhagen and Jancke. Their model consists of interacting excitatory and inhibitory cell populations. The intrinsic network dynamics explain mislocalization of the final position of a moving target by assuming that the population response to the moving stimulus continues to travel in the direction of stimulus motion even after stimulus offset. However, in the absence of stimulus input, the dynamic extrapolation of trajectory information decays. The strength of extrapolation depends on thresholds for recurrent interactions. The lower the threshold, the further the forward shift of the final target position. It is assumed that cognitive factors and eye movements may adjust these thresholds.

Thornton and Hayes extend previous work on the mislocalization of the final target position to complex scenes. The authors used movies simulating self-motion through an artificial landscape or movies of realistic scenes, such as passengers boarding a train. They found that the position in the final image of these movies was shifted in the direction of the self-motion. This effect was independent of the nature of the probe stimulus. Regardless of whether observers compared the final image of the movie to a static or a dynamic test image, a forward shift was observed. This paper broadens our understanding of memory for dynamic events by showing that mislocalization of the final target position is not confined to the highly impoverished stimuli used in previous studies (typically disks or rectangles), but may occur in more realistic scenarios as well.

In the final section of this issue, two studies with higher demands in motor control discuss how intended or executed manual actions influence spatial perception. Johnson-Frey, McCarty, and Keen asked their participants to grasp an object. Consistent with Fitts' law, they found that movement time increased with decreasing object size. Additionally, movement times were shorter when observers intended to transport it to a new location compared to when they only had to lift the object. This effect was independent of the difficulty of the task following the initial grasping movement. These results indicate that both the immediate and the future goal of a movement determine movement speed.

Finally, Humphreys, Riddoch, Forti, and Ackroyd review recent literature on two interesting symptoms: Neglect and Balint's syndrome. They show that tool use may alleviate neglect for the contralesional side because of visual and

visuomotor cueing. If patients explore space with a tool, objects close to the tool may be detected even if they fall within the neglected area of space. Also, neglect is reduced if two objects, one falling into the neglected part of space, are placed relative to each other such that they have the correct place for action. For instance, a hammer may be placed above the nail. Similarly, correct action relations may allow for improved binding of object properties in a patient with Balint's syndrome.

Overall, the 12 papers included in this Special Issue present a number of exciting findings and raise a number of interesting questions for future research. In addition the papers make clear that space perception and action is a central component of human perception and performance.

REFERENCES

- Aglioti, S., DeSouza, J. F. X., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology*, *5*(6), 679–685.
- Allport, D. A. (1987). Selection for action: Some behavioral and neurophysiological consideration of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395–419). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*(6), 719–722.
- Fehrer, E., & Raab, D. (1962). Reaction time to stimuli masked by metacontrast. *Journal of Experimental Psychology*, *63*, 143–147.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25.
- Gordon, I. E. (1997). *Theories of visual perception*. New York: John Wiley & Sons.
- Haffenden, A. M., & Goodale, M. A. (1998). The effect of pictorial illusion on prehension and perception. *Journal of Cognitive Neuroscience*, *10*(1), 122–136.
- Haffenden, A. M., & Goodale, M. A. (2000). Independent effects of pictorial displays on perception and action. *Vision Research*, *40*(10–12), 1597–1607.
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(5), 869–937.
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford, UK: Oxford University Press.
- Milner, D., & Dyde, R. (2003). Why do some perceptual illusions affect visually guided action, when others don't? *Trends in Cognitive Sciences*, *7*(1), 10–11.
- Müsseler, J., & Wühr, P. (2002). Response-evoked interference in visual encoding. In W. Prinz & B. Hommel (Eds.), *Attention and performance XIX: Common mechanisms in perception and action* (pp. 520–537). Oxford, UK: Oxford University Press.
- Nakayama, K. (1994). James J. Gibson: An appreciation. *Psychological Review*, *101*(2), 329–335.
- Norman, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. *Behavioral and Brain Sciences*, *25*, 73–144.
- Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. New York: Oxford University Press.
- Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.

- Thelen, E., & Smith, L. B. (1994). *A dynamics systems approach to the development of perception and action*. Cambridge, MA: MIT Press.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.
- von Helmholtz, H. (1866). *Handbuch der physiologischen Optik* [Handbook of physiological optics]. Hamburg, Germany: Voss.