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

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
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 indicated how or why people typically perform each of the described actions (e.g., drink coffee). 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

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>How do people do it?</th>
<th>Why do people do it?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brush teeth</td>
<td>use a toothbrush</td>
<td>to clean teeth</td>
</tr>
<tr>
<td>Diet</td>
<td>by eating less</td>
<td>to lose weight</td>
</tr>
<tr>
<td>Drink coffee</td>
<td>in a mug/cup</td>
<td>to stay awake</td>
</tr>
<tr>
<td>Eat ice cream</td>
<td>with a spoon</td>
<td>because it tastes good</td>
</tr>
<tr>
<td>Lift weights</td>
<td>use hands/arms</td>
<td>to be stronger</td>
</tr>
<tr>
<td>Listen to music</td>
<td>with headphones</td>
<td>for enjoyment</td>
</tr>
<tr>
<td>Read newspaper</td>
<td>with their eyes</td>
<td>to stay informed</td>
</tr>
<tr>
<td>Surf Internet</td>
<td>with a computer</td>
<td>to find information</td>
</tr>
<tr>
<td>Take aspirin</td>
<td>swallow a pill</td>
<td>to relieve pain</td>
</tr>
<tr>
<td>Vote in an election</td>
<td>with a ballot</td>
<td>to express an opinion</td>
</tr>
<tr>
<td>Watch TV</td>
<td>turn on television</td>
<td>to relax/entertain oneself</td>
</tr>
<tr>
<td>Write in a diary</td>
<td>use a pen</td>
<td>to express thoughts</td>
</tr>
</tbody>
</table>
eM’s Stimulus Software (MSS; Falk, 2009). Video goggles 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, η² = .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; Peelen & Downing, 2007). 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 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 activity 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 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


