

Dissociable Neural Systems Support Retrieval of *How* and *Why* Action Knowledge

Robert P. Spunt, Emily B. Falk, and Matthew D. Lieberman

University of California, Los Angeles

Psychological Science
 21(11) 1593–1598
 © The Author(s) 2010
 Reprints and permission:
sagepub.com/journalsPermissions.nav
 DOI: 10.1177/0956797610386618
<http://pss.sagepub.com>



Abstract

In everyday discourse, people typically represent actions in one of two ways: how they are performed or why they are performed. In the present study, we determined the neural systems that support these natural modes of representing actions. Participants underwent functional magnetic resonance imaging while identifying how and why people perform various familiar actions. Identifying how actions are performed produced activity in premotor areas that support the execution of actions and in higher-order visual areas that support the perception of action-related objects; this finding supports an embodied view of action knowledge. However, identifying why actions are performed preferentially engaged areas of the brain associated with representing and reasoning about mental states; these areas were right temporoparietal junction, precuneus, dorsomedial prefrontal cortex, and posterior superior temporal sulcus. Our results suggest that *why* action knowledge is not sufficiently constituted by information in motor and visual systems, but requires a system for representing states of mind, which do not have reliable motor correlates or visual appearance.

Keywords

semantic memory, motor processes, theory of mind, social cognition, cognitive neuroscience

Received 12/11/09; Revision accepted 5/17/10

Now what is an action? Not one thing, but a series of two things. . . . The volition or intention to produce the effect, is one thing: the effect produced in consequence of the intention, is another thing: the two together constitute the action. (Mill, 1871/1961, pp. 68–69)

Actions are constituted by both movement and mind—sitting down in a chair is something quite different from falling into a chair, even if the position and motion of the body are identical in the two cases. Actions possess both a *how*—the executed movements of the body and its mechanical interactions with the physical world around it—and a *why*—the relatively disembodied motives, beliefs, and intentions of the actor. We used functional magnetic resonance imaging to identify the neural systems that support *how* and *why* action knowledge.

The majority of research on the organization of action knowledge in the brain has operated within an embodied-cognition framework, which proposes that action knowledge is grounded in the motor and visual systems that underlie the execution of actions and the perception of action-related objects (Pulvermüller, 2005). This proposal is supported by behavioral studies demonstrating motor activity during comprehension of linguistic action stimuli (e.g., Glenberg & Kaschak, 2002), neuropsychological studies demonstrating

action-knowledge impairments associated with damage to visual and motor areas of the brain (e.g., Tranel, Manzel, Asp, & Kemmerer, 2008), and neuroimaging studies demonstrating activation of visual and motor systems during the comprehension of linguistic action stimuli or the retrieval of actions associated with objects (e.g., Hauk, Johnsrude, & Pulvermüller, 2004; Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995; for reviews, see Mahon & Caramazza, 2009; Noppeney, 2008). Such research reliably implicates a left lateralized network of regions at the junction of the occipital and temporal cortices, in the lateral parietal cortex, and in the motor and premotor cortices in the frontal lobes.

Research on the neural organization of action knowledge has focused on concrete motor actions and tools. Such research provides a clear picture of the neural systems involved in knowing how to execute an action, but may provide less information regarding the neural representation of why actions are executed in the first place. A dissociation between representations of how and why actions are performed is suggested by

Corresponding Author:

Robert P. Spunt, Department of Psychology, 1285 Franz Hall, University of California, Los Angeles, CA 90095-1563
 E-mail: spunt@psych.ucla.edu

work showing that action intentions are represented by inferior parietal cortex and the supplementary motor area (SMA; Desmurget et al., 2009; Fried et al., 1991; Lau, Rogers, Haggard, & Passingham, 2004) and by research showing that, at least in parietal cortex, intentions are represented abstractly, independent of movement execution (Anderson & Buneo, 2002; Desmurget & Sirigu, 2010).

Social-cognitive psychologists also distinguish *how* and *why* in the context of the conceptual representations of the actions both of the self and of other people, with *how* focusing attention on the concrete mechanics of an action and *why* focusing attention on the relatively abstract mental states that causally explain an action (Kozak, Marsh, & Wegner, 2006; Vallacher & Wegner, 1987). Work in the social-cognitive neurosciences has identified a set of regions collectively called the mentalizing, or theory-of-mind, system that reliably responds when individuals are induced to represent or reason about mental states: These regions include right temporoparietal junction (TPJ), dorsomedial prefrontal cortex, precuneus, posterior superior temporal sulcus (pSTS), and the temporal poles (e.g., Gallagher et al., 2000; for reviews, see Frith & Frith, 2003; Lieberman, 2010). We recently showed that during action observation, areas of this mentalizing system are engaged to the extent that individuals think about why the action was being performed (Spunt, Satpute, & Lieberman, 2010). In the present study, we used an ecologically valid free-recall task to investigate the neural bases of retrieving *how* and *why* knowledge for linguistic action stimuli.

Method

Participants

Sixteen right-handed, native-English-speaking participants (8 female and 8 male; mean age = 21.67 years, $SD = 3.64$) were recruited from the University of California, Los Angeles

(UCLA) subject pool and received financial compensation for participating.

Experimental stimuli

Stimuli consisted of 20 descriptions of familiar actions adapted from the Behavior Identification Form (Vallacher & Wegner, 1989) and from descriptions provided by participants in our previous study (Spunt et al., 2010). Table 1 features example stimuli and typical responses collected in a separate pilot study regarding *how* and *why* questions.

Experimental paradigm and procedure

On different trials, participants were asked how and why people typically perform each action. For *how* trials, participants were instructed to think of one necessary part of performing the action; for *why* trials, participants were instructed to think of one plausible motive for performing the action. In each trial, participants were instructed to silently think of a single response and to press a button with their right index finger once they had their response in mind. There were 20 described actions, each of which appeared twice in the session, once for *how* and once for *why*, for a total of 40 trials. The order of trials was pseudorandomized, with the constraint that half the described actions would first appear in a *how* trial, and the other half would first appear in a *why* trial. Participants had 5 s to make each response, and successive trials were separated by an interval that varied from 1 to 6 s.

Prior to entering the scanner, participants were introduced to the task and asked to answer how and why for three actions not used in the experimental session. Immediately prior to the functional scan, participants performed a demonstration version of the task featuring four actions not used in the experimental session. Stimulus presentation was implemented using eM's Stimulus Software (MSS; Falk, 2009). Video goggles

Table 1. Sample Stimuli With Typical Responses From *How* and *Why* Trials in the Behavioral Pilot Study ($N = 26$)

Stimulus	How do people do it?	Why do people do it?
Brush teeth	Use a toothbrush	To clean teeth
Diet	By eating less	To lose weight
Drink coffee	In a mug/cup	To stay awake
Eat ice cream	With a spoon	Because it tastes good
Lift weights	Use hands/arms	To be stronger
Listen to music	With headphones	For enjoyment
Read newspaper	With their eyes	To stay informed
Surf Internet	With a computer	To find information
Take aspirin	Swallow a pill	To relieve pain
Vote in an election	With a ballot	To express an opinion
Watch TV	Turn on television	To relax/entertain oneself
Write in a diary	Use a pen	To express thoughts

compatible with magnetic resonance imaging and a button box were used to present the task and record responses.

Image acquisition

Imaging data were acquired using a Siemens (New York, NY) Trio 3-T magnetic resonance imaging scanner at the UCLA Ahmanson-Lovelace Brain Mapping Center. For each participant, we acquired 152 functional T2*-weighted echo-planar images (EPIs) with the following parameters: slice thickness = 4 mm, 33 axial slices, repetition time (TR) = 2,000 ms, echo time (TE) = 30 ms, flip angle = 75°, matrix = 64 × 64, field of view (FOV) = 220 mm. We also acquired a high-resolution T2-weighted matched-bandwidth anatomical scan (same parameters as EPIs, except TR = 5,000 ms, TE = 34 ms, flip angle = 90°) and a magnetization-prepared rapid-acquisition gradient echo anatomical scan (slice thickness = 1 mm, 160 slices, TR = 2,300 ms, TE = 2.47 ms, flip angle = 8°, FOV = 256 mm).

Image analysis

Functional data were analyzed using Statistical Parametric Mapping Version 5 (Functional Imaging Laboratory, 2005) software. Image volumes were slice-time corrected, realigned to correct for head motion, normalized into Montreal Neurological Institute space (resampled at 3 × 3 × 3 mm), and smoothed with a 10-mm Gaussian kernel, full width at half maximum. The resulting time series was high-pass filtered to 1/100 Hz.

First-level models were specified with *how* and *why* trials modeled as events convolved with a canonical hemodynamic response function. To control for differences in time on task, we parametrically modulated the *how* and *why* regressors as a function of response time (RT; time of button press minus time of stimulus onset). We also included the six motion parameters as covariates of no interest. Linear contrasts were then applied to the design to determine regions more active in the *how* trials than in the *why* trials (*how* > *why* contrast) and regions more active in the *why* trials than in the *how* trials (*why* > *how* contrast). The resulting contrast images were then gray-matter masked and subjected to a random-effects analysis to investigate effects at the group level. Whole-brain interrogations were conducted using a false-discovery-rate-corrected *p* value of .05, with a cluster-size threshold of 5 voxels. For visual presentation, lateral clusters were surface rendered using the SPM SurfRend toolbox Version 1.0.2 (Kahn, 2008), and medial clusters were overlaid on the average of the participants' normalized anatomical images.

Results

Consistent with previous findings using a similar task (Spunt et al., 2010), results showed that participants took longer identifying how actions are performed ($M = 3.14$ s, $SD = 0.87$) than

why actions are performed ($M = 2.88$ s, $SD = 0.76$), $t(15) = 3.112$, $p = .007$, $\eta_p^2 = .39$. Although we included RT as a regressor in the primary analysis, we conducted a secondary analysis on a subset of *how* and *why* trials that were collectively matched on RT. All effects reported were also observed in the secondary analysis (details are available in the Supplemental Material available online).

We first performed a whole-brain search for regions showing a larger response when identifying how than when identifying why actions are performed (Table 2). As depicted in Figure 1, we observed a primarily left lateralized network of regions associated with the representation of motor actions, body parts, and tools. These regions included posterior middle temporal gyrus at the lateral occipitotemporal junction, fusiform gyrus at the medial occipitotemporal junction, posterior parietal cortex, and dorsal and ventral aspects of premotor cortex.

Next, we performed a whole-brain search for regions showing a larger response when identifying why than when identifying how (Table 2). As depicted in Figure 1, we observed four regions associated with the representation of mental states: right TPJ, precuneus, dorsomedial prefrontal cortex, and right pSTS.

Discussion

Using a novel, ecologically valid paradigm for eliciting conceptual knowledge about actions, we found that *how* action knowledge was associated with areas of the brain's motor system for executing actions and the brain's visual system for recognizing action-related objects. This finding replicates past research associating these systems with the conceptual representation of motor actions, body parts, and tools (Mahon & Caramazza, 2009; Noppeney, 2008). Conversely, *why* action knowledge was associated with the brain's system for representing agency and for reasoning about mental states (Decety & Lamm, 2007; Desmurget & Sirigu, 2010; Frith & Frith, 2003). This finding confirms the proposition that high-level action knowledge involves mental-state attribution (Kozak, Marsh, & Wegner, 2006). Together, these findings show that the neural system activated when people represent actions depends on whether they are considering how or why the action is performed.

Embodied-cognition frameworks have dominated the theoretical discussion of action knowledge (Mahon & Caramazza, 2008). The results of the present study support these frameworks but suggest an important qualification. Our findings provide support by offering another instance of motor- and visual-system engagement during conceptual processing of action stimuli. However, our data qualify support by identifying a boundary condition on such engagement: When individuals identify *how* action knowledge, motor and visual systems are preferentially engaged; however, when individuals identify *why* action knowledge, a system for mental-state representation shows preferential engagement. We suggest that in representing actions, the involvement of motor and visual

Table 2. Activations Observed in Whole-Brain Analyses for the Contrasts *How* > *Why* and *Why* > *How*

Contrast and brain region	MNI coordinates			t(15)	No. voxels
	x	y	z		
<i>How</i> > <i>why</i>					
Occipitotemporal junction					
Fusiform gyrus (L)	-36	-36	-15	6.80	197
Posterior middle temporal gyrus (L)	-57	-63	-3	5.86	—
Parietal cortex					
Posterior parietal cortex (L)	-39	-75	30	6.83	209
Dorsal precuneus (L)	-12	-60	60	5.67	58
Frontal cortex					
Dorsal premotor cortex (L)	-21	6	60	6.10	159
Dorsal premotor cortex (R)	24	6	66	4.51	15
Ventral premotor cortex (L)	-33	6	33	4.90	54
Dorsolateral prefrontal cortex (L)	-45	33	18	4.38	20
Dorsolateral prefrontal cortex (R)	30	42	33	4.99	27
<i>Why</i> > <i>how</i>					
Parietal cortex					
Temporoparietal junction (R)	57	-63	30	8.39	83
Precuneus (R)	6	-57	36	5.97	157
Frontal cortex					
Dorsomedial prefrontal cortex	0	48	42	5.48	178
Medial prefrontal cortex (R)	12	60	18	5.44	—
Temporal cortex					
Superior temporal cortex (L)	-63	-12	-3	5.32	10
Posterior superior temporal sulcus (R)	63	-33	-3	4.24	5
Cerebellum (L)	-18	-87	-36	5.58	76

Note: L and R refer to the left and right brain hemispheres, respectively; x, y, and z refer to Montreal Neurological Institute (MNI) coordinates in the left-right, anterior-posterior, and inferior-superior dimensions, respectively; t refers to the t score at those coordinates (local maxima).

systems depends on the extent to which the representation of the action can be embodied in motor events or in concrete objects. As is clear from the responses listed in Table 1, identifying how an action is performed entails representing an action in terms of its concrete embodiments, that is, the tools, body parts, and motor actions necessary for its completion. However, identifying why the same action is performed entails representing the action in terms of relatively more abstract, disembodied intentional and motivational states that causally explain performance of the action. Such states do not have reliable motor correlates or visual appearances, and thus rely less on motor and visual systems during their representation.

Thus, these results support recent propositions that embodied cognition provides only a framework for characterizing the neural bases of concrete conceptual knowledge and cannot account for abstract concepts (Mahon & Caramazza, 2008; Toni, de Lange, Noordzij, & Hagoort, 2008). With this said, it should be emphasized that these data do not rule out a constitutive role for sensory information in *why* action knowledge. Indeed, both right TPJ and pSTS are associated with encoding higher-order sensory properties of stimuli, such as spatial

properties (in the case of right TPJ; Decety & Lamm, 2007) and biological motion (in the case of pSTS; Allison, Puce, & McCarthy, 2000). Thus, it may be artificial to posit “embodied” and “disembodied” as discrete categories of concepts; rather, like the concrete-abstract dimension, they may best be characterized as two poles of a continuum.

These results converge with research findings demonstrating that neural activity underlying the representation of an intention to move is independent of the execution of movement (Anderson & Buneo, 2002; Desmurget & Sirigu, 2010). Recently, Desmurget et al. (2009) found that the experience of intending to move could be induced by electrically stimulating a region of right inferior parietal cortex overlapping with the right TPJ cluster associated with *why* action knowledge in the present study. This observation accords with neuroimaging research showing modulation of agency in right TPJ in response to manipulations of agency (e.g., Farrer & Frith, 2002; Ruby & Decety, 2001). Moreover, in a recent meta-analysis, Decety and Lamm (2007) found that the area of right TPJ activation observed in studies of agency shows substantial overlap with the area of right TPJ activation observed in studies of

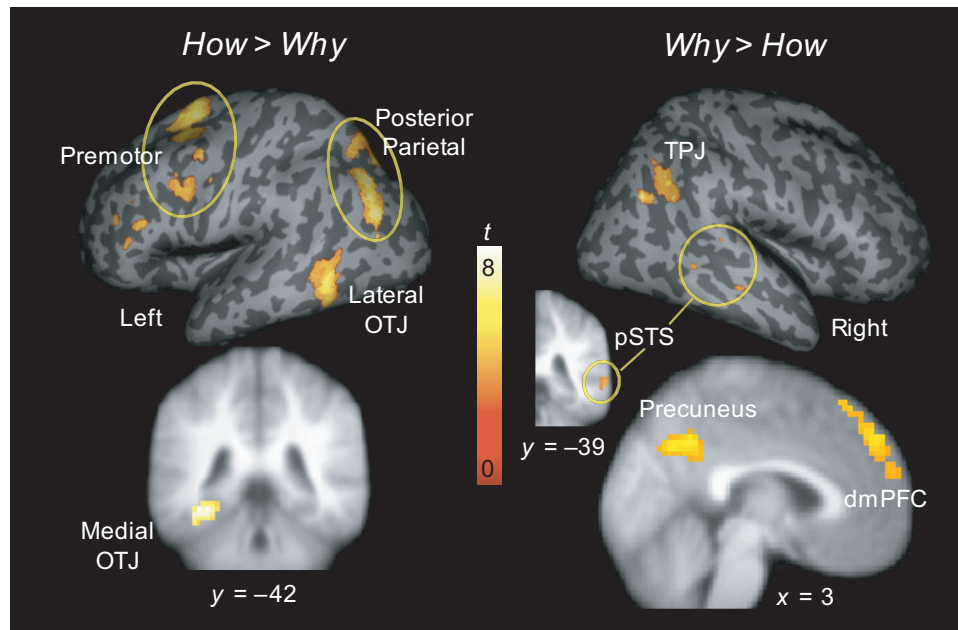


Fig. 1. Contrast maps showing activations in the *how* > *why* and *why* > *how* whole-brain analyses. For all images, the threshold used was a false-discovery-rate-corrected p value of .05, except for the coronal image displaying posterior superior temporal sulcus (pSTS), for which the threshold was .005 (uncorrected) to show spatial extent. OTJ = occipitotemporal junction; dmPFC = dorsomedial prefrontal cortex; TPJ = temporoparietal junction.

mental-state reasoning. Given this, we speculate that during the retrieval of why action knowledge, right TPJ may be involved in representing either the specific intentional state described by the action stimulus or the sense of agency more generally. The SMA has also been associated with the representation of motor intentions (Fried et al., 1991; Lau et al., 2004), but was not associated with why action knowledge in the present study. This might be explained by the observation that SMA activity produces an involuntary urge to move and is more tightly coupled to movement execution (Fried et al., 1991), and thus may be less relevant to agentic representations of action (Desmurget & Sirigu, 2010).

Two limitations of the present study suggest valuable directions for future research. First, participants answered all questions from a generic third-person perspective (e.g., How do *people* do X ?). Given that first- and third-person perspectives recruit different brain regions during mental simulation of actions (Ruby & Decety, 2001), answering *why* and *how* questions from a first-person perspective (e.g., How do *you* do X ?) might produce at least partially different results.

Second, the present study does not speak to the question of whether the neural systems engaged during the retrieval of *how* and *why* action knowledge are modulated by dimensions of the action stimulus. We controlled for such stimulus effects by using identical stimuli in *how* and *why* trials. However, such effects present a valuable direction for future research. For example, the extent to which the action stimulus is represented concretely (e.g., “grip a toothbrush”) or abstractly (e.g., “maintain dental hygiene”) may modulate engagement of

motor and visual systems, regardless of whether individuals represent how or why an action is performed. This is because the conceptual content of how and why varies as a function of the prepotent level of action identification (Vallacher & Wegner, 1987). This is illustrated by the fact that a response to *how* for one action description (e.g., How does one *maintain hygiene*? Answer: *brush teeth*) can be identical to a response to *why* for a different action description (e.g., Why does one *grip a toothbrush*? Answer: *brush teeth*). This suggests an important question for future research: In the neural representation of action knowledge, do how and why refer to mutually exclusive mind-sets (Freitas, Gollwitzer, & Trope, 2004), or is their only difference in the concreteness of the concepts they tend to evoke?

Our results confirm the proposition that there is more to action than acting. Actions imply both movements of the body and belief-desire states of the mind. To the extent that individuals represent one or the other, dissociable neural systems are involved.

Acknowledgments

The authors thank Catherine Mulvenna for assistance with data collection and two anonymous reviewers for generous feedback. For support, the authors thank the Brain Mapping Medical Research Organization, Brain Mapping Support Foundation, Pierson-Lovelace Foundation, Ahmanson Foundation, William M. and Linda R. Dietel Philanthropic Fund, Tamkin Foundation, Jennifer Jones-Simon Foundation, Capital Group Companies Charitable Foundation, Robson Family, and Northstar Fund.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information may be found at <http://pss.sagepub.com/content/by/supplemental-data>

References

- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, 4, 267–278.
- Anderson, R.A., & Buneo, C.A. (2002). Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience*, 25, 189–220.
- Decety, J., & Lamm, C. (2007). The role of right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *The Neuroscientist*, 13, 580–593.
- Desmurget, M., Reilly, K.T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science*, 324, 811–813.
- Desmurget, M., & Sirigu, A. (2010). A parietal-premotor network for movement intention and motor awareness. *Trends in Cognitive Sciences*, 13, 411–418.
- Falk, E. (2009). eM's Stimulus Software [Computer software]. Retrieved September 22, 2009, from <http://psych2.dnsalias.org/MSS/>
- Farrer, C., & Frith, C.D. (2002). Experiencing oneself vs. another person as being the cause of an action: The neural correlates of the experience of agency. *NeuroImage*, 15, 596–603.
- Freitas, A.L., Gollwitzer, P., & Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of Experimental Social Psychology*, 40, 739–752.
- Fried, I., Katz, A., McCarthy, G., Sass, K.J., Williamson, P., Spencer, S.S., & Spencer, D.D. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *Journal of Neuroscience*, 11, 3656–3666.
- Frith, U., & Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358, 459–473.
- Functional Imaging Laboratory. (2005). Statistical Parametric Mapping (Version 2) [Computer software]. Retrieved September 22, 2008, from <http://www.fil.ion.ucl.ac.uk/spm/software/spm5/>
- Gallagher, H.L., Happé, F., Brunswick, N., Fletcher, P.C., Frith, U., & Frith, D.C. (2000). Reading the mind in cartoons and stories: An fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, 38, 11–21.
- Glenberg, A.M., & Kaschak, M.P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9, 558–565.
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41, 301–307.
- Kahn, I. (2008). SPM SurfRend (Version 1.0.2) [Computer software]. Retrieved September 12, 2010, from <http://spmsurfrend.sourceforge.net/Downloads/Downloads.html>
- Kozak, M.N., Marsh, A.A., & Wegner, D.M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, 90, 543–555.
- Lau, H.C., Rogers, R.D., Haggard, P., & Passingham, R.E. (2004). Attention to intention. *Science*, 303, 1208–1210.
- Lieberman, M.D. (2010). Social cognitive neuroscience. In S.T. Fiske, D.T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 143–193). New York, NY: McGraw-Hill.
- Mahon, B.Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology (Paris)*, 102, 59–70.
- Mahon, B.Z., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review of Psychology*, 60, 27–51.
- Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., & Ungerleider, L.G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, 270, 102–104.
- Mill, J.S. (1961). *A system of logic: Ratiocinative and inductive* (8th ed.). London, England: Longmans, Green, & Co. (Original work published 1871)
- Noppeney, U. (2008). The neural systems of tool and action semantics: A perspective from functional imaging. *Journal of Physiology (Paris)*, 102, 40–49.
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6, 576–582.
- Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, 4, 456–450.
- Spunt, R.P., Satpute, A.B., & Lieberman, M.D. (2010). Identifying the what, why, and how of an observed action: An fMRI study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience*. Advance online publication. doi:10.1162/jocn.2010.21446
- Toni, I., de Lange, F.P., Noordzij, M.L., & Hagoort, P. (2008). Language beyond action. *Journal of Physiology (Paris)*, 102, 71–79.
- Tranel, D., Manzel, K., Asp, E., & Kemmerer, D. (2008). Naming dynamic and static actions: Neuropsychological evidence. *Journal of Physiology (Paris)*, 102, 80–94.
- Vallacher, R.R., & Wegner, D.M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, 94, 3–15.
- Vallacher, R.R., & Wegner, D.M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, 57, 660–671.