



Comment

The role of action semantics in action control  
Comment on “Action semantics: A unifying conceptual framework  
for the selective use of multimodal and modality-specific object  
knowledge” by Michiel van Elk, Hein van Schie and  
Harold Bekkering

Bernhard Hommel\*

*Leiden University, Cognitive Psychology Unit & Leiden Institute for Brain and Cognition, Leiden, The Netherlands*

Received 22 January 2014; accepted 23 January 2014

Available online 23 January 2014

Communicated by L. Perlovsky

In this issue, van Elk, van Schie and Bekkering [4] succeed in presenting a well-organized, highly structured, and impressively complete summary of the available research on how knowledge may affect action control. However, their theoretical efforts remain unconvincing for three reasons.

First, even though the suggested framework is presented as a general “model” of how people acquire, store, and make use of action knowledge, the authors restrict their theoretical reasoning to unimanual reaching for/grasping of objects—a surprisingly small subset of possible actions. Other actions often do not fit into the authors’ schema: walking, talking, singing, gesturing, typing, looking, driving, dancing, playing music, doing sports, and reading do not refer to particular objects and/or have goals other than being directed towards particular spatial locations—and yet they all rely on “what” and “how” knowledge. This suggests that the “multimodal object representation” box of the model is too specific to apply to action in a broader sense; it should rather refer to all sorts of events that people can aim at, generate, and imply by means of voluntary actions—as suggested by the Theory of Event Codes [2]. Event representations may also refer to the action itself, such as with a ballet figure, suggesting that the “object representation” box might sometimes merely consist of representations of the sensory consequences the action is intended to produce. It thus makes little sense to distinguish between “multimodal object representations”, “proprioceptive consequences” (which, strangely enough, are distinguished from other sensory consequences), and “sensory consequences”—this may just be different examples of “event representations” [2].

Second, the remaining concepts of the framework are ill-defined: Distinguishing between functional and manipulation knowledge makes sense for object grasping but not for many other actions, including gesturing, dancing, and singing. Worse, the authors’ definition that the action-related knowledge comprises of all “the procedural or manipulation knowledge” that enables us to carry out meaningful actions renders the model tautological: if humans can carry

---

DOI of original article: <http://dx.doi.org/10.1016/j.plrev.2013.11.005>.

\* Correspondence to: Leiden University, Department of Psychology, Cognitive Psychology Unit, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands.

E-mail address: [hommel@fsw.leidenuniv.nl](mailto:hommel@fsw.leidenuniv.nl).

out such actions, which they do, they must have and apply that knowledge—which renders the main claim of authors logically necessary and, thus, empirically unfalsifiable.

Third, the other major claim regarding the concept of hierarchy is also questionable. The concept comes in one of two versions and refers to either a control hierarchy, where higher levels control lower levels, or to differential temporal stability, where higher-level representations are more stable over time than lower-level representations are [3]. But it makes little sense to assume that functional knowledge controls sensory consequences (as they simply differ in content) or that the former has greater temporal stability than the latter. Once functional knowledge has served to select appropriate action, it can be deactivated immediately, while representations of the sensory consequences would need to be kept active until the action is actually carried out—otherwise it would be impossible to evaluate the outcome. Hence, if anything, representations of sensory consequences should be more temporally stable than functional knowledge. Moreover, the distinction between proprioceptive and sensory consequences is not only semantically odd (as proprioception provides sensory data) but also empirically unjustified: Studies comparing the contribution of inconsistent proprioception and visual sensory representations to action control show that visual action representations easily overwrite proprioceptive representations (e.g., [1]). This leaves only two boxes in the suggested action-hierarchy model (functional knowledge and sensory consequences), which in the case of object-unrelated actions are likely to be strongly dependent on each other. It is difficult to see what sense it would make to assume some hierarchical relationship between these two remaining sources of knowledge—be it in the sense of control or temporal relationships.

## References

- [1] Hommel B. Inverting the Simon effect by intention: determinants of direction and extent of effects of irrelevant spatial information. *Psychol Res* 1993;55:270–9.
- [2] Hommel B, Müsseler J, Aschersleben G, Prinz W. The theory of event coding (TEC): a framework for perception and action planning. *Behav Brain Sci* 2001;24:849–78.
- [3] Uithol S, van Rooij I, Bekkering H, Haselager P. Hierarchies in action and motor control. *J Cogn Neurosci* 2012;24:1077–86.
- [4] van Elk M, van Schie H, Bekkering H. Action semantics: A unifying conceptual framework for the selective use of multimodal and modality-specific object knowledge. *Phys Life Rev* 2014;11(2):220–50 [in this issue].