European Journal of Social Psychology
Eur. J. Soc. Psychol. **39**, 1191–1193 (2009)
Published online in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/ejsp.682

Commentary

Taking the grounding problem seriously

BERNHARD HOMMEL*
Institute for Psychological Research & Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands

Cognitivistic approaches to human thought and behavior have been very successful in tracing, identifying, and analyzing the processes and mechanisms underlying human cognition. However, they have created and strengthened the image of humans as couch potatoes that have incredibly interesting thoughts and mental simulations but hardly translate any of those into real action. Moreover, the strategy to construe human cognition as symbol manipulation, that is, as the processing of abstract units of information that are not related to, and grounded in real-world events in straightforward and theoretically well-understood ways, has led to a growing dissatisfaction with traditional cognitivistic approaches. A promising alternative is the embodied-cognition approach that construes cognition and cognitive representations as emerging from, and as being grounded in perceptual, affective, and action-related states and processes (see Pecher & Zwaan, 2005). Ideally, the meaning of a perceived or produced event can be reduced entirely to the sensorimotor (and affective) states and processes directly involved in its perception or production, so that cognitive representations lose their explanatory overhead and become mere summaries of, or pointers to well-understood sensorimotor component processes—as in the Theory of Event Coding (TEC; Hommel, 2009; Hommel, Müsseler, Aschersleben, & Prinz, 2001).

Rueschemeyer, Lindemann, van Elk, and Bekkering (2009; henceforth RLvEB) make an attempt to apply an embodied-cognition approach to the interface between language and action, and they put forward two major claims: That the new concept of ''semantic resonance'' is needed to understand how language and action control interact and that a dedicated cognitive control mechanism is needed to regulate this interaction and tailor it to the situation at hand. I strongly sympathize with the general approach defended by RLvEB because the embodied-cognition approach is healthy in forcing us (more than traditional cognitivistic approaches) to think of how mind, brain, and body interact, and how our cognitions relate to our physical and social environment, and because relating nonverbal perception and action to verbal perception and action is likely to be very productive both theoretically and empirically. At the same time, however, I have doubts whether the concrete suggestions RLvEB make really advance our understanding of embodied cognition in general and of the relationship between language and action in particular. In fact, I believe that their approach actually represents a significant setback on the way to a comprehensive theory of embodied cognition. As I will explain in the following, this is because their approach increases, rather than decreases, the gap between cognition and the sensorimotor processes that according to be embodied-cognition perspective should represent its basis and substrate. It, thus, effectively disembodies cognition and, as I will also explain, it does so without any need, that is, in the face of obvious theoretical alternatives that perfectly fit with the notion of embodied cognition.

For the sake of the argument, let us take an extreme alternative and assume that the semantics of human perception and action, and of any language-related cognitive process, can be entirely reduced to sensorimotor (and affective) states and processes, as implied by TEC. Hence, let us assume that there are no separate semantic representations or processes or resonance—the meaning of a perceived or produced event would thus consist of and be embodied by the total of activated sensorimotor states. If so, one would clearly expect that semantically related objects and actions interact, simply because

*Correspondence to: Bernhard Hommel, Department of Psychology, Cognitive Psychology Unit, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands. E-mail: hommel@fsw.leidenuniv.nl

representing their meaning involves overlapping components. One would also expect that processing action words activates motor areas of the brain, as it is this activation that realizes the representation of meaning, or at least contributes to this representation. However, rather than interpreting these and other observations as pointing to the motoric (or sensorimotor) embodiment of meaning, RLvEB take them to ''demonstrate effects of semantic processing on the motor system.'' Even though they briefly consider ''that motor and semantic representations operate on the same cognitive codes,'' they eventually contrast ''semantic resonance'' with ''motor resonance'' (excluding the possibility of ''perceptual resonance'' for unclear reasons). For this contrast to be meaningful, one would need to assume that these two types of resonance involve different codes, which amounts to giving up the idea that semantic representations are no more than the sum of sensorimotor representations.

This theoretical move is not enforced by any data discussed in this context. Take the observation that the processing of words like ''mouth'' is facilitated by preparing a semantically related action like moving a cup to the mouth (Lindemann, Stenneken, van Schie & Bekkering, 2006). The authors insist that this effect must be semantic, as detecting a letter in these words was not affected by action preparation. From an embodied-cognition perspective, processing the meaning of the word is likely to include, or even require, the activation of perceptual, and perhaps even sensorimotor representations of the word's referent, such as visual and proprioceptive representations of a mouth. Action preparation also uses such codes: Actions are cognitively represented in terms of their sensory consequences (action effects, see Hommel, 1996) and codes of these consequences are involved in planning the action (Hommel, 2009; Melcher, Weidema, Eenshuistra, Hommel & Gruber, 2008). The codes that produced the facilitation reported by Lindemann et al. are therefore likely to be of sensorimotor nature—preparing a mouth-related action involves activating the same codes that embody the meaning of the word ''mouth''—so that the critical ''resonance'' was ''sensorimotor'' but not ''semantic''. Note that codes of action effects are assumed to mediate intentional action planning but not passive movements, which fits with the observation that action-induced priming of words is restricted to intentional actions (Rueschemeyer, van Rooij, Lindemann & Bekkering, submitted).

Similar considerations apply to the study of Rueschemeyer, Pfeiffer and Bekkering (submitted). They showed that preparing an action towards the body facilitates the semantic processing of words denoting objects that are typically moved towards the body (such as ''cup''), whereas, preparing an action away from the body facilitated the processing of words denoting objects that are typically moved away from the body (such as ''hammer''). Again it is clear that facilitation was associated with, and presumably due to, an overlap of the sensorimotor representations of the actions and the word's referents, suggesting that what mattered was ''sensorimotor resonance.''

It is more difficult to evaluate the implications of the study by van Elk, van Schie, and Bekkering (2008), where participants prepared meaningful or meaningless actions (e.g., moving a cup to the mouth or to the eye, respectively) before judging the category of a visually presented word. Preparing meaningful actions elicited a larger N400 for goal-incongruent words (e.g., the word ''mouth'' when having prepared a movement towards the eye) but preparing meaningless actions did not. RLvEB take this to demonstrate that ''a comprehensible goal end-state activates semantic information, while a nonsensical goal end-state does not,'' but this conclusion suffers from a number of problems. First, the manipulation of meaningfulness was entirely confounded with novelty and practice: Participants must have moved many more cups to their mouth than to their eye. As compared to practiced actions, the planning of novel actions is known to be more on-line (rather than relying on overlearned plans retrieved from memory), more in terms of sequences of movement components (rather than one coherent whole), and thus more resource demanding. The lack of a N400 effect may have been due to any of these factors: Unfamiliar actions may have been programmed in terms of end locations rather than body parts, so that planning a movement towards the eye did not involve any representation of the eye; planning those actions may have exhausted too many resources to detect the semantic incongruence of the words, etc. Another major problem is that the congruence between words and action goals had no impact on the speed in which the words were semantically processed. If the N400 effect in the meaningful condition would really indicate semantically-based interactions between word and action goal, and if these interactions would have really been induced by the obligatory activation of semantic information, it is difficult to see why the performance of the participants was entirely independent of these interactions. Whatever process was causing the N400 effect, this process cannot have played an essential role in evaluating the meaning of words.

RLvEB conclude by arguing that the meaning of actions is context dependent and that context-sensitive coding presupposes the operation of a dedicated semantic-selection mechanism. Even though this mechanism is assumed to ''closely mirror'' the operation of mechanisms responsible for perceptual and action selection, it is claimed to be different

DOI: 10.1002/ejsp

from those. An example from an old study of mine (Hommel, 1993) may help to clarify that this is not a necessary, and not even the most obvious conclusion. The study was on the Simon effect, which refers to the observation that people respond faster to stimuli that spatially correspond to the response. But what is a response? I connected left and right response keys to lights on the opposite side, so that pressing the left key, say, switched on a right light. When participants were instructed to ''press a key'' in response to a stimulus, they were faster if stimulus and response keys corresponded—the Simon effect. However, when participants were instructed to ''switch on a light'' in response to a stimulus, they were faster if stimulus and light corresponded—an inversion of the Simon effect! Obviously, participants coded the actions in terms of their contextual meaning—the same action that was ''left'' in the key-pressing context was ''right'' in the light-switching context. As elaborated elsewhere (Hommel, 2009, in press; Hommel et al., 2001), this and other observations suggest two cognitive mechanisms. First, every action results in the integration of the motor patterns underlying it and the sensory consequences it produces—this is the ideomotor principle (Lotze, 1852; James, 1890). Second, the codes representing a to-be-produced or to-be-perceived event are weighted with respect to the task relevance of the dimension they are defined on—this is TECs intentional-weighting principle (Hommel et al., 2001). A key instruction makes keys more relevant than lights, while a light instruction does the opposite, so that the former induces stronger weighting of key positions and the latter stronger weighting of light positions in the cognitive representation of the alternative actions. Hence, selective attention to sensorimotor attributes (like features of action effects) changes the meaning of an action, suggesting that the context-sensitive coding of action meanings does not rely on or require any dedicated mechanism that operates on semantic information.

Taken altogether, the evidence discussed by RLvEB provides strong support for an embodied-cognition approach but does not suggest, or even require the assumption of semantic representations separate from sensorimotor codes, or of any special mechanisms dedicated to operate on some semantic representations. This is not to say that every possible semantic aspect is necessarily based on sensorimotor representations (even though this may well be the case; see Pecher & Zwaan, 2005), but sensorimotor coding seems sufficient to account for the examples RLvEB have considered. Given that one of the major objectives of the embodied-cognition perspective is tackling the grounding problem, which again calls for the successful reduction of abstract symbols to sensorimotor processes, it seems important to resist the unnecessary postulation of abstract representations as far as possible and to treat semantic representations as *explanandum* but not as *explanans*.

REFERENCES

Hommel, B. (1993). Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information. *Psychological Research*, *55*, 270–279.

Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research*, *59*, 176–186.

Hommel, B. (2009) Action control according to TEC (theory of event coding). *Psychological Research*, *73*, 512–526.

Hommel, B. (in press) Grounding attention in action control: The intentional control of selection. In B. J. Bruya (Ed.), *Effortless attention: A new perspective in the cognitive science of attention and action* Cambridge, MA: MIT Press.

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*, 849–937.

James, W. (1890). *The principles of psychology* (Vol. 2). New York: Dover Publications.

Lindemann, O., Stenneken, P., van Schie, H. T., & Bekkering, H. (2006). Semantic activation in action planning. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 633–643.

Lotze, R.H. (1852). *Medicinische Psychologie oder die Physiologie der Seele*. Leipzig: Weidmann'sche Buchhandlung.

Melcher, T., Weidema, M., Eenshuistra, R. M., Hommel, B., & Gruber, O. (2008). The neural substrate of the ideomotor principle: An event-related fMRI analysis. *NeuroImage*, *39*, 1274–1288.

Pecher, D., & Zwaan, R. A. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language and thinking*. Cambridge: Cambridge University Press.

Rueschemeyer, S.-A., Lindemann, O., van Elk, M., & Bekkering, H. (2009). Embodied cognition: The interplay between automatic resonance and selection-for-action mechanisms. *European Journal of Social Psychology, 39*, 1180–1187. DOI: 10.1002/ejsp.662

Rueschemeyer, S.-A., Pfeiffer, C., & Bekkering, H. (submitted). Goal representations of manipulable object words.

Rueschemeyer, S.-A., van Rooij, D., Lindemann, O., & Bekkering, H. (submitted). Effects of intentional action on language processing.

van Elk, M., van Schie, H. T., & Bekkering, H. (2008). Semantics in action: An electrophysiological study on the use of semantic knowledge for action. *Journal of Physiology Paris*, *102*, 95–100.