Journal of Memory and Language 59 (2008) 256–265

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

# Kicking calculators: Contribution of embodied representations to sentence comprehension ☆

Michael E.J. Masson *, Daniel N. Bub *, Christopher M. Warren

*Department of Psychology, University of Victoria, Cornett Building Room A234, P.O. Box 3050 STN CSC, Vic., Canada BC V8W 3P5*

ARTICLE INFO

*Article history:*
Received 28 February 2008
revision received 23 May 2008
Available online 14 July 2008

*Keywords:*
Action representations
Embodied cognition
Sentence comprehension

ABSTRACT

Evocation of motor representations during sentence comprehension was examined by training subjects to make a hand action in response to a visual cue while listening to a sentence. Sentences referred to manipulable objects that were either related or unrelated to the cued action. Related actions pertained either to the function of the object or to its volumetric properties (e.g., shape). The results demonstrate priming of hand actions even though the sentences referred to non-manual interactions with manipulable objects. When sentences described an attentional interaction (looking at the calculator), only functional actions were primed. Sentences describing a non-manual physical interaction (kicking the calculator) primed volumetric as well as functional actions. We describe how seemingly irrelevant motor representations can play a role in constructing sentence meaning.
© 2008 Elsevier Inc. All rights reserved.

There is mounting evidence that the sensory-motor system is intimately involved in conceptual operations (Barsalou, 1999; Gallese & Lakoff, 2005; Glenberg, 1997; Wilson, 2002). In some respects this idea is not so counterintuitive. For example, we can answer questions using our "mind's eye" (e.g., Would this couch look good in our living room?). Research has also demonstrated, however, that we use other sensory-motor imagery, including our "mind's hands", and we do so in a covert and involuntary fashion. For example, Klatzky, Pellegrino, McCloskey, and Doherty (1989) found that judgment of the sensibility of a sentence was facilitated by performing a relevant hand action in advance. Glenberg and Kaschak (2002) had subjects judge the sensibility of sentences by pressing one of two buttons, to indicate sensible, vs. non-sensible. One of the buttons required the subject to move an arm inward in order to press it, whereas the other required the subject to move an arm outward. Subjects responded faster when the sentence described an action congruent with the movement required by the correct response. This study and many others support the view that mental representations of motor activity are automatically evoked by sentences describing actions, and further, that these representations constitute a mental simulation of an action as a necessary component of understanding the sentence (Barsalou, 1999; Gallese & Lakoff, 2005; Glenberg, 1997; Glenberg & Kaschak, 2002; Klatzky et al., 1989; Pecher, Zeelenberg, & Barsalou, 2003; Tucker & Ellis, 2004; Wilson, 2002; Zwaan & Taylor, 2006).

In this article, we wish to develop a more precise understanding of mental simulation as applied to the dynamics of sentence comprehension. Consider the sentence, The lawyer kicked aside the calculator. On one view, simulating the implied action requires a literal depiction of a standing person kicking a small rectangular object with his or her foot. MacWhinney (2005) refers to this mode of simulation as "depictive". But most models of language comprehension assume that meaning is developed piecemeal over time (e.g., Kintsch, 1988; Townsend & Bever, 2001). How, then, are mental simulations generated in relation to the

☆ This research was supported by discovery grants from the Natural Sciences and Engineering Research Council of Canada to Michael Masson and to Daniel Bub, and by a grant from the Perceptual Expertise Network which is funded by the James S. McDonnell Foundation. We are grateful to Marnie Jedynak and Genevieve Roberts for assistance in conducting the experiments.
* Corresponding authors. Fax: +1 250 721 8929.
*E-mail addresses:* mmasson@uvic.ca (M.E.J. Masson), dbub@uvic.ca (D.N. Bub).

0749-596X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2008.05.003

ongoing comprehension of a sentence? The word calculator refers to an object with its own habitual actions including actions pertaining to its conventional use. When the word is encountered in the above sentence, does the meaning of calculator trigger these action sequences along the way to a simulation of a lawyer kicking a calculator? Alternatively, kicking a calculator has nothing to do with manual interactions. Would the representations of hand actions not be evoked as part of sentence comprehension in this context? (See MacWhinney, 2005, on perspective taking.) Finally, how do we go about representing the act of kicking a calculator? After all, we have no experience in performing such actions. What sensory-motor representations do we draw upon to simulate sentence meaning in these cases?

We raise the possibility that sensory-motor representations evoked on-line are unlikely to stand in some simple relationship to the literal meaning of a sentence. Indeed, there is suggestive evidence from neuroimaging studies consistent with this proposition. For example, in an fMRI study, Tettamanti et al. (2005) had subjects listen to sentences that described actions involving either the mouth (I bite the apple), hand (I grasp the knife) or foot (I kick the ball). For each type of sentence, they found that the relevant area of the motor cortex (e.g., the mouth area for bite) showed significant activation. Very similar results were demonstrated by Hauk, Johnsrude, and Pulvermüller (2004) using single words that denoted an action (e.g., chew, step, punch). Importantly, however, both of these studies also showed consistent, though weaker, activation of the supposedly non-relevant motor areas as well (e.g., hand activation for kick). This kind of result raises questions for simulation-based theories of embodied cognition. What part of the verb kick or the phrase kick the ball might the hand be simulating?

Here, we ask this question specifically in regard to sentences describing non-manual interactions, like kicking or stepping on, with manipulable objects (e.g., calculator, toothbrush) that typically evoke well-defined hand actions. We propose that because the sentences describe unfamiliar actions with familiar objects, embodied simulation of these actions is arrived at indirectly, by evoking relevant physical properties derived from previous manual experience. For example, the rigidity of a manipulable object, its inertial properties, size, and weight are all experienced through manual interaction. We argue that part of the meaning of a sentence like John kicked the calculator includes a representation of the physical forces required to accomplish the act and that these forces are arrived at by a process of analogy, by consulting representations built up through handling objects. Indeed, other researchers have made similar arguments in regard to the generality of mental simulation by extension or analogy, particularly as applied to abstract words (Borghi, Glenberg, & Kaschak, 2004; Boroditsky & Ramscar, 2002; Glenberg & Kaschak, 2002; Zwaan, 2004).

If hand actions are evoked by these types of sentences, then it is important to determine which among the variety of different manual actions applicable to a particular object are called into play. Objects are inherently ambiguous with respect to the actions they afford. Carrying out the intention of moving a calculator when clearing a shelf, for example, requires an open grasp with the fingers pointed downward. Adding numbers, in contrast, involves poking keys in a particular order, usually with a forefinger. In previous work we have distinguished between functional and volumetric grasps associated with objects (Bub & Masson, 2006; Bub, Masson, & Cree, 2008; Masson, Bub, & Newton-Taylor, 2008; see also Daprati & Sirigu, 2006). Functional grasps have to do with hand shapes applied when using an object for its intended purpose. Volumetric grasps are those employed when lifting or moving an object. For some objects (e.g., water bottle), these two types of action are virtually the same, but for many others, they are distinct (e.g., calculator, thimble). To enable us to distinguish between these two types of actions, we use only words denoting objects that have clearly distinct functional and volumetric actions (e.g., palm gesture and power grasp for a stapler). Which of these action types, functional or volumetric, might be evoked during the comprehension of a sentence like Mary kicked the calculator? Our assumption is that volumetric grasps generate the knowledge about weight, shape, and rigidity, among other attributes, that would serve as the means to simulate a novel kicking action.

## Measuring hand representations in real time

If the representations of these actions are evoked dynamically during sentence comprehension, we need some way of measuring their presence on-line as meaning is constructed. We describe next a novel methodology that accomplishes this goal. We have developed a procedure that allows us to measure the dynamic representation of hand actions elicited by sentences describing interactions with manipulable objects. We constructed a multi-element response apparatus consisting of generic shapes associated through training with eight distinct hand actions (see Fig. 1). Subjects learned to produce distinct hand actions cued by a photograph of a hand posture, where each cue signaled a particular action carried out on the apparatus. Speeded responses were made to the cue by lifting the preferred hand from a button box and manually carrying out the action on the correct element of the apparatus.

Immediately before carrying out a cued hand action, subjects heard a sentence referring to a manipulable object. The objects were chosen so that they were associated with distinct functional and volumetric hand actions. We assume that any mental representation of hand actions triggered by the meaning of the sentence will prime the execution of similar actions on the response apparatus, provided that the parameters of the cued action approximate those evoked by the object referred to in the sentence. For example, we assume that under certain circumstances, calculator evokes the representation of a volumetric action consistent with a power grasp in which the palm faces downward. If the subject is now cued to grasp an element of the response apparatus requiring a similar hand shape, then we should observe faster motor performance relative to making an unrelated response (a type of motor priming). In support of this assumption,

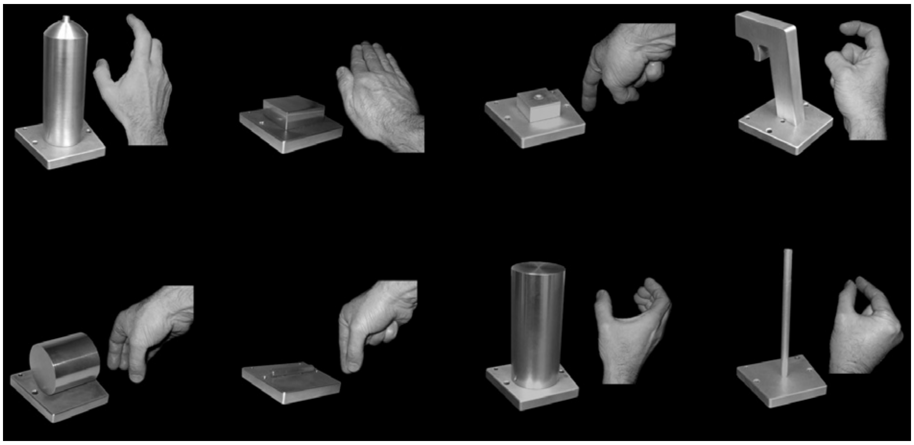

**Fig. 1.** Cues for the eight hand actions and response elements used to perform the actions. The top row shows the hand cues for the four functional actions (aerosol, palm, poke, and trigger) and the associated response elements. The bottom row shows the hand cues for the four volumetric actions (horizontal grasp, horizontal pinch, vertical grasp, vertical pinch) and the associated response elements. A hand cue consisted of a photo of a hand alone.

many cells in the primate (and, presumably, human) anterior–inferior parietal cortex code generic aspects of hand shape (e.g., pinch) without specifying full details of the object that is being grasped (Fagg & Arbib, 1998; Sakata, Taira, Kusunoki, Murata, & Tanaka, 1997). Thus, the parametric specification for the representation for grasping a small cylinder will overlap substantially with that for grasping a small block. Grasping an element of the apparatus will be primed to the extent that there is sufficient overlap between this grasp and the representation of the grasp evoked by the object concept referred to in the sentence. Motor tasks generally show influences of semantic knowledge on early stages of preparation and response execution (Gentilucci, Benuzzi, Bertolani, Daprati, & Gangitano, 2000; Glover, Rosenbaum, Graham, & Dixon, 2004; Lindemann, Stenneken, van Schie, & Bekkering, 2006), whereas later stages of movement are affected more by the actual physical parameters of the object being grasped (Goodale, Milner, Jakobson, & Carey, 1991). Response latency was therefore measured from the onset of the hand cue to contact with the correct response element.

The procedure we have described bears a direct analogy to a paradigm used to examine resolution of lexical ambiguity. In that procedure, an auditory sentence context contains an ambiguous word (e.g., bank) followed at various intervals by a visually presented probe word requiring a classification response. The probe (e.g., money or river) is related to one or the other meaning of the ambiguous word or it is unrelated (Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979). If multiple meanings of an ambiguous word like bank are activated, then both related probes should yield shorter response latencies than an unrelated probe. But if only a single meaning is selected, then priming will be confined to the probe related to that meaning.

A similar logic applies to the motor priming task we have developed. If the word calculator in a sentence evokes both functional and volumetric representations, then cuing either hand action (poke or inverted power grasp) should produce shorter response times than cuing an unrelated action. If selection of one of these representations has occurred, however, then priming should be correspondingly selective.

## Sentence context and hand representations

Our interest is in two types of sentence context, neither of which explicitly entails hand actions: (1) verb phrases like look at or move toward that clearly involve physically orienting to an object or its location but without making contact (referred to here as attention verbs); (2) verb phrases involving physical contact, like step on or kick aside, with no implied hand action (referred to as interaction verbs). This choice of verb phrase types was motivated by the following logic. If mental simulation is simply depictive of the action described by a sentence, then there is no reason to expect hand representations to be evoked as part of comprehending sentences like these. Alternatively, if a word like calculator automatically generates associated hand representations during comprehension, then we should see evidence of their presence despite the ostensive meaning of the sentence.

The two classes of verbs we have chosen are interesting if we consider their demands on spatial processing. Jackendoff and Landau (1992) note that objects referred to in spatial terms (e.g., The book is on the table) are encoded as relatively schematic geometric descriptions. They provide an analogy based on the conventional representation of maps. A map contains location information and the objects populating the map are not distinguished by their shape but are represented in a more stylized manner (e.g., large dots representing cities, smaller dots for towns). Jackendoff and Landau argue that many of the same design principles apply to the representation of spatial relationships in sentences, in particular the need to maintain "...objects as tokens in the representation but to compress their encoding by eliminating most information about their form" (p. 123). If this conjecture is true, then sentences like John looked at the calculator should not evoke detailed volumetric (or shape-based) information. Of course, the meaning of calculator must still be part of the sentence. Some evidence suggests that the function of an object is a core component of its meaning. For example, young children classify novel artifacts on the basis of their intended use (Greif, Kemler Nelson, Keil, & Gutierrez, 2006) rather than appearance. Moreover, in a study similar to that reported here, using visual presentation of sentences describing mental contemplation of objects (e.g., thought about, remembered), we found evidence for priming of functional but not volumetric actions (Masson et al., in press). We may expect, then, that action representations based on the function of an object, but not on its volumetric properties, are evoked by sentences that refer to an agent orienting toward an object without physical contact.

Sentences that describe actions like kicking aside an object or stepping on it entail physical contact, and object properties such as size, shape, mass, and rigidity now become relevant (see Jackendoff & Landau, 1992, for a similar argument with respect to situations in which the object's orientation or distribution through a region is emphasized). We have argued above that if these properties are invoked, they may call up representations of volumetric hand actions and we should see evidence for such action representations during sentence comprehension. At first glance, it may appear somewhat outlandish to expect volumetric hand representations to be evoked by sentences describing actions with feet. But note that listeners presumably have had little or no experience stepping on manipulable objects like calculators or spray cans. The question is, what kind of prior experience is the basis for the motor representations that are constructed in response to such sentences? A plausible possibility is that manual experience of volumetric properties is the means by which the listener simulates these unusual actions. The presence of volumetric hand-action representations in response to sentences dealing with non-manual physical interactions would be of considerable relevance to the question of what mental simulation entails, especially if no such volumetric representations are evoked by sentences that merely refer to glancing toward or looking at an object.

Before investigating the influence of these two types of sentence context, we sought to determine whether hand-action representations are evoked when subjects listen to the name of a manipulable object presented in isolation. Any priming effects found under these circumstances can be used as a benchmark against which to compare the effects of sentence context. In Experiment 1, then, we presented subjects with a series of object names in the auditory modality and presented a visual cue consisting of a hand shaped into a grasp representing the action to be carried out on one of the response elements. The cue was presented either halfway between the onset and offset of the auditory word or immediately after the word's offset. The cued hand action was either related or unrelated to the functional or volumetric action associated with the named object. Testing for priming at two probe points provided a means of investigating possible differences in the time course of evocation of functional and volumetric hand-action representations.

## Experiment 1

### Methods

#### Subjects

Twenty undergraduate students at the University of Victoria participated for extra credit in a psychology course. All subjects were fluent in English and had normal or corrected to normal vision and hearing.

#### Materials

Eight hand actions and names of 18 manipulable objects (e.g., stapler, thimble) related to at least one of those actions were used. Each object name was recorded as a digitized audio file in a female voice. In the context of these objects, four of the hand actions were considered to be functional and four were volumetric. A greyscale photograph was taken for each hand action (see Fig. 1) and these images were used to cue execution of hand actions by the subjects. Mirror-image versions of the hand cues shown in Fig. 1 were used for left-handed subjects.

#### Design

Six of the object names were related to one or another of the functional actions, six were related to one of the volumetric actions, and six were related to one functional and one volumetric action, such that each action had three related objects. On related-prime trials, a cued hand action was primed with a related object name. On unrelated-prime trials, the cued action was unrelated to the object name but was of the same action type (functional or volumetric) as that object's related action. Examples of related and unrelated name-action pairs are shown in Table 1. There were 256 critical trials, with 64 trials in each of the conditions defined by probe point (middle or end of auditory word prime) and type of action (functional or volumetric). Half of the trials in each condition presented a related prime and half presented an unrelated prime. Object names associated with only one action were presented

**Table 1**
Examples of related and unrelated pairs of objects and hand actions used in the experiments

| Action type and object | Related action(s) | Unrelated actions |
|---|---|---|
| *Functional only* | | |
| Keyboard | Poke | Aerosol, palm, trigger |
| Water pistol | Trigger | Aerosol, palm, poke |
| *Volumetric only* | | |
| Thimble | Horiz. pinch | Horiz. grasp, vert. grasp, vert. pinch |
| Toothbrush | Vert. pinch | Horiz. pinch, horiz. grasp, vert. grasp |
| *Both* | | |
| Calculator | Poke, horiz. grasp | Aerosol, palm, trigger, horiz. pinch,vert. grasp, vert. pinch |
| Spray paint can | Aerosol, vert. grasp | Palm, poke, trigger, horiz. grasp, horiz. pinch, vert. pinch |

*Note.* horiz., horizontal; vert., vertical.

as primes approximately 10 times each and object names associated with both a functional and a volumetric action were presented approximately 20 times each (about equally often with each action type). Each of the eight actions were tested equally often in each condition defined by the factorial combination of probe point and prime relatedness.

*Procedure*

Subjects were tested individually using a Macintosh G3 to present object names auditorially over headphones and to display hand cues on a monitor. A second monitor displayed information about the target response to an experimenter who coded the accuracy of each response. Hand actions were made by the subject using a response apparatus that we call a Graspasaurus. This apparatus consists of a curved tray holding up to eight different response elements made of aluminum. Each element is mounted on its own base and the elements are placed side by side on the curved tray, forming a semi-circle in front of the subject so that all elements can be reached easily. Each element accommodated one of the hand actions (e.g., a thick upright cylinder for the vertical grasp, a flat block for the palm action; see Fig. 1). The elements could be arranged in any order on the tray and order of elements on the tray was varied across subjects. A weak electric current passing through the apparatus was disrupted by hand contact, providing a means of measuring response time. A correct response occurred if the subject grasped the cued response element. An error was defined as the execution of a grasp other than the cued grasp or accidental contact with the wrong element of the Graspasaurus, erroneously closing the electrical circuit and signaling the end of the trial.

Subjects were first given 40 trials of training to learn which hand action to make in response to each hand cue and how to execute that action on the corresponding Graspasaurus element. A trial began with the subject using the index finger of her or his preferred hand to depress a key mounted on a response box. The hand cue then appeared on the computer monitor and the subject raised her or his hand and grasped a response element as dictated by the cue. Subjects were then told that they would hear a series of object names while at the same time being cued to make one of the learned hand actions. To ensure that subjects attended to the object names, on a randomly selected 25% of the trials they were prompted after making the manual response to report the object name for that trial. Subjects were instructed to listen to each object name and to make a hand response as quickly and as accurately as possible, as soon as a visual cue appeared. No information was given regarding the relation between the listening and hand-action tasks. Subjects were presented 16 practice and 256 critical trials. On half of the related and unrelated trials, the cue appeared at a point halfway between the onset of the object name and completion of its auditory presentation. On the remaining trials, the cue was presented immediately after the object name had been spoken. Response time was measured from cue onset to the moment of contact with the response apparatus.

*Results and discussion*

Subjects responded correctly on 99.4% of the trials when prompted to report the object name that had been presented, indicating that they had attended to the auditory input. In post-session debriefing, subjects sometimes indicated that they had noticed that for some of the object names the cued hand action seemed relevant to the object, but that this relation did not hold consistently and therefore they made no strategic use of this information. Across the 20 subjects, errors occurred when making hand actions only 21 times out of over 5000 critical trials (0.4%) and were too infrequent to allow a meaningful analysis. These trials were excluded from the analysis of response times. In addition, any response time exceeding 1200 ms was classified as an outlier and was not included in the analyses we report. This cut-off was established so that no more than 0.5% of the observations would be excluded (Ulrich & Miller, 1994). Application of the response time limit resulted in the exclusion of 0.4% of correct responses.

Mean correct response time was computed for each subject as a function of probe point, action type, and prime relatedness. The means computed across subjects are shown in Fig. 2. The error bars in the figure are 95% within-subject confidence intervals (Loftus & Masson, 1994) based on the MSE for each pair of related and unrelated conditions. Inspection of the pattern of means and confidence intervals indicates that response times were shorter when the probe occurred at the end of the object name rather than at the midpoint and that there was a priming effect, but only for functional hand actions. An analysis of variance (ANOVA) with type I error rate set at .05 and probe point, action type, and prime as repeated measures factors yielded significance tests consistent with these conclusions. All three main effects were significant: response times were shorter when the cue appeared at the end of the object name, $F(1,19) = 56.58$, $MSE = 2211$, when sub-

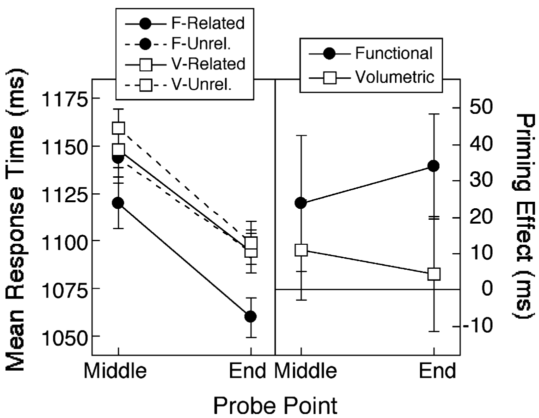


**Fig. 2.** Mean response time (left panel) for functional (F) and volumetric (V) hand actions in Experiment 1 as a function of probe point (middle or end of object name) and prime (related, unrelated) and mean priming effect as a function of action type and probe point (right panel). Error bars for condition means represent 95% between-subject confidence intervals (Loftus & Masson, 1994; Masson & Loftus, 2003) based on the MSE computed for each pair of related–unrelated conditions and error bars for priming effects are standard 95% confidence intervals.

jects were cued to make a functional rather than a volumetric hand action, $F(1,19) = 6.61$, MSE = 2649, and when the cued action was related to the object serving as the prime, $F(1,19) = 28.94$, MSE = 463. The only other significant effect was the interaction between action type and prime, $F(1,19) = 7.58$, MSE = 596. A simple effects analysis indicated that there was significant priming for functional actions (29 ms), $F(1,19) = 28.08$, but not for volumetric actions (8 ms), $F(1,19) = 1.98$.

The results of this experiment clearly indicate that when the names of manipulable objects were presented as auditory primes, action representations corresponding to the functional grasps associated with those objects were evoked, whereas representations of volumetric hand actions were not elicited. With this benchmark pattern of priming in place, we turn to an examination of the influence of sentence context, particularly verb type, on the evocation of hand-action representations.

## Experiment 2

In Experiment 2, we assessed the influence of two types of verbs describing non-manual engagement with manipulable objects. One set of verbs referred to actions that implied attention being directed toward an object (e.g., looked at), and the other set specified an interaction between an actor's foot and an object (e.g., kicked aside). As explained above, we anticipated that the latter type of interaction, by virtue of one's lack of experience in performing such actions, would evoke volumetric hand actions as a means of providing information about what the interaction might feel like and what consequences it might generate. Priming of volumetric hand actions by sentence contexts of this type would form a striking contrast to the lack of such priming when object names alone were presented in Experiment 1. Two versions of Experiment 2 were conducted. In the object-flanking version, we presented hand cues at two different probe points: (1) immediately after the post-verb preposition and before the article that preceded mention of the manipulable object and (2) immediately after the object name had been spoken. The early probe point was included to determine whether priming might be seen when preparation of a hand action was initiated slightly in advance of an object name. The late probe point corresponds to the end probe point in Experiment 1. The second version of Experiment 2, the object-middle version, was a replication of the object-flanking version in which the hand cues were presented at the midpoint of the enunciation of the object name (corresponding to the middle probe point in Experiment 1).

### Methods

#### *Subjects*

Eighty-four undergraduate psychology students drawn from the same pool as in Experiment 1 took part in the experiment. Forty-eight of the subjects were tested in the object-flanking version and 36 were tested in the object-middle version.

#### *Materials*

The hand actions, hand cues, and manipulable object names from Experiment 1 were used in Experiment 2. In addition, 288 sentences spoken in a male voice were digitally recorded. Each sentence conformed to the structure: The [adjective] [agent] [verb] the [object] (e.g., The young scientist looked at the stapler). Six different human agents (e.g., lawyer, professor, accountant) and various adjectives were used across the sentences. Two categories of verbs were used. For half the sentences, a verb was used that implied an agent was attending to an object. The following four verbs, which we call these attention verbs, were used equally often in these sentences: approached, glanced towards, looked at, and moved near. For the other half of the sentences, four verbs depicting non-manual, physical interaction (kicked aside, slipped on, stepped on, and walked over) were used equally often. We refer to these as interaction verbs. Within each set of 144 sentences representing a verb condition, 12 of the 18 manipulable object names served as the sentence object 6 times each. Among these 12 object names, six were related to one of the critical functional hand actions (e.g., water pistol: trigger) and six were related to one of the critical volumetric hand actions (e.g., thimble: horizontal pinch). The remaining six object names were related to one functional and to one volumetric hand action in our set (e.g., calculator: poke and horizontal grasp), and these object names appeared in 12 sentences each.

#### *Design*

In both versions of Experiment 2, each subject was tested with a pseudo-randomly selected set of 128 critical sentences taken from each verb condition (attention, interaction), for a total of 256 sentences. Selection was arranged so that, within each verb condition, each object name associated with either a functional or a volumetric action in our set of actions was used with about equal frequency across the experiment and each object name associated with both a functional and a volumetric action was used about equally often (and twice as frequently as object names in the former group). Within each verb condition, half of the sentences were paired with a functional action cue and half were paired with a volumetric action cue. These pairings were random with the constraints that each action was used equally often and that on half of the trials the action be paired with a sentence mentioning a related object name (e.g., drill paired with the trigger hand action), and with an unrelated object name on the other half of the trials. In selecting unrelated action cues, care was taken always to pair a cued action with an object name that was related to another hand action of the same kind. For example, the related action for eraser was horizontal pinch, a volumetric grasp. For unrelated sentence-action pairs involving sentences with eraser as the object name, any one of the other three volumetric actions could be used as the cued action. See Table 1 for additional examples. For the six object names that had both a related functional and a related volumetric action (e.g., stapler: palm and horizontal grasp), their sentences could be paired with any of the eight actions, with the constraint that half of the pairings resulted in a related name-action pair and half produced an unrelated pair.

*Procedure*

Subjects were tested using the same apparatus, general procedure, and instructions as in Experiment 1. After 40 trials of training on the eight hand actions, subjects were presented 16 practice and 256 critical trials in which they listened to a sentence and were cued to make a hand action. To ensure that subjects attended to the sentence content, they were presented with a probe question on 25% of the trials. These questions tested various aspects of the sentences (e.g., Who looked at the service bell? What did the teacher kick aside?) to encourage subjects to attend to all sentence components. Subjects were instructed to listen to each sentence and to make a hand response as soon as a visual cue appeared. In the object-flanking version, on half of the related and unrelated trials the cue appeared immediately after the post-verb preposition had been spoken and on the other half of the trials, the cue appeared immediately after the object noun had been spoken. In the object-middle version, all cues were presented as soon as the midpoint of the spoken object name was reached. As in Experiment 1, response time was measured from cue onset to the moment of contact with the response apparatus, and subjects responded orally to the probe questions.

*Results and discussion*

In the object-flanking version, trials on which the hand cue was presented in advance of the object name failed to generate reliable priming effects, so the results we report for that experiment are confined to those trials on which the cue occurred after the target object was spoken. Debriefing comments by subjects regarding the relation between hand actions and objects mentioned in the sentences were similar to those described in Experiment 1. Performance on the probe questions indicated that in both versions of the experiment, subjects attended to the content of the sentences; mean correct responding to the probe questions was 96.1% in the object-flanking version and 96.4% in the object-middle version. Trials were classified as spoiled if the response time was less than 200 ms. Only two trials in the object-flanking version and 14 trials in the object-middle version were excluded from our analyses by this criterion. As in Experiment 1, errors were very rare in both versions of Experiment 2 (0.5% and 0.7% in the object-flanking and object-middle versions, respectively), so no analysis of errors is reported. Correct response latencies longer than 2280 ms in the object-flanking version and longer than 2300 ms in the object-middle version were classified as outliers. These cut-offs excluded slightly less than 0.5% of critical trials in each case.

Mean response time as a function of verb type (attention, interaction), action type (functional, volumetric), and prime relatedness (related, unrelated) are shown in Fig. 3, separately for the two versions of Experiment 2. The pattern of means and 95% within-subject confidence intervals suggests that priming of hand actions was modulated by the type of verb contained in the context sentence.

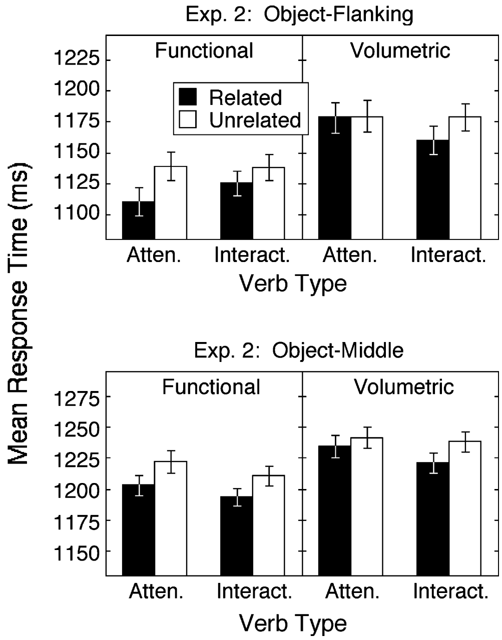


**Fig. 3.** Mean response time for functional and volumetric hand actions in the object-flanking and object-middle versions of Experiment 2 as a function of verb type (attentional, interaction) and prime (related, unrelated). Data from only the end of object probe point are shown for the object-flanking version. Error bars represent 95% within-subjects confidence intervals based on the MSE computed for each pair of related–unrelated conditions.

In particular, priming effects were apparent in both versions of the experiment for functional actions when attention verbs were used. In addition, both versions showed priming of volumetric actions when sentence contexts contained interaction verbs, but not for sentences that presented attention verbs. Priming of functional actions by sentences that included interaction verbs was somewhat equivocal in that it appeared in the object-middle version but was less robust in the object-flanking version.

To obtain a clear and statistically powerful analysis of priming effects, the data from the two versions of Experiment 2 were combined and entered into an ANOVA with type I error rate set at .05. The combined means for the two versions of Experiment 2 are shown in the upper panel of Fig. 4. The lower panel of this figure shows the magnitude of the priming effect for each combination of action type and verb type. The ANOVA yielded main effects of action type (functional actions were executed faster than volumetric actions), $F(1,83) = 40.00$, MSE = 5970, and prime relatedness, $F(1,83) = 35.10$, MSE = 1176. The only other significant effect was the three-way interaction, $F(1,83) = 4.72$, MSE = 1315. The interaction indicated that the influence of verb type had different effects on priming of functional and volumetric actions, as can be seen in the lower panel of Fig. 4. Specifically, changing from attention to interaction verbs reduced priming for functional actions (25 ms vs. 15 ms), but increased priming for volumetric actions (4 ms vs. 18 ms).

Pairwise comparisons indicated that priming was significant for functional actions in both verb conditions and

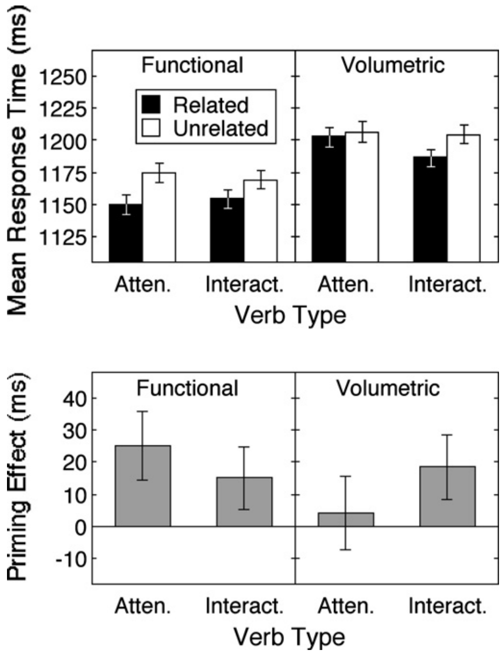


**Fig. 4.** Mean response time (upper panel) and mean priming effect (lower panel) as a function of action type and verb type combined across the two versions of Experiment 2. Error bars for condition means represent 95% between-subject confidence intervals based on the MSE computed for each pair of related–unrelated conditions and error bars for priming effects are standard 95% confidence intervals.

for volumetric actions only in the interaction verb condition ($F$s > 9), but not significant for volumetric actions in the attention verb condition ($F$ < 1). To check the consistency of the three significant priming effects across the four different verbs used in each verb condition, we computed mean response time in the related and unrelated-prime conditions for each set of trials associated with a particular verb. All four attention verbs yielded a shorter response time mean in the related-prime condition when functional actions were considered. Three of the four interaction verbs produced a priming effect for functional actions and three of those four verbs showed a priming effect for volumetric actions. Thus, the priming effects were reasonably consistent across the different verbs in each category.

It may be surprising to observe that priming effects as small as 15 ms are reliable in our procedure, given that mean response times are on the order of 1100–1200 ms. Typically, response times that long would require effects at least double the size of some of our highly reliable differences. Note, however, that a substantial part of the response time in our task is taken up with the transport of the response hand through space from a resting position to grasp a response element. These movements require approximately 400–500 ms and are usually conducted very smoothly and consistently, contributing very little variance to our performance measure.

To summarize the results of Experiment 2, we observed that sentences that have manipulable object names as their referents evoke representations of hand actions despite the fact that these sentences do not refer to manual interactions. The absence of reliable priming effects in the object-flanking version of Experiment 2 when hand action cues were presented just prior to the beginning of an object name suggests that substantial preparation of the hand action had been made before any meaningful processing of the object name occurred. Because our priming effects are defined by the relation between the cued hand action and the object, information leading up to but not including the object name (e.g., the verb) is not capable, by itself, of generating measurable priming effects in our task.

The overall pattern of results across the experiments reported here makes it unlikely that subjects formed and acted on overt expectations that a particular action would be cued in conjunction with a specific object. There are three aspects of the data that weigh strongly against strategic factors as the basis for the motor priming effects we report. First, in Experiment 1, priming occurred even when the hand cue was presented well before the word was fully articulated and the size of the priming effect in that case was no different than the priming effect found when the hand cue was presented at the end of the word. If subjects had been generating expectancies, we should have seen an increase in priming with the longer cue delay (e.g., Neely, 1977). We also note that in Experiment 2, where priming of volumetric actions was found with interaction verbs but not with attention verbs, mean response time for the unrelated-prime condition was nearly identical in the two verb conditions (1204 and 1206 ms, respectively). Had the priming of volumetric actions found with interaction verbs been the result of conscious expectancies, we should have observed not only a speed up in the related-prime condition, but also a slowing of responses in the unrelated condition (Neely, 1977), but no such effect materialized. More strikingly, if subjects deliberately anticipated functional hand actions in response to object names, either presented in isolation or in one type of sentence context, then it is difficult to explain why priming of volumetric actions suddenly emerges in another verb context, especially when that context has nothing to do with manipulating or using an object. Clearly, a more reasonable explanation of the pattern of priming effects is that they reflect the natural dynamics of sentence comprehension interacting with motor-based representations.

## General discussion

We have examined the evocation of hand actions by sentences that refer to manipulable objects. The verbs we used denoted either attention/movement toward an object, or a non-manual interaction with it. Our interest is in the kind of motor representations that may play a role in mental simulation as part of the construction of sentence meaning. On an account that assumes that motor simulation is confined to a literal depiction of the events described in the sentence, we should not observe activation of hand representations for either type of verb. The fact that we show clear evidence for evocation of both functional and volumetric hand actions to words like calculator and stapler is relevant to the question of how simulations are constructed over time and what exactly is being simulated.

*Functional actions*

In the introduction, we raised the possibility that attention verbs would activate action representations pertaining to object function but not representations based on the volumetric properties of objects. Our results confirm that functional representations alone are elicited dynamically as part of comprehending these sentences. This outcome makes sense if no information about general shape, weight, and rigidity is computed unless demanded by the context. We have argued that spatial reference requires only schematic representations of objects without volumetric details (Jackendoff & Landau, 1992). The evidence from single words showing priming of only functional actions, at least in the auditory domain (Experiment 1), invites a similar conclusion.

If part of the core meaning of an object is its intended function (e.g., Bloom, 1996; Malt & Johnson, 1992), the question remains as to whether hand actions for implementing that function are also part of the core meaning or are automatically derived once an abstract representation of function has been processed. For example, the meaning of calculator may depend on knowing that the device is used for arithmetic computation, but does understanding this word depend also on knowing the actions for operating it? Our evidence indicates that attending to the meaning of a word referring to a manipulable object automatically evokes knowledge associated with functional hand actions (see also Masson et al., 2008). Even sentences that describe actions such as kicking or stepping on, implying no intention to use an object or even interact with it manually, invoke functional representations. We infer from these findings that evocation of hand actions associated with object use (sometimes referred to as manipulation knowledge; Buxbaum & Saffran, 2002) are an inevitable consequence of accessing word meaning. Further work will be required to clarify the causal role of functional hand actions in semantic tasks.

*Volumetric actions*

By contrast, verbs denoting interactions like stepping on or kicking aside evoke volumetric hand actions, even though no such representations are evoked either by auditory words in isolation or by sentences that include verbs such as looking at and moving toward. This result raises important questions about sensory-motor representations constructed as part of sentence comprehension. A depictive simulation of The lawyer kicked aside the calculator presumably requires that the motor system represents the act of kicking with the leg. This assumption provides the rationale for demonstrations that specific motor regions are activated when subjects hear sentences or verbs describing actions with the hand, foot, or mouth (Hauk et al., 2004; Tettamanti et al., 2005). Sentences that convey routine interactions with objects (e.g., kicking a ball) may evoke representation of actions commensurate with prior experience in a manner that suggests depictive mental simulation. The deeper question is whether the same mechanism underlies the representation of action in less familiar contexts.

*Derivation of hand actions*

The evidence we have obtained suggests that motor simulation of actions may not necessarily stand in a direct relationship to the underlying meaning of a sentence. Rather, motor simulation draws on whatever experience best captures the relevant properties of the object referred to in relation to the act described. We may not know exactly what it feels like to kick a calculator, but we can rely on sensory-motor representations based on what it felt like in the past to lift and move such an object by hand. Thus, in certain situations, the simulation provides a kind of analogy for the listener using whatever prior motor representations are available (Barsalou, 1999; Zwaan, 2004).

This view of mental simulation implies that novel object affordances may not be directly available to the perceiver (cf. Gibson, 1979), but must be derived from known instances that can be used to draw inferences about properties of the desired object. Kaschak and Glenberg (2000) claim that understanding language includes the derivation of affordances for an object referenced in a sentence. Among the affordances of crutch they list, for example, an aid for walking, striking something, and pushing something through a long narrow crevice. They further propose that ease of derivation of one or another affordance depends on the actor's goals. We assume that on this account, deriving an unusual function is based directly on the appearance of the object in combination with knowledge about how one's body can interact with it. Thus, the affordance of a long and narrow object, for example, is immediately available from the word crutch if an actor's intention is to push a small object through a narrow opening. We distinguish between the "proper function" of an object (Millikan, 1984) and its derived functions. The proper function depends on the purpose for which the object has been designed, and we argue that the representations of hand actions immediately evoked by the name of the object in a sentence will be those associated with the proper function. A derived function would be constructed via a process of analogy to other objects more directly suited to the goal of the user. For example, the sentence John rolled the dough with the spray can requires a derived function using the shape of the spray can with a different goal in mind than the one for which it was originally designed. The listener understands this function indirectly, by analogy between the shape of the can and that of a rolling pin, the object whose proper function is indeed that of flattening dough. If this account is correct, then spray can should evoke the habitual functional and volumetric actions associated with a spray can even in a sentence about rolling dough. We predict that the action of pressing a palm on a cylinder would evolve only after the conventional actions have been evoked.

Returning to the sentences used in our experiment that described non-manual interactions with manipulable objects, such as The lawyer kicked aside the calculator, our analysis suggests an interesting possibility that could be evaluated using functional imaging methodology. Recall that Tettamanti et al. (2005) found that sentences describing leg actions (e.g., kicking a ball) activated areas in the motor system corresponding to the control of leg movement. There was less activation of the hand area for such sentences, but the reverse was true for sentences describing manual interactions with objects (e.g., grasping a knife). We found substantial priming associated with hand representations to sentences describing kicking or stepping on a manipulable object. Sentences of this type, referring to an atypical but plausible interaction, were of course not examined by Tettamanti et al. If our claim is correct that representations of volumetric hand actions are the means by which subjects derive information about how the foot interacts with such objects, then a functional imaging experiment should reveal at least as much activation of the hand as the foot area to a sentence like The lawyer stepped on the calculator.

## References


Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–660.

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition, 60*, 1–29.

Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition, 32*, 863–873.

Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science, 13*, 185–189.

Bub, D. N., & Masson, M. E. J. (2006). Gestural knowledge evoked by objects as part of conceptual representations. *Aphasiology, 20*, 1112–1124.

Bub, D. N., Masson, M. E. J., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition, 106*, 27–58.

Buxbaum, L. J., & Saffran, E. M. (2002). Knowledge of object manipulation and object function: Dissociations in apraxic and nonapraxic subjects. *Brain and Language, 82*, 179–199.

Daprati, E., & Sirigu, A. (2006). How we interact with objects: Learning from brain lesions. *Trends in Cognitive Science, 10*, 265–270.

Fagg, A. H., & Arbib, M. A. (1998). Modeling parietal–premotor interactions in primate control of grasping. *Neural Networks, 11*, 1277–1303.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*, 455–479.

Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., & Gangitano, M. (2000). Language and motor control. *Experimental Brain Research, 133*, 468–490.

Gibson, J. J. (1979). *The ecological approach to visual perception*. New York: Houghton Mifflin.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20*, 1–55.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558–565.

Glover, S., Rosenbaum, D. A., Graham, J., & Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research, 154*, 103–108.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). Perceiving the world and grasping it. A neurological dissociation. *Nature, 349*, 154–156.

Greif, M. L., Kemler Nelson, D. G., Keil, F. C., & Gutierrez, F. (2006). What do children want to know about animals and artifacts? *Psychological Science, 17*, 455–459.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301–307.

Jackendoff, R., & Landau, B. (1992). Spatial language and spatial cognition. In R. Jackendoff (Ed.), *Languages of the mind: Essays on mental representation* (pp. 99–124). Cambridge, MA: MIT Press.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language, 43*, 508–529.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95*, 163–182.

Klatzky, R. L., Pellegrino, J. W., McCloskey, B. P., & Doherty, S. (1989). Can you squeeze a tomato? The role of motor representations in semantic sensibility judgments. *Journal of Memory and Language, 28*, 56–77.

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.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490.

MacWhinney, B. (2005). The emergence of grammar from perspective taking. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 198–223). Cambridge, England: Cambridge University Press.

Malt, B. C., & Johnson, E. C. (1992). Do artifact concepts have cores? *Journal of Memory and Language, 31*, 195–217.

Masson, M. E. J., Bub, D. N., & Newton-Taylor, M. (2008). Language-based access to gestural components of conceptual knowledge. *Quarterly Journal of Experimental Psychology, 61*, 869–882.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*, 203–220.

Millikan, R. (1984). *Language and other abstract objects*. Cambridge, MA: MIT Press.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226–254.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science, 14*, 119–124.

Sakata, H., Taira, M., Kusunoki, M., Murata, A., & Tanaka, Y. (1997). The TINS lecture—The parietal association cortex in depth perception and visual control of hand action. *Trends in Neurosciences, 20*, 350–357.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*, 489–537.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645–660.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience, 17*, 273–281.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta Psychologica, 116*, 185–203.

Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General, 123*, 34–80.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*, 625–636.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.). *The psychology of learning and motivation* (Vol. 44, pp. 35–62). New York: Elsevier.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*, 1–11.