Husain, 2005). This is the case when reaching for extra-foveal targets but not for foveated targets (see Jackson et al., 2009). Furthermore, the "dorsal stream" should not be viewed as "visual" but instead as a brain region in which different sensory signals are dynamically integrated to produce multimodal task-dependent spatial representations. In the case of the POJ, it appears to play a role in integrating spatial signals within a "global tuning field" network that relate to the direction of gaze and the direction of reaching movements, and may compute an error or displacement vector based on the angular difference between gaze direction and reach direction that may be particularly important when executing movements away from the direction of gaze (Jackson et al., 2009).

* * *

Vision for action is not veridical

Jeroen B. J. Smeets and Eli Brenner
Research Institute MOVE, Faculty of Human Movement Sciences, VU University Amsterdam, van der Boechorststraat 9, NL-1081 BT Amsterdam, The Netherlands
E-mail: J.Smeets@fbw.vu.nl

DOI: 10.1080/17588920903490939

Abstract: We agree with Schenk and McIntosh that the human brain can better be described in terms of task-specific functional networks than in terms of a division between (dorsal) egocentric vision for action and (ventral) viewpoint-independent vision for perception. However, by concentrating on the lack of experimental support for the latter division, the authors neglect an important reason for postulating that there is a separate vision-for-action system: needing veridical metric information to guide one's actions. We argue that considering this reason would support the authors' conclusion because the visual information that guides our actions does not have to be veridical.

The notion of task-to-task functional networks that Schenk and McIntosh propose is a very likely neural implementation of the behavioral view of task-dependent use of spatial attributes (Smeets, Brenner, de Grave, & Cuijpers, 2002). In this view, the reliability of information determines which information will be used for the task at hand. Aiming for a high reliability can also explain why we use positions and not size for shaping our hand during grasping (Smeets & Brenner, 2008), and why one shifts from using egocentric toward allocentric information after a delay if the target is removed from view.

The idea that we rely on the most *reliable* information seems to imply that we need *veridical* metric information to control our actions (Aglioti, DeSouza, & Goodale, 1995). However, it is clear that not all aspects of vision that are used to control our actions are veridical: Illusions of size, orientation, and speed have all been shown to influence certain aspects of movements (Smeets, et al., 2002). One might even argue that the only aspect of vision for action that needs to be veridical is the information about the target's location.

A target's position can be determined based on extraretinal information about eye-orientation, but often also from other sources, such as pictorial cues. We know that pictorial depth illusions can affect the manual tracking of a moving target with one's invisible hand to the same extent as perception of depth (López-Moliner, Smeets, & Brenner, 2003). Schenk and McIntosh correctly mention that the reliability of information for online control depends on the latency at which it is available: The shorter the latency, the more reliable it is. We have recently shown that pictorial depth cues can be used at a latency that is 40 ms shorter than that for binocular depth cues (van Mierlo, Louw, Smeets, & Brenner, 2009). This means that a cue that need not provide veridical information can be the most important for the online control of actions.

The whole idea that veridical information is available is probably wrong. Even extraretinal information about the position of a fixated target is not veridical: Subjects show biases that remain stable across days when moving their invisible hand to isolated visual targets (Smeets, van den Dobbelsteen, de Grave, van Beers, & Brenner, 2006). A careful analysis of the way in which information is used shows why this bias is not a problem: We combine all information optimally, both for the hand and for the target (Smeets et al., 2006). For locating our hand, we combine several visual cues with proprioception; for locating the target we combine visual cues with *extended proprioception*. This term refers to allocentric (visual) information about the target's position relative to the hand, converted into an egocentric judgment by combining it with the proprioceptive location of the hand. In this way, non-veridical visual or proprioceptive information affects the perceived location of the hand and that of the target in the same way, which makes the scheme robust for errors in either modality.

Our conclusion is that there is neither experimental evidence nor a theoretical need for veridical vision-for-action. This makes the absence of a separate vision-for-action stream (as proposed by Schenk & McIntosh) easy to accept.

* * *