Cognitive representations of functional interactions with objects

ROBERTA L. KLATZKY
University of California, Santa Barbara, California

JAMES PELLEGRINO
Vanderbilt University, Nashville, Tennessee

BRIAN P. McCLOSKEY
Nike, Inc., Beaverton, Oregon

and

SUSAN J. LEDERMAN
Queen’s University, Kingston, Ontario, Canada

Two studies addressed people’s knowledge about the movements underlying functional interactions with objects, when the interactions were described by simple verbal labels expressing environmental goals. In Experiment 1, subjects rated each action with respect to six dimensions: which portion of the limb moved, distance moved, forcefulness, effectors involved, size of the contact surface, and resemblance to grasp. Ratings were systematic and fell on two distinct underlying factors related to limb movement and effector (usually the hand) configuration. In Experiment 2, subjects sorted a subset of the actions by similarity of movement. Clustering and multidimensional scaling solutions indicated that the six initial dimensions contributed to similarity judgments, along with additional parameters. The results support the existence of cognitively accessible, but still relatively specific, representations of functional actions, with potential implications for motor and memory performance.

Functional interactions with objects require a variety of movement patterns. We lift boxes, pound nails, staple papers, crumble cookies, and squeeze lemons in our daily lives. The purpose of this research was to investigate how people cognitively represent such actions, and it specifically addresses whether movement-related parameters (i.e., distinguishing dimensions) of interactions with objects can be retrieved from memory and function in processes like similarity judgments. This topic is important, because representations of action at cognitively accessible levels are thought to underlie a number of phenomena related to motor performance, as described below.

For the present purposes, a functional object interaction is a motor act that purposively changes the state (particularly the physical state) of an object (animate or inanimate) in the environment.¹ These movements constitute an interesting domain for the study of action representations, because environmental constraints tend to make them relatively stereotyped, promoting a stable encoding in memory. In addition, they can generally be denoted by simple verbal expressions for the action and object involved. Given that the lexicon tends to partition the world into ecologically important categories, the nature of the linguistic terms we have available for designating actions may tell us something about their most salient, cognitively discriminated features (Miller, 1972).

At the start of the present research, we identified a number of parameters that seemed likely to be incorporated into a cognitive representation of functional action—one that is memory-based and retrievable from verbal cues. The studies were intended to verify the relevance of those parameters, which are described below, and to discern others that might be used to distinguish between verbally presented actions. Further, we investigated the relative importance or salience of such distinctions.

As was noted above, cognitive representations of action have been assumed to serve a number of functions. For example, they may constitute high levels in a hierarchical motor program that are activated in initial stages of motor control (see Keele, 1986, for a review). They may be used in motor learning. For example, Schmidt’s (1975, 1988) schema theory assumes that knowledge of environmental consequences of a movement is used to adjust a representation of its parameters. They may be used to evaluate potential actions, enabling an individual to select the best among candidate movement sequences or to detect potential errors in advance. Cognitive levels of

¹ We thank Kristina Schamberger for her assistance with data collection. This paper profited considerably from interactions with participants at an ISAS (International School for Advanced Studies, Trieste, Italy) conference on mentally represented actions, directed by Mark Jeannerod. Address correspondence to R. L. Klatzky, Department of Psychology, University of California, Santa Barbara, CA 93106-9660.
movement representation also appear to be important for mental practice—that is, to enhance performance by rehearsing movements without actually producing muscular effects (Feltz & Landers, 1983; Richardson, 1967). They may also play a role in mediating verbal memory performance: A body of research on subject-performed tasks suggests that enactment of simple actions (e.g., bouncing a ball) leads to the encoding of a highly specific motor representation into memory—one that facilitates verbal memory but is also relatively inaccessible to nonmotoric cues (e.g., Engelkamp & Zimmer, 1984; Helstrup, 1989; see also Engelkamp & Cohen, 1991, for a review).

Although the foregoing purposes have been suggested, relatively little research has directly addressed the nature of cognitive movement parameters. Various systems for differentiating hand shapes have been presented (e.g., Cutkosky & Howe, 1990; Iberall, Bingham, & Arbib, 1986; Kroemer, 1986; Malek, 1981; Napier, 1956). However, those systems are generally motivated by biomechanical, physiological, or functional concerns. In contrast, our focus is on features of movement that are distinguished in everyday language and might function in cognitively governed contexts.

The present study is an extension of previous research that focused on shaping of the hand during functional actions. It classified hand shapes in terms of two binary parameters: size of the hand surface contacting the object (large or small) and whether the posture was prehensile (i.e., grasping) or nonprehensile. There are distinct lexical entries for each of the resulting four classes: clench (large and prehensile), pinch (small and prehensile), palm (large and nonprehensile), and poke (small and nonprehensile).

There is evidence that knowledge about interactions with objects incorporates such hand-shape distinctions. People can reliably report which hand shape they use in interacting with a given object and, conversely, when a hand shape is designated, they show substantial agreement on objects and functions that would elicit it (Klatzky, McCloskey, Doherty, Pellegrino, & Smith, 1987). Other evidence comes from a task in which subjects were asked to indicate whether or not verbally described interactions were “sensible” (Klatzky, Pellegrino, McCloskey, & Doherty, 1989). For example, it is sensible for a person to “crumple a newspaper,” but not to “climb a grape.” The time to make a positive decision was speeded by about 50 msec when subjects were given an immediately preceding cue as to the appropriate shape of the hand for the described action (e.g., a cue indicating a full-handed grasp, prior to a judgment about “crumple a newspaper”). These results suggest that evaluation of the action’s sensibility referred to a representation of the hand—the priming of which presumably facilitated the sensibility judgment.

Although these findings suggest a cognitively accessible level of knowledge about the movements performed during interactions with objects, additional results indicate that such knowledge is closely associated with the motor system. For one thing, the effectiveness of the handshape cues (e.g., for a poking shape, the symbol!) was found to depend on the nature of initial training as to their meanings, which preceded the sensibility judgments. The cues were effective when the training task required responding to each cue with overt movement (e.g., extending the index finger when the symbol appeared), but not when subjects were trained to respond with a verbal label (e.g., saying “poke”). Second, the priming effect was found to be disrupted when trials calling for sensibility judgments were randomly intermixed with trials calling for a manual tapping response, and subjects did not know until after the appearance of the prime which task would be required (McCloskey, Klatzky, & Pellegrino, 1992). Apparently, planning for the tapping task interfered with the representation invoked by the hand-shape cue. However, a comparable task, in which vocalization replaced tapping, produced no such interference.

Others have shown that cognitively accessible motor representations can penetrate the motor system to rather deep levels, in the absence of overt movement initiation. In a study by Gandevia and Rothwell (1987), subjects were taught to focus selectively on either of two muscles in the hand without moving the hand or recruiting motor units. At some time during the focus period, the motor cortex was weakly stimulated through the scalp in order to send diffuse activation to the hand. The cortical stimulation was found to produce a measurable response from motor units only in the muscle that was focused upon. Apparently, the conscious focus on a particular muscle produced activation of central nervous system motor neurons specific to that muscle, increasing its excitability relative to the other. This activation, combined with the stimulation of the scalp, yielded selective responses at the periphery.

The present studies were intended to extend our understanding of cognitive movement parameters by considering two general aspects of interactions with objects: how the limb is moved and how the end-effector contacts the object. (We define the end-effector as the body part contacting the object, and the limb as the appendage supplying and carrying the end-effector, together with its joints; most frequently these are the hand and the arm, respectively.) The literature on visually guided reaching has made a related distinction between arm transport and hand configuration (Jeannerod, 1981, 1986). Transport is generally described in terms of kinematic variables (related to position over time), whereas hand configuration is described in terms of geometric ones (e.g., the size of the aperture between fingers and thumb in grasping). Dynamic variables such as force at contact are usually minimized in simple reaching, because the arm markedly decelerates prior to contact, but they become important when describing object interactions beyond the reaching stage.

We will now consider a number of kinematic, geometric, and dynamic parameters that might be incorporated into knowledge about functional object interactions. Two variables concerning the movement of the limb are (1) which
limb segments are in motion during the action (i.e., whether movement is around a proximal or distal joint), and (2) the distance through which the end-effector moves. These will generally be highly correlated, because there is a preference for movement from more proximal joints as the required displacement of the end-effector increases (Rosenbaum, Slotta, Vaughan, & Plamondon, 1991).

As was noted above, force is important in achieving successful object interactions. Force is used to accelerate a mass, whether it is the end-effector alone (e.g., patting a back or trampling grass) or the end-effector and an object (e.g., hammering a nail or threading a needle). Force is obviously closely related to limb movement. Because limbs rotate around joints, we are often concerned with rotational dynamics and with torque, the rotational analogue of linear force. In the simple case of a point mass at the end of a lever arm of length $R$ (i.e., the distance from the mass to the point of rotation), with angular acceleration $a$, the torque $= mR^2a$, where $mR^2$ is the moment of inertia. Thus, the torque increases with acceleration of the limb, the mass of the limb and effector (which, although not a point mass, is directly related to torque for a given shape and axis of rotation), and the distance between the effector and the joint around which rotation occurs.

Types of object contact can be distinguished by which effectors are involved. For example, actions may involve the hand or the foot, and we may use one or both hands. Other geometric distinctions can be made in the way the effector interacts with the object, as exemplified by the hand-shape parameters described above—size of contact surface andprehension. These aspects of hand-shape map directly onto object geometry: When function is not an overriding factor, the tendency to select a prehensile hand shape increases with the object’s projection, and the tendency to select a shape with a large surface area increases with the size of the object’s projecting surface (Klatzky et al., 1987).

A number of considerations suggest that an important feature of a functional action is whether it involves a full-handed prehensile hand shape, which we term the clench, and others have called the grasp. Full-handed grasping is phylogenetically and ontogenetically primitive, appearing in lower evolutionary species than does grasping with a lesser number of fingers (MacNeilage, 1990). Even newborn infants can be observed to make rudimentary full-handed grasping motions in response to objects (Trevathan, 1984; von Hofsten, 1984). There is also some evidence that prehension (with full or partial hand) can be mediated by neural structures that are dissociated from conscious perceptual experience and hence may be relatively primitive. Goodale, Milner, Jakobson, and Carey (1991) reported that a subject with severe agnosia for visual form, subsequent to brain damage in the occipital region, was nonetheless able to adjust the aperture size of the hand to variations in object dimensions. In functional encounters with objects, the full-handed grasp is fundamental because many such actions begin with grossly picking up and orienting the object. However, the importance of grasping appears to extend to object-specific functional actions as well; we found that when subjects were asked to generate objects associated with the four hand shapes described above, the clench shape produced the most responses (Klatzky et al., 1987).

The present studies directly address whether features such as those just described are incorporated into the cognitive representation of functional object interactions. In Experiment 1, subjects rated a large number of such actions, denoted by simple verb phrases, with respect to the dimensions just described—which portion of the limb moved, the distance moved in space, the amount of force used, the effectors involved, the size of the hand surface contacting the object, and the similarity of the hand shape to a grasp. The substantial agreement on these ratings and reasonable correlational patterns indicated that the subjects were able to evaluate actions along the given dimensions. In Experiment 2, we investigated whether these same dimensions were reflected in free sorting of actions into similar groups. A similarity measure derived from sorting was used in clustering and multidimensional scaling (MDS) algorithms, in order to recover the dimensions used in representing functional actions and to determine their importance.

**EXPERIMENT 1**

**Dimension Ratings**

**Method**

**Subjects.** The subjects were 15 undergraduates at the University of California, who participated as partial fulfillment of an assignment.

**Stimuli.** Initially, we exhaustively examined verbs in Webster’s Collegiate Dictionary, with the intention of including those that (1) were transitive; (2) if polysemous, had a strong verbal meaning; and (3) were associated with functional interactions with objects. We also wished to sample verbs in order to vary the nature of the action on parameters such as those considered above. We could not, however, identify a very substantial number of actions that might be imposed on an object using the feet. Although our original intention was to use verbs alone, it soon became clear that knowledge of the object of action is necessary to specify the kinds of variables we were considering. For example, the verb stir implies very different forces when applied to coffee, paint, or mortar. Ultimately, a set of 92 verb phrases was constructed. Most explicitly identified a noun object, as shown in Appendix A.

**Rating scales.** A set of six scales was constructed, which assessed the parameters described above. Three (2, 4, and 6, below) pertained primarily to the limb, describing its movement and resulting contact force. The other three pertained to the configuration of the end-effector, describing which effectors were used and how they contacted the object. The scales used wording that was intended to be uncomplicated and therefore understandable to unsophisticated respondents. Hence, the word limb was used loosely to refer to the limb and end-effector, with response categories and examples used for clarification. All six scales were described on the cover page of a questionnaire and each was prefaced by a label (in the form of a question) that was used to identify it on subsequent pages. The label introducing each scale below will hereafter be used to denote it in this paper; the label given to the subject for that scale follows in parentheses. In slightly abbreviated form, the scales were as follows:

1. **Effector size** ("Which limbs?"). Which limbs of the body are involved in the action? Answer to indicate one hand; one foot; two
hands; two feet. (This was scored as 1-4, respectively, to reflect increasing effector size.)

2. *Amount of limb* ("How much limb?"). How much of the limb or limbs are in motion during the action? For example, if one hand is used, is there movement just of the fingers? The wrist? The elbow? The shoulder? Example given: Padding a canoe uses the whole arm; playing a trumpet confines the motion to the fingers. Answer with 1–5 scale, where 1 = motion is primarily in outermost parts of limb; 5 = substantial motion throughout limb.

3. *Amount of surface contact area* ("How much surface area?"). In most actions, the hand or foot makes contact with some object, which may involve a small amount of the surface of the hand or foot, a large amount, or something in between. How much of the hand/foot surface contacts the object? Answer with 1–5 scale, where 1 = little surface; 5 = full surface of hand or foot.

4. *Distance moved* ("How much movement?"). How much does the hand or foot move through space as the action is performed? Answer with 1–5 scale, where 1 = little or no motion; 5 = as much movement as the body will allow.

5. *Resemblance to grasp* ("Grasping?"). If the hand was used, does it grasp the object or operate without grasping? Example given: Squeezing a tomato involves grasping; picking one’s nose does not. Answer "grasp" [scored 1] or "no grasp" [scored 0].


**Procedure.** The rating scales were used by all the subjects in the order given, in order to facilitate comprehension (e.g., the concept of a limb was introduced before asking questions about its movement). Each page after the cover sheet describing the scales showed a single scale label at the top, below which were listed all 92 phrases. The subjects were to rate each of the phrases with respect to that scale, returning to the cover page if needed for clarification. Decimal numbers were allowed for numerical scales. The subjects could also mark any actions with which they were unfamiliar. Although we did not explicitly forbid the subjects to mime actions either here or in Experiment 2, we cannot rule out covert or small-scale actions. However, gross miming was generally not observed and, of course, the subjects did not have tools or target objects at their disposal. Hence, we assume that the subjects’ responses reflect cognitive evaluation with minimal, if any, motor cuing.

**Results.** The subjects seemed to have little difficulty interpreting and using the rating scales for most items. For the 1–5 scales, standard errors of the mean for the individual actions were usually below .3 (often well below), and with the scales having nominal responses (effector size and grasp), agreement was reached by 13 or more of the 15 subjects on about two thirds of the actions.

The correlations among scales were all significant, as shown in Table 1. Included are correlations with an additional variable that was subjectively coded by two experimenters: prehension, defined as the extent to which the hand flexed, regardless of whether another object was enclosed or not (a 1–3 scale, indicating increasing flexion). We added this variable because it was clear that the subjects in Experiment 1 had used the grasp scale as if closure on another object (cf. closing the empty hand on itself) was required in order to term the action a grasp.

As expected, the correlations were very high among the limb-related variables of force, amount of limb involved in the action, and distance moved through space. As indicated above, a greater amount of limb corresponds to a longer lever arm, which produces greater torque. The fact that the extent of the limb involved in an action was highly related to the amount of movement in space is consistent with the findings of Rosenbaum et al. (1991); these two scales produced the highest correlation (.93). Lower correlations were observed among variables related to the end-effector (i.e., area, effector size, and grasp). Further, the relatively low correlations relating the limb variables to the effector variables suggest that these are cognitively treated, to some extent, as independent. However, the amount of surface area in contact with the object did correlate substantially with the limb variables.

A principal-components factor analysis with orthogonal rotation was performed on the data for the 88 actions involving the hand (thus contributing a value to the prehension variable). Table 2 shows the rotated factor loadings. The first factor, accounting for 48% of the variance, received loadings on the arm variables of limb, distance, and force, and also on one hand-related variable, the surface area contacting the object. The second factor, accounting for 29% of the variance, had high loadings for effector size (which here reflects the number of hands used in the action), grasp, and prehension. Thus, nearly 80% of the variance was accounted for by factors that seem to represent arm movement/force and hand configuration.

In conclusion, people are able to evaluate a number of dimensions related to functional actions with objects when they are verbally interrogated about them. Correlations among those variables are generally consistent with principles of biomechanics and physics. We now turn to the possibility that these same dimensions will underlie people’s judgments of similarities among the movements used in functional actions with objects. If these parameters de-

---

### Table 1

<table>
<thead>
<tr>
<th>Item</th>
<th>Limb</th>
<th>Distance</th>
<th>Force</th>
<th>Effector</th>
<th>Area</th>
<th>Grasp</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distance</strong></td>
<td>.93†</td>
<td>.88‡</td>
<td>.31†</td>
<td>.35‡</td>
<td>.63‡</td>
<td>.37‡</td>
</tr>
<tr>
<td><strong>Force</strong></td>
<td>.87†</td>
<td>.34‡</td>
<td>.63‡</td>
<td>.18</td>
<td>.37‡</td>
<td>.69‡</td>
</tr>
<tr>
<td><strong>Effector</strong></td>
<td>.53‡</td>
<td>.64‡</td>
<td>.37‡</td>
<td>.41‡</td>
<td>.06</td>
<td>.08</td>
</tr>
<tr>
<td><strong>Area</strong></td>
<td>.26*</td>
<td>.37‡</td>
<td>.35‡</td>
<td>.37‡</td>
<td>.06</td>
<td>.08</td>
</tr>
<tr>
<td><strong>Grasp</strong></td>
<td>.05</td>
<td>.06</td>
<td>.18</td>
<td>.06</td>
<td>.08</td>
<td>.08</td>
</tr>
</tbody>
</table>

*p < .05, †p < .01, ‡p < .001.

---

### Table 2

<table>
<thead>
<tr>
<th>Scales</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>.958*</td>
<td>.081</td>
</tr>
<tr>
<td>Limb</td>
<td>.944*</td>
<td>.108</td>
</tr>
<tr>
<td>Force</td>
<td>.915*</td>
<td>.219</td>
</tr>
<tr>
<td>Area</td>
<td>.758*</td>
<td>.207</td>
</tr>
<tr>
<td>Effector</td>
<td>.302</td>
<td>.603*</td>
</tr>
<tr>
<td>Grasp</td>
<td>.247</td>
<td>.856*</td>
</tr>
<tr>
<td>Prehension</td>
<td>-.077</td>
<td>.911*</td>
</tr>
</tbody>
</table>

Note—Variance explained by Factor 1 = 48%; variance explained by Factor 2 = 29%. *Higher loading.
scribe important components of such actions, they should be recovered from the judgments of linguistic descriptions. Similarity judgments also allow us to ask about the relative importance or salience of such dimensions.

**EXPERIMENT 2**

**Similarity Judgments**

**Method**

**Subjects.** Twenty-three subjects from the same pool as Experiment 1 participated.

**Stimuli.** A subset of 50 verb phrases was selected from the 92 scaled in Experiment 1, as indicated in Appendix A. Initially, we eliminated a few phrases that had produced relatively high standard errors of the mean on some variable; for example, the subjects tended to disagree on whether ‘‘crush a coke can’’ involved the hand or the foot. We also eliminated the only two phrases in which there was strong agreement that the foot was used (‘‘tamp mare grass’’ and ‘‘kick a can’’), because there were few actions with the foot. To select the remaining phrases, we initially converted the mean for each action on each of the six variables to the nearest integer value. We then examined the frequency with which each combination of mean values occurred over the 92 items. We chose items with the goals of representing each dimension to the fullest extent possible, eliminating redundant combinations, and sampling as equally as possible from the available combinations of levels of force, limb, and distance moved. Essentially, we attempted to maximize the spread of items within a six-dimensional space corresponding to the scales used in Experiment 1. In the ultimate pool, only one combination of integer-valued means was represented by more than one item (hold and pinch, which we included to increase the number of items having low forces, limb, and distance). Sampling did not, of course, eliminate the natural correlations between scales that were observed in Experiment 1, as can be seen from Table 3, which shows correlations among the variables of Experiment 1 for the 50 actions that were retained in Experiment 2. Appendix B gives the mean value of each variable for each action used in Experiment 2 (indexed by the verb).

**Procedure.** The 50 verb phrases were written on individual file cards. The subjects were instructed to read the phrase on each card and then sort the cards into groups so that actions that were physically similar were sorted together. They were told that they should consider how a person performs the action, not its meaning. As an example, they were told that inserting a key in a door and putting out a cigarette in an ashtray might be thought of as similar, to some extent, on the basis of the action. After the initial sorting, the subjects were asked to give each group a title or descriptive phrase. They then combined their original groups into secondary groups that were similar, with the goal of reducing the total number of groups to half without restricting the number of original groups that were combined into each new one. Finally, they repeated the second step, forming third-level groups.

**Results**

The sorting data were converted to similarities by tallying, for each pair of verb phrases, the number of subjects who placed the two together. This was done for each of the three levels of sorting (recall that the subjects formed initial groups, then combined them twice successively, forming three levels). Two similarity measures were ultimately produced: similarity defined by the first level of sorting, and the summed similarity over the three levels. (The summed similarity produces a higher value, the earlier the level at which two verb phrases were sorted together.)

The similarities were submitted to complete-link cluster and MDS analyses. We focus on the cluster analysis for similarities summed over all three levels of sorting, because it can take advantage of the broader similarity scale that results. The MDS analyses are presented for the initial level of sorting.

Figure 1 shows the cluster analysis for the three-level sorting. The filled circles mark clusters where both members of a division have a similarity of less than 25; below that point, distinctions become relatively fine-grained. Suggestive labels have been given to clusters with relatively straightforward interpretations. The topmost distinction appears to divide actions with substantial movement of the arm through space (on the right branch) and those (left branch) where action concentrates on the hand or fingers (any arm movement occurs under spatial constraints). The ‘‘arm’’ branch then splits into a more voluntary, usually repetitive, motions and those where the arm swings more freely. Most of the latter are single movements involving the shoulder joint; they break further into two groups (unlabeled) that might be differentiated by whether movement is away from the body (left branch) or not (right branch). The ‘‘hand’’ cluster appears to incorporate the four shape classes studied in our previous research (e.g., Klatsky et al., 1987). It breaks down first into full-handed grasping actions (clench) and other hand actions; the latter then further differentiates actions without prehension from those using a pinching hand shape. Although the similarity values then become greater than our criterion, the no-prehension branch can be seen to break into actions involving the poke and palm shapes. The clenched actions bifurcate further into two groups (unlabeled)—those using the hand only (e.g., squeeze) and those involving arm motion within a limited range (e.g., stir).

In short, a subjective interpretation of the cluster analysis indicates that the subjects were most sensitive to a distinction between actions of the arm (particularly those involving free arm movement) versus the hand. Further distinctions were made with respect to the direction of the arm and the configuration of the hand, particularly full-handed grasps versus other hand shapes. We note that the cluster analysis using the first level of sorting data was generally quite similar to that combining levels; perhaps the most notable exception was that rake and hoe

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Correlation Between Scales Over the Mean Values of the Items in Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limb</td>
<td>Distance</td>
</tr>
<tr>
<td>Distance</td>
<td>.94‡</td>
</tr>
<tr>
<td>Force</td>
<td>.87‡</td>
</tr>
<tr>
<td>Effector</td>
<td>.20</td>
</tr>
<tr>
<td>Area</td>
<td>.52‡</td>
</tr>
<tr>
<td>Grasp</td>
<td>.34‡</td>
</tr>
<tr>
<td>Prehension</td>
<td>.21</td>
</tr>
</tbody>
</table>

*p < .05. †p < .01. ‡p < .001.
Figure 1. Cluster solution to the sorting data of Experiment 2. Filled circles mark nodes having components with similarity less than 25. Node labels suggest the nature of component groups.

combined with dunk, bounce, and hammer somewhat later.

Figure 2 presents the two-dimensional MDS solution to the first-level similarities, minimizing Kruskal stress. (A two-dimensional solution to the three-level similarity data yielded results similar to those of data from the first level of sorting; the two sets of coordinates correlated at .90 for one dimension and .92 for the other. Note also that in the figure, the scales of the dimensions differ to accommodate writing out the labels.) After 30 iterations, the stress level was .219 (cf. .144 for the threedimensional solution and .98 for four dimensions). The horizontal dimension appears to represent arm variables (force, limb, and distance, which are highly correlated), with pitching a horseshoe and rubbing a shoulder at opposite extremes. The vertical dimension appears to be related to the hand. For example, the extremes differ in grasp force, with pointing at a star and pressing a doorbell at one end, and crumpling, choking, and grabbing at the other.

These impressions were confirmed with two stepwise multiple regressions, using the x and y coordinates of each item in the two-dimensional scaling solution as the two predicted variables and the mean values for the item on the rating scales of Experiment 1 as predictors. Table 4 shows the results of these regressions. The variables "amount of limb involved in the action" and "resemblance to grasp" entered the x-axis regression first and second, respectively. The y-axis entered these variables in the reverse order, following with a third variable, effector size (which, for these 50 items, translates into the number of hands used for the action). Thus, the x-axis was best predicted by an arm variable and the y-axis by one relating to hand configuration, paralleling the factoranalytic solution in Experiment 1.

Figure 2. Two-dimensional scaling solution to the sorting data of Experiment 2 (first-level sorting only). Note that the x-axis has been expanded to accommodate item labels.
The relationship between the scaling solutions in Experiment 2 and the underlying factors obtained in Experiment 1 was confirmed with regressions predicting an item’s value on each axis in the two-dimensional solution from its factor scores. The x-axis values were predicted primarily by scores on the first factor (standardized coefficient of .69, \( p < .01 \)), although the contribution of the second was also significant (standardized coefficient of .22, \( p < .05 \)); the resulting regression accounted for 53% of the variance. For the y-axis values, the second factor was the better predictor (standardized coefficient of -.82, \( p < .01 \); cf. a coefficient of .17 for the first factor, \( p < .05 \)), and the regression accounted for 69% of the variance. Thus, the two factor scores were able to capture a substantial proportion of the variance in the similarity judgments.

The higher dimensional solutions for similarities derived from the first level of sorting are described in Table 5, which depicts significant correlations between an item’s coordinate on each dimension and the various predicting scales. It can be seen that the first MDS dimension is dominated by arm variables, regardless of the dimensionality of the solution space. (Whereas hand contact area, which loaded on the arm/force factor in Experiment 1, correlates strongly with this dimension, it tends to correlate with other dimensions as well.) In the two-dimensional MDS solution, the second dimension is an undifferentiated “hand” dimension. However, successive dimensions break this, at least to some extent, into other components.

**Table 4**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Step Entered</th>
<th>Multiple ( R^2 ) at Step</th>
<th>Standardized Coefficient (In Final Equation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>x-Axis</strong></td>
</tr>
<tr>
<td>Limb</td>
<td>1</td>
<td>.545</td>
<td>.643</td>
</tr>
<tr>
<td>Grasp</td>
<td>2</td>
<td>.603</td>
<td>.260</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>y-Axis</strong></td>
</tr>
<tr>
<td>Grasp</td>
<td>1</td>
<td>.546</td>
<td>-.839</td>
</tr>
<tr>
<td>Limb</td>
<td>2</td>
<td>.753</td>
<td>.499</td>
</tr>
<tr>
<td>Effector</td>
<td>3</td>
<td>.787</td>
<td>-.202</td>
</tr>
</tbody>
</table>

**GENERAL DISCUSSION**

The present results represent an initial step in understanding cognitive representations of functional interactions with objects. They clearly show that when interactions are described by simple labels representing environmental goals, a number of dimensions distinguishing among the component movements are cognitively accessible. Moreover, these dimensions have predictable relations to one another. The relevant aspects of movement have been revealed in two ways—by direct ratings of parameters of functional interactions and by judgments of the actions’ physical similarity.

Several of our findings point to a fundamental distinction between the hand and arm as components of cognitively represented action. The dimensions we examined a priori loaded on two distinct factors; the first largely represented variables relating to the arm and the second was dominated by the geometry of the hand. Analyses of the similarity data reinforced the arm/hand distinction. There is a parallel to this distinction in the literature on manipulatory control, in which arm transport and hand shaping are theorized to be independent components that are centrally coordinated over time (Jeannerod, 1981, 1984, 1986).

The present sorting data also reflect a number of more refined dimensions as well, relating to temporal and spatial aspects of movement (e.g., repetitiveness and direction relative to the body). With respect to the hand, there appears to be a major cognitive distinction between the full-handed grasp (or clench) and all other hand shapes. The cluster analysis, in particular, indicated that a grasp/nongrasp distinction was more critical than one differentiating shapes on the basis of prehension (which would cluster pinch and clench together and contrast them with the open shapes of poke and palm). The introduction reviewed a number of findings suggesting that grasping

**Table 5**

<table>
<thead>
<tr>
<th>Coordinate</th>
<th>Arm Variables</th>
<th>Hand Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Limb</td>
<td>Distance</td>
</tr>
<tr>
<td><strong>Two-Dimensional</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td>.74</td>
<td>.73</td>
</tr>
<tr>
<td>y</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Three-Dimensional</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td>.73</td>
<td>.75</td>
</tr>
<tr>
<td>y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Four-Dimensional</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td>.80</td>
<td>.80</td>
</tr>
<tr>
<td>y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
is a fundamental manipulatory category. At a more fine-grained level, beyond the grasp/no-grasp distinction, we do find evidence for the importance of the four-way classification of hand shapes used in our previous research, which makes distinctions based onprehension and contact area.

Although the present sample of action phrases was of course limited, it was constructed to provide a reasonable, ecological representation of the domain of functional interactions with objects. Toward this end, we attempted to sample very broadly from action verbs in constructing the materials of Experiment 1, and the phrases retained for Experiment 2 were selected in order to use the full range of each a priori dimension and to avoid biasing any one region of the multidimensional space under consideration. Thus, the present identification and ordering of movement parameters would be expected to have substantial generality over this action domain.

An important aspect of the present analysis is that the actions were described by simple phrases. Such labels may be useful for capturing functional interactions with objects, because by their very nature these interactions begin with an abstract level of representation describing an environmental goal, which the verb phrase denotes. In the context of real actions, the goal presumably calls up a more specific plan for movement, which institutes a chain of commands that ultimately culminates at a muscular level. It is this chain of events from intention to movement that has been termed a motor program. At its higher levels, a program is thought to be general enough to represent families of movement, members of which differ with respect to specific parameter values. The importance of this level of representation is that it is concrete enough to be used in governing action, yet is potentially accessible to conscious analysis and modification. This would allow it to function in the variety of roles considered in the introduction, including error evaluation in advance of overt movement, correction using knowledge of results, strengthening and priming actions, and encoding motor information into memory.

The importance of how action is represented at relatively high levels is evident in certain neurological impairments. Marcel (1992) described patients with symptoms resembling those of ideomotor apraxia, who were able to perform actions in the course of a familiar function (e.g., lifting a glass of milk to drink it) but not when they were designated as purely motoric goals (e.g., lifting a cylinder of the same size and weight). Brouchon, Joanette, and Samson (1986) found that both normal and brain-lesioned subjects produced different error patterns to biomechanically equivalent actions, depending on how the goal was designated (e.g., touching a near target with the fingertip or pointing to indicate a far target).

The present results may prove useful for understanding phenomena assumed to rely on cognitive movement representations. For example, the tendency of subject-performed tasks to enhance verbal memory performance may reflect the distinctiveness of the motor representation that is encoded. The advantage for overt performance should then decline if the to-be-remembered items are similar on cognitively salient movement parameters, such as the tasks within a low-level cluster in Figure 1. Interference from similarity of the objects of actions (cf. the movements) has been demonstrated (Nilsson & Backman, 1991). An interesting question is whether motric similarity effects would increase if subjects were induced to consciously focus on the parameters of movement underlying an object interaction.

As a final point, we consider the relationship between cognitive knowledge about dimensions of movement, as revealed with the present paradigms, and the concept of motor imagery. Are the movement representations we have described something that might be called motor images? Although there is no consensus on the definition of a motor image, it seems reasonable to consider criteria that have been applied to visual imagery. Thus, one might require that motor images have the content of actions, that they function to some extent like motor commands or motor output, and that they share processing avenues with the motor system. The present results speak to the first of these criteria, by showing that individuals have knowledge about rather specific components of action. On this basis, there seems little reason to distinguish between what we have termed cognitively accessible knowledge about actions and motor images.

REFERENCES


NOTES

1. Such movements should be distinguished from gestures that accompany speech for the purposes of reinforcement or semantic expansion (see Ekman & Friesen, 1969; McNeill, 1985; Wiener, Devoe, Rubinow, & Geller, 1972). The present corpus includes two marginal cases, wave and point, which have common verbal labels and are used in conjunction with objects but are also communicative and do not involve physical contact.

2. For the same reason, we did not attempt to instruct the subjects in how to judge the relative contributions of two limbs/effectors involved in an action, when assessing parameters such as distance moved and surface in contact. This would rarely be problematic, because most cases involve either a single limb/effecter or two that act similarly.

APPENDIX A

<table>
<thead>
<tr>
<th>Verb Phrases Used in the Experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>bend a coat hanger</td>
</tr>
<tr>
<td>*press a doorbell</td>
</tr>
<tr>
<td>*pump water</td>
</tr>
<tr>
<td>*punch a stomach</td>
</tr>
<tr>
<td>pull on a handle</td>
</tr>
<tr>
<td>*push a drawer in</td>
</tr>
<tr>
<td>*rake leaves</td>
</tr>
<tr>
<td>rattle a baby rattle</td>
</tr>
<tr>
<td>reel in a fish</td>
</tr>
<tr>
<td>*rub a shoulder</td>
</tr>
<tr>
<td>saw a board</td>
</tr>
<tr>
<td>*scoop flour</td>
</tr>
<tr>
<td>scour a pot</td>
</tr>
<tr>
<td>scratch an itch</td>
</tr>
<tr>
<td>*screw with a screwdriver</td>
</tr>
<tr>
<td>scribble on paper</td>
</tr>
<tr>
<td>scrub a floor</td>
</tr>
<tr>
<td>*slam a door</td>
</tr>
<tr>
<td>slap a cheek</td>
</tr>
<tr>
<td>*slash clothing</td>
</tr>
<tr>
<td>*slice meat</td>
</tr>
<tr>
<td>*spike a football</td>
</tr>
<tr>
<td>splash water</td>
</tr>
<tr>
<td>*squeeze a toothpaste tube</td>
</tr>
<tr>
<td>stab a back</td>
</tr>
<tr>
<td>stamp with a rubber stamp</td>
</tr>
<tr>
<td>staple pages</td>
</tr>
<tr>
<td>*stir a soup</td>
</tr>
<tr>
<td>*stroke a cheek</td>
</tr>
<tr>
<td>*swat a bug</td>
</tr>
<tr>
<td>sweep a floor</td>
</tr>
<tr>
<td>*swing a tennis racket</td>
</tr>
<tr>
<td>*tap a shoulder</td>
</tr>
<tr>
<td>tear a piece of paper</td>
</tr>
<tr>
<td>*thread a needle</td>
</tr>
<tr>
<td>*throw a baseball</td>
</tr>
<tr>
<td>tickle a friend</td>
</tr>
<tr>
<td>*toss a beanbag</td>
</tr>
<tr>
<td>*touch wet paint</td>
</tr>
<tr>
<td>trampoline grass</td>
</tr>
<tr>
<td>tug on a rope</td>
</tr>
<tr>
<td>twist a jar top</td>
</tr>
<tr>
<td>*type a letter</td>
</tr>
<tr>
<td>*wave one’s hand</td>
</tr>
<tr>
<td>whip cream</td>
</tr>
<tr>
<td>*wring a towel</td>
</tr>
<tr>
<td>wind a watch</td>
</tr>
</tbody>
</table>

*Used in Experiment 2.

APPENDIX B

<table>
<thead>
<tr>
<th>Mean Value of Each Variable for Verbs Used in Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>limb</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>bounce</td>
</tr>
<tr>
<td>braid</td>
</tr>
<tr>
<td>choke</td>
</tr>
<tr>
<td>chop</td>
</tr>
<tr>
<td>clap</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>club</td>
</tr>
<tr>
<td>clutch</td>
</tr>
<tr>
<td>crank</td>
</tr>
<tr>
<td>crumple</td>
</tr>
<tr>
<td>dunk</td>
</tr>
<tr>
<td>fasten</td>
</tr>
<tr>
<td>fingerprint</td>
</tr>
<tr>
<td>grab</td>
</tr>
<tr>
<td>grasp</td>
</tr>
<tr>
<td>grate</td>
</tr>
<tr>
<td>hammer</td>
</tr>
<tr>
<td>hit</td>
</tr>
<tr>
<td>hoc</td>
</tr>
<tr>
<td>hold</td>
</tr>
<tr>
<td>knock</td>
</tr>
<tr>
<td>pet</td>
</tr>
<tr>
<td>pinch</td>
</tr>
<tr>
<td>pitch</td>
</tr>
<tr>
<td>pluck</td>
</tr>
<tr>
<td>point</td>
</tr>
<tr>
<td>poke</td>
</tr>
<tr>
<td>polish</td>
</tr>
<tr>
<td>press</td>
</tr>
<tr>
<td>pump</td>
</tr>
<tr>
<td>punch</td>
</tr>
<tr>
<td>push</td>
</tr>
<tr>
<td>rake</td>
</tr>
<tr>
<td>rub</td>
</tr>
<tr>
<td>screw</td>
</tr>
<tr>
<td>slam</td>
</tr>
<tr>
<td>slash</td>
</tr>
<tr>
<td>slice</td>
</tr>
<tr>
<td>spike</td>
</tr>
<tr>
<td>squeeze</td>
</tr>
<tr>
<td>stir</td>
</tr>
<tr>
<td>stroke</td>
</tr>
<tr>
<td>sweat</td>
</tr>
<tr>
<td>swing</td>
</tr>
<tr>
<td>tap</td>
</tr>
<tr>
<td>thread</td>
</tr>
<tr>
<td>throw</td>
</tr>
<tr>
<td>toss</td>
</tr>
<tr>
<td>touch</td>
</tr>
<tr>
<td>type</td>
</tr>
<tr>
<td>wave</td>
</tr>
</tbody>
</table>

Note—dist = distance; eff = effector; pre = prehension.

(Manuscript received February 3, 1992; revision accepted for publication October 9, 1992.)