

# Perceiving Imitatable Stimuli: Consequences of Isomorphism Between Input and Output

Margaret Wilson  
North Dakota State University

For more than a century, psychologists have been intrigued by the idea that mental representations of perceived human actions are closely connected with mental representations of performing those same actions. In this article, connections between input and output representations are considered in terms of the potential for imitation. A broad range of evidence suggests that, for imitable stimuli, input and output representations are isomorphic to one another, allowing mutual influence between perception and motoric planning that is rapid, effortless, and possibly obligatory. Thus, the cognitive consequences of imitability may underlie such diverse phenomena as phoneme perception; imitation in neonates; echoic memory; stimulus–response compatibility; conduction aphasia; maintenance rehearsal; and a variety of developmental and social activities such as language acquisition, social learning, empathy, and monitoring one's own behavior.

The ability to copy what we see others do is a ubiquitous feature of human activity. It occurs spontaneously in infants, surfaces frequently in humor and play, and is necessary for the all-important task of social learning. What sort of cognitive architecture gives rise to this talent for imitation?

There are two a priori reasons to expect that imitable stimuli may have a special status in the human cognitive system. The first is the paramount importance of conspecifics in the normal daily activity of a social species. The investment of special resources for processing information about conspecifics would not be extravagant, given its potential return in terms of survival value. Second, the fact that we as humans already have expertise about human bodies, derived from having bodies ourselves, provides a ready-made resource that could be co-opted for perception. The perception of human figures may be unlike the perception of bicycles or octopuses simply because extra resources already exist.

The idea has been advanced many times that there may be deep connections between the mental representation of posture and movement of one's own body, on the one hand, and perception of posture and movement of others' bodies, on the other (for reviews, see Prinz, 1987; Viviani, 1990). It has been suggested, for example, that perception of human actions activates a motoric or kinesthetic representation of performing those actions oneself (e.g., Dimberg, Thunberg, & Elmehed, 2000; Rizzolatti & Arbib, 1998); that certain voluntary actions that physically copy a perceived action are driven by direct, privileged connections from perception to production (e.g., McLeod & Posner, 1984); and that a single, amodal *body schema* may underlie the representation of one's own body and of the perceived bodies of others (e.g., Berlucchi & Aglioti, 1997; Reed & Farah, 1995).

Although recurrent, this type of proposal is considered speculative, if not outright controversial (e.g., Ohala, 1996), and even supporters may conclude that the evidence is insufficient (e.g., Decety & Grezes, 1999). Two obstacles have stood in the way of serious consideration and widespread acceptance of this idea. The first is the range of evidence considered. The relevant evidence is often considered only within a limited domain, such as language perception or the perception of curvilinear point movement. At the same time, the literature on this topic is often embedded in broader theoretical contexts concerning sensorimotor interactions, such as perception for action (e.g., Decety & Grezes, 1999; Viviani & Stucchi, 1992b) or motor activity in the service of gathering perceptual information (e.g., Scheerer, 1984; Viviani, 1990).

A second major obstacle has been disagreement over the exact theoretical commitments involved. As noted above, some have postulated a body schema, an amodal representation of the human body that can serve for both perception of others' bodies and movement of one's own body. Other authors have proposed that two body schemas may exist, one perceptual and one motoric, with a translation process between the two (see Prinz, 1987, for a review). Still others have proposed distinguishing between a *body schema* and a *body image*, the latter involving conscious access (Gallagher & Meltzoff, 1996). It has also been suggested that there are multiple levels of perceptual and motoric representation, the highest of which contains partial but not complete overlap between the two types of coding (Prinz, 1997). It is unclear, though, that the data sufficiently constrain theory to decide among these alternatives, nor is it clear that we must do so in order to make theoretical progress. Instead, the proposal that perception of imitable stimuli is tied to the representation of one's own body can be captured in the following definition and two claims.

*Definition:* There is a class of stimuli that are *imitatable*, in the sense that the human body can engage in activity in which its configuration and movement can be mapped onto the configuration and movement of the stimulus. Although these stimuli are usually human actions, imitability may extend to include cases where the mapping is not perfect (e.g., imitating the movement of

---

Margaret Wilson, Department of Psychology, North Dakota State University.

Correspondence concerning this article should be addressed to Margaret Wilson, who is now at the Department of Psychology, University of California, Santa Cruz, California 95064. Electronic mail may be sent to [margaret\\_wilson@nds.u.nodak.edu](mailto:margaret_wilson@nds.u.nodak.edu).

a dog) or where the mapping only applies to a limited set of properties of the stimulus (e.g., copying the path of a shooting star with one's finger). However, it does not extend to cases of creating artifacts that resemble the stimulus (e.g., drawing a house, carving a wooden model of a boat, copying a printed word).

*Claim 1:* Imitable stimuli are not represented solely in terms of their surface perceptual properties, which are presumably incommensurable with motoric representations. Instead, imitable stimuli are parsed and encoded in terms of *articulatory gestures*, which reflect knowledge of the biomechanics of the human body—how its parts are arranged and how it moves. These biomechanically structured representations are isomorphic in their content to representations of the structure and movement of one's own body.

*Claim 2:* This isomorphism facilitates the flow of information processing between perception and action, particularly when the perceived actions are highly familiar and practiced in the observer. Such facilitation has observable consequences, which may include perceptual events triggering the activation of motoric representations, motorically derived knowledge influencing perception, and cognitive activities that involve commerce between representations of perception and production taking place in a rapid and near-automatic fashion.

### Perception of Articulatory Gestures

#### *Recovering the Distal Stimulus*

In 1785, Thomas Reid defined perception as “sensation that carries with itself an inevitable reference to an external object” (Reid, 1785/1965, p. 172). It has since become a commonplace within psychology that, while our sense organs receive only *proximal* stimulus (patterns of light falling on the retina, sound waves striking the eardrum, pressure against the skin), the job of our various perceptual systems is to recover information about the *distal* stimulus, the external objects and events that give rise to the proximal stimulus.

The point is most obvious for visual perception—examples abound in which the visual system actively reconstructs the external world. The principle of shape constancy, for instance, enables us to see rectangular objects as rectangular, despite the distortions of the proximal stimulus that result from changes in viewing perspective. Other examples include size, color, and brightness constancy; depth perception based on monocular cues; Gestalt principles such as figure-ground segregation and grouping; and modal completion effects, such as illusory contours and apparent motion.

The same principle holds in the other sensory modalities. Haptic exploration of objects with the hands is not perceived as a bewildering play of pressure and touch across noncontiguous patches of skin, but is instead perceived as a coherent pattern indicating the shape of the object. The point is slightly less obvious in the case of audition, since auditory perception does not recover the shape of the stimulus's source in the same way that vision and touch do. Nevertheless, audition does provide rich information about the distal stimulus (is it hard or soft, light or heavy, stiff or flexible?) and what is happening to it (is it being struck, scraped, or squashed?). Audition, no less than vision, aims at recovering the distal stimulus (Cabe & Pittenger, 2000; Gaver, 1993; Kunkler-Peck & Turvey, 2000; Romanski, Tian, Fritz, Mishkin, Goldman-Rakic, & Rauschecker, 2000).

When people perceive human movement or human speech, as with other stimuli, we attempt to recover the distal stimulus. In this case, the distal stimulus consists of movements or postures whose characteristics (e.g., range of motion, temporal dynamics, structural relations of parts) are determined by the biomechanics of the human body. That is, the distal stimulus to be recovered in the case of human stimuli is arguably a set of articulatory gestures, defined in terms of human biomechanics, and therefore commensurable with the articulatory gestures of one's own body. The following sections examine the evidence for this claim.

#### *Language Perception*

The issue of whether we perceive articulatory gestures comes into sharpest relief when we consider speech perception. There is considerable evidence that the phonemes of speech are not perceived in terms of their physical acoustic patterns, but rather they are perceived in terms of the articulatory events—movements of the lips, tongue, and so on—that give rise to those phonemes (for reviews, see Fowler, 1986, 1994; Liberman, 1996; Liberman & Whalen, 2000; Miller, 1990).

One source of evidence is the impact of visual information on phoneme perception. Watching the mouth of a speaker can assist in speech comprehension, often so seamlessly that people are unaware they are lip-reading as well as listening. Evidence that this is an integral part of speech perception itself, and not a form of guesswork, comes in part from the McGurk effect. In this effect, visual perception of a mouth speaking a syllable that conflicts with a heard syllable can alter what syllable is heard (Massaro & Cohen, 1983; McGurk & MacDonald, 1976). For example, participants who hear the syllable *da* but see the syllable *ba* are likely to report *hearing ba* and are not aware of any conflict between the auditory and visual signals. In other words, what is perceived by the participant is the output of the perceptual system's “best guess” based on multiple sources of information. Thus, phonemes are not perceived solely in terms of acoustic patterns but are influenced by other sources of information about the articulatory gestures involved.

Further evidence for the perception of speech gestures is the phenomenon of *duplex perception*. Segments of a phoneme sound qualitatively different in isolation than when embedded in the context of the phoneme (e.g., Mann & Liberman, 1983; Whalen & Liberman, 1987). For example, when part of the early waveform of the syllable *da* or *ga* (specifically, the pitch glide of the third formant) is isolated, it sounds like a chirp that differs in quality from human speech. Yet when the pitch glide is presented to one ear and the remaining portion of the syllable is presented to the other ear, the two sounds are integrated so that the whole syllable is heard. Strikingly, the chirp sound vanishes. Further, when the intensity of the pitch glide is increased, participants report hearing the syllable *and* the chirp. That is, this sound fragment is interpreted by the auditory system in two radically different ways simultaneously—as an integrated part of the syllable, and as an isolated and qualitatively different nonhuman sound. These phenomenal shifts may reflect the fact that, in the context of a syllable, the segment forms part of the signal of a coherent articulatory gesture, whereas the segment in isolation could not be produced by human articulation. (A related phenomenon may be the shift in sound quality that occurs when a noise is interpreted as human or

not human—for example, a trumpet note heard as a human voice or a squeaky piece of machinery heard as a baby crying.)

Duplex perception has in fact been observed for nonhuman sounds as well. When the sound of a metal door slamming is filtered to create a high-frequency “excerpt” and a low- and medium-frequency “base,” the base sounds like a wooden door slamming (that is, the metallic quality is gone), while the excerpt sounds like something metallic being shaken or rattled. When the base is presented to one ear and the excerpt to the other, the shaking sound vanishes and participants report hearing the metal door. Further, when the intensity of the excerpt is increased, a duplex percept emerges just as with speech, and participants report hearing both the metal door and the shaking sound (Fowler & Rosenblum, 1990). Although this contradicts claims that duplex perception is due to a special phonetic module (e.g., Liberman & Mattingly, 1989), it fits well with the concept of recovering the distal stimulus. Of particular interest here, though, is that in the case of speech what is recovered appears to be a set of articulatory gestures, representing the biomechanical properties of the human mouth.

Perhaps supporting this interpretation of duplex perception is a recent report that human newborns (and cotton-top tamarin monkeys as well) can distinguish between two unfamiliar languages—for example, Dutch and Japanese—when they are played normally, but not when they are played backward (Ramus, Hauser, Miller, Morris, & Mehler, 2000). This suggests that the discrimination is based on properties of the signal that are only discernible when they can be interpreted as arising from vocal articulation.

A third source of evidence on the recovery of articulatory gestures is the effect of articulatory context on perceived phoneme categories (e.g., Mann & Repp, 1980). For example, when synthetic speech is used to create stimuli that vary between *sha* and *sa*, participants are more likely to perceive the *sh* sound than when the stimuli vary between *shu* and *su*. This shifting of phoneme category boundaries is needed to compensate for *coarticulation* (the altering of phoneme pronunciation caused by the articulatory demands of the immediately preceding and following phonemes). Because coarticulation from the *a* vowel tends to shift the pronunciation of *sh* in the direction of *s*, listeners must use a more lenient criterion for detecting *sh* in *sha* than in *shu*. In other words, phoneme categories cannot be defined in terms of strictly physical acoustic characteristics. Instead, they appear to be defined in terms of the sound produced by a particular articulatory gesture as shaped by the articulatory demands of neighboring phonemes.

The evidence summarized here has sometimes been taken to support a strong version of perceptual–motor interaction. According to the original formulation of the motor theory of speech perception, articulatory gestures are recovered by actually engaging speech production mechanisms. For the present purpose it is possible to remain neutral on this point and instead simply note that the mental representations generated during speech perception are commensurable with representations of speech production, because both are structured in terms of movement of the human mouth.

### *Immediate Memory*

Further evidence that speech is perceived in terms of articulatory gestures comes from immediate serial recall. Two classic

effects are relevant here: the recency effect and the suffix effect. The *recency effect* refers to the characteristic superior performance on the last few items of a list of spoken words to be recalled. The *suffix effect* refers to a drop in performance on the last few items (i.e., reduction of the recency effect) when the presented list is followed by a single irrelevant word or syllable, which the participant knows is not part of the material to be recalled. With immediate serial recall, both the recency effect and the suffix effect are greatly reduced or absent for nonspeech stimuli such as printed words. (This phenomenon has been called the *modality effect*, referring to a difference between auditory and visual stimuli, but as we shall see, this is a misnomer.) In early theories, this pattern of effects was attributed to a precategorical acoustic store (echoic memory). The last few items heard were believed to be retained briefly in a relatively unprocessed form, preserving their acoustic properties but highly vulnerable to disruption by further incoming acoustic material (see Surprenant, Pitt, & Crowder, 1993, for a review). However, later results challenged this interpretation.

First, it has been found that the recency and suffix effects occur for lip-read material, which is linguistic but not auditory (Campbell & Dodd, 1980; Greene & Crowder, 1984; Spoehr & Corin, 1978). Conversely, these effects are small or nonexistent for non-linguistic sounds (Foreit, 1976; Greene & Samuel, 1986; Roberts, 1986). This indicates that the critical factor is not sensory modality but something about the structure of the stimulus.

Additional evidence comes from the fact that the suffix effect can be instated or abolished without changing the stimulus properties at all, but simply by influencing how the participant interprets the suffix. In a study by Ayres, Jonides, Reitman, Egan, and Howard (1979), participants were presented with sequences of spoken words to be remembered, followed by an ambiguous suffix. One group of participants was told that the suffix was the sound of a trumpet; the other group was told that it was a human voice saying the syllable *wa*. Only the participants who believed they were hearing a syllable produced by a human vocal tract showed a suffix effect. In addition, Surprenant et al. (1993) found that the recency effect can also be instated or abolished depending on how the to-be-remembered items themselves are interpreted. Degraded speech was either described to participants as the numbers 1 through 4 and accompanied by other nondegraded digits or was described as the sounds “bubbles,” “owl,” “squeal,” and “whistle” and accompanied by other nondegraded sounds. The recency effect occurred in the former condition, but not the latter. Thus, as Crowder and Surprenant (1995) suggested, these effects may be “associated with the recovery of phonetic gestures” (p. 62). When participants interpret an acoustic pattern as coming from a trumpet or bubbles, there simply are no gestures to be recovered, and the signature effects of this kind of immediate memory do not occur.

### *Interpreting Point-Light Displays and Apparent Motion Displays*

Another approach to the recovery of articulatory gestures is to study the perception of impoverished displays that imply human motion. It has been found that a vivid and immediate percept of human motion can be obtained from an extremely minimal display of point lights attached to the joints of a moving human model

(e.g., Johansson, 1973, 1977). Even attributes such as gender can be perceived, and complex movements can be extracted from these displays (see Thornton, Pinto, & Shiffrar, 1998, for a review). Furthermore, the perception of these figures is not accomplished solely by low-level visual analysis, such as extracting the movement of pairs of points indicating rigid segments. Instead, high-level, global visual processing is implicated in the perception of point-light human figures (e.g., Thornton et al., 1998). These findings suggest that even very minimalist displays, whose surface, proximal form bears little resemblance to the human body, are interpreted in terms of structural relations and movements of human body parts.

In addition, point-light displays have been used to examine in a more fine-grained manner the nature of the information that can be extracted from the dynamic spatiotemporal properties of such displays. Viviani and Stucchi (1992a) conducted experiments on the perception of curvilinear movement trajectories of a single point light. They found that trajectories obeying the kinematics of biological motion (expressed by an equation in which speed varies with the degree of curvature) were perceived as having uniform speed, even though the variations in speed in fact exceeded 200% (see also Viviani, Baud-Bovy, & Redolfi, 1997; Viviani & Mounoud, 1990; Viviani & Stucchi, 1989). These findings suggest that curvilinear movements are perceived, not in terms of actual changes in velocity, but in terms of wholistic gestures, which appear smooth and uniform when executed as a biological system would articulate them.

In a related study, Kandel, Orliaguet, and Viviani (2000) showed participants excerpts of three-letter sequences being handwritten in cursive. Only the middle letter was shown, and the participants' task was to guess which of two letters would be the third letter, based on the movement they perceived in the writing of the middle letter. Accuracy was high when the trajectories obeyed the equation mentioned above and decreased as violations of that equation become greater. The authors concluded that "perceptual anticipation of human movements involves comparing the perceptual stimulus with an internal dynamic representation of the ongoing event" (p. 706).

A similar conclusion can be drawn from a study by Klima, Tzeng, Fok, Bellugi, and Corina (1996), in which participants were shown a point-light display of a finger tracing an invented Chinese character in the air. The task was to immediately reproduce the target figure on paper. Of particular interest for the current purpose is the fact that the target figure did not consist of a continuous trajectory but rather of discrete "strokes." However, the point light that was tracing the character remained lit continuously throughout the duration of the display. Thus, part of the problem set for these participants was to distinguish the strokes from the transitional movements between strokes, in order to recover the target character. The primary variable of interest for Klima et al. was a comparison between deaf and hearing participants, which showed that deaf participants not only were better at reproducing the characters in general but, more specifically, were better at extracting the strokes from the transitions. This enhanced performance was presumably due to experience with sign language, which likewise involves dynamic specification of strokes connected by transitions. However, a more basic lesson can be drawn from these data, which is the fact that the stroke versus transition

information was recoverable at all by human observers. This information was presumably carried by the dynamics of the display—that is, by the pattern of accelerations and decelerations that characterized the strokes versus the transitions. The ability to distinguish these two types of characteristic dynamic patterns, and to know which one constituted the task-relevant information, is presumably based on recognition of the dynamics involved in intentional or goal-directed biological movement. Once again, this suggests that the display is being processed in terms of the gestures that produced it.

Evidence for the recovery of articulatory gestures can also be seen in how the visual system "completes" a display containing partial information. In the phenomenon of apparent motion, two stimuli shown alternating rapidly at two distinct locations are perceived as a single object moving back and forth. Normally, apparent motion obeys a shortest-path principle. However, Shiffrar and colleagues (Shiffrar & Freyd, 1990, 1993; Shiffrar, Lichtey, & Chatterjee, 1997) have demonstrated that, with appropriate timing parameters, apparent motion for displays involving human figure stimuli obeys principles of biological motion rather than the shortest-path principle. For example, if two photographs are rapidly alternated, one showing a human figure with the hand in front of the torso, the other showing the hand behind the torso, participants tend to perceive the hand moving *around* the torso, rather than *through* it as would be required by the shortest-path principle. Similarly, if the two photographs show a human arm with bent elbow at the two extremes of its range of rotation, participants tend to perceive the arm rotating all the way through its range of rotation from the one extreme to the other, rather than taking the shortest (but biomechanically impossible) path between the two extremes.

Further evidence comes from an experiment by Kourtzi and Shiffrar (1999), in which participants were sequentially shown two depictions of a human figure, the difference between the two being rotation of the figure in the depth plane and also changed position of one limb. For example, Frame 1 might show a human figure facing 45° to the left and with a leg extended forward with straight knee. Frame 2 might show the figure facing 45° to the right and with the leg bent at the knee. Participants showed priming immediately afterward not only for the two figures they had seen but also for figures that represented a biomechanically possible midpoint between the two—for example, facing front and with the knee partially bent. This occurred both when there was apparent motion between the two prime figures and when there was not. Priming did not occur for biomechanically impossible midpoint figures. This implies that the two seen figures activated a mental representation of how the human body would move from the first posture to the second.

In short, stimuli generated by the human body appear, as a general principle, to result in perceptual representations that are structured in terms of the biomechanics of the human body. Unlike purely acoustic or visual information, this kind of representation is parsed into the same kinds of components and structural relationships that characterize the representation of one's own body. The two types of representations—others' bodies and one's own body—should therefore be commensurable and capable of being mapped isomorphically onto one another. If this is so, one should expect to find telltale consequences in a variety of cognitive tasks that involve representing the human body.

### Facilitated Information Flow

#### *Automatic Mapping of Perception to Response*

It is well-known within the field of psychology that certain types of stimuli and responses seem to pair naturally. Various called *stimulus–response compatibility* or *ideomotor compatibility*, this phenomenon may be driven by the extent to which the required response imitates the stimulus.

The point is perhaps most obvious in the domain of verbal repetition. *Shadowing* is a task in which participants hear an ongoing stream of words or sentences through headphones and are required to repeat the message word for word as they hear it. Although shadowing requires attentional focus on the source of the stimulus, it appears to be relatively automatic in the sense that it does not require an effortful translation process between the auditory input and the articulatory output. Shadowing can in fact be performed rapidly with extremely short lags by untrained participants (McLeod & Posner, 1984). The connection between verbal input and output is further highlighted by the phenomenon of *conduction aphasia*, in which verbal repetition can be selectively impaired, even when both speech perception and production remain intact. Patients with conduction aphasia can speak and can understand spoken language, but they have great difficulty when asked to repeat back words that are spoken to them. This suggests selective damage to a pathway whose function is to translate heard speech into corresponding produced speech.

McLeod and Posner (1984) proposed that these phenomena are the result of a “privileged loop from percept to act” (p. 55) that is distinct from more general information-processing systems that mediate between stimulus and response. They suggested that this privileged channel is not an endpoint on a continuum of compatibility, but rather a special pathway dedicated to the repetition of spoken words. They also suggested, however, that other, as yet undiscovered privileged loops exist. (One can imagine that repetition of sign language, for example, might also make use of such privileged connections.)

The apparently unique status of verbal repetition, though, may in fact rely on two properties of language that are not shared by most other stimuli: the high degree of familiarity of the stimulus elements (at both the phonemic and lexical levels), with the consequent automaticity of both perception and production, and the degree to which the stimulus can be imitated by the human body. Thus, what appears to be a sharp break between language repetition and other forms of stimulus–response mapping may in fact mask an underlying continuum. For example, repetition of nonlinguistic body movements or vocal sounds, although not as overpracticed as language, might exhibit a weaker form of privileged connection due to their imitability.

This conceptualization also opens the door to explaining the broader phenomenon of stimulus–response compatibility in the same context as verbal repetition. It has long been known that certain stimulus–response combinations afford more rapid, effortless processing than other combinations (for reviews, see Klapp, Porter-Graham, & Hoifjeld, 1991; Kornblum, Hasbroucq, & Osman, 1990; Prinz, 1997). For example, participants are faster to respond to a stimulus that appears on the left or right side of space by making a response with their left or right hand than they are to respond by saying the words *left* or *right*. Further, participants are

slower still to respond with the hand that conflicts with the location of the stimulus. Similar effects of “compatibility” (a concept that has usually appealed more to intuition than to a consistent definition) have been noted across a wide variety of stimuli and tasks. However, Kornblum et al. described these as “an atheoretical collection of seemingly unrelated phenomena” (p. 255) observed in widely different experimental paradigms addressing disparate theoretical issues. Kornblum et al. attempted to unify the phenomena of stimulus–response compatibility in a coherent framework by proposing that the critical factors are degree of *dimensional overlap* between a stimulus set and a response set and the *mapping* of individual stimuli and responses within those sets. To take a simple example, spatially varied stimuli, and responses that are also spatially varied, have dimensional overlap with one another and thus lend themselves to stimulus–response compatibility. However, the compatibility will also be determined by whether the relative positions of various stimuli are mapped onto corresponding or conflicting relative positions of the responses. This analysis is not restricted to spatial position but can apply to a variety of other measurable dimensions such as size, direction of motion, timing, pitch, or loudness.

Another way of phrasing this concept of dimensional overlap is that some stimulus–response combinations constitute imitation, at least on a restricted set of dimensions or properties. One can imitate the spatial layout of stimuli with the positions of one’s hands, the timing and rhythm of a repetitive tactile stimulus with the tapping of one’s foot, or the pitch of a stimulus with the pitch of one’s voice.

Furthermore, when stimulus–response compatibility is considered in terms of imitation, we can see that analysis in terms of number of shared dimensions may capture only part of the phenomenon. A more complex analysis may be required for imitating a curvilinear path movement, a configuration of limbs, or speech. For example, tracing a path with a finger or limb might need to be defined in terms of replicating the number of points of maximum curvature and discontinuities as well as the relative horizontal and vertical positions of these key points. For posture, an analysis of joint angles as acute, right, or obtuse, and where the plane defined by the angle of the joint lies relative to the *x*, *y*, and *z* planes, may be required. For speech, imitation requires reproducing an ordered sequence of phonemes. Stimulus–response compatibility may then be defined in terms of proportion of match on these kinds of elements. (A number of predictions proceed from this—for example, responding to a seen hand shape with the same hand shape rotated in space should be easier than responding with a different hand shape, and responding to a stimulus word with a word that rhymes should be easier than responding with a phonologically unrelated word.)

Thus, conceiving of stimulus–response compatibility in terms of imitation not only offers a new perspective on dimensional overlap, but also provides a way to broaden that perspective to include more complex kinds of compatibility. There may indeed be privileged connections, as McLeod and Posner (1984) suggested, offering a distinctive kind of processing advantage due to rapid and automatic translation between input and output; however, this privilege may be a matter of degree, varying as a function of the imitability of the stimulus.

### *Neurological Substrates of the Connection*

Stimuli that strongly suggest a particular response can generate activity in motor cortex, even if no overt response is made. For example, motor cortex becomes active when participants observe objects such as tools, which are strongly associated with specific activities (Grafton, Fadiga, Arbib, & Rizzolatti, 1997; Murata et al., 1997). This kind of highly automatized connection between a stimulus and a single overlearned response may underlie the phenomenon of utilization behavior in patients with frontal lobe damage, who cannot inhibit behaviors such as combing their hair with a comb or making eating motions with a fork (cf. Craighero, Fadiga, Umiltà, & Rizzolatti, 1996).

In contrast to these cases of single-use objects, observation of one's conspecifics is not usually associated with a fixed response. For a single observed behavior (e.g., a conspecific eating), cooperation may be appropriate in one set of circumstances, aggression in another, and polite inattention in a third. In any case, sheer imitation of the conspecific is only sometimes an appropriate response.

It is surprising and important, then, that motor activation can be observed which does not correspond to using or interacting with a stimulus but instead corresponds to imitating it. In monkeys, particular neurons dubbed *mirror neurons* fire both when the monkey performs a particular action and when the monkey observes a human or a monkey performing a similar action (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Further, these motor neurons are in an anatomically distinct region from those neurons that respond to stimuli in terms of appropriate uses. The function of the mirror neurons is apparently not to plan a response to observed movement but to mentally simulate observed movement, and the purpose of this simulation may be to encode and understand the behavior of conspecifics (cf. Miklósi, 1999; Rizzolatti & Arbib, 1998, 1999).

Similar kinds of activity may also occur in the motor cortex of humans. In one study, participants were given transcranial magnetic stimulation (TMS), a procedure in which a magnetic pulse is applied to the surface of the head over a particular region of cortex in order to observe the effects of transient disruption or stimulation of that area. In this case, the TMS was applied over the motor cortex, and motor evoked potentials were recorded from the muscles of the participants' hands (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995). The motor activity in the hands induced by the TMS increased when participants simultaneously watched an experimenter handling an object or tracing a shape in the air. Furthermore, the pattern of muscle activity matched that of actually performing the movements. Similarly, Hari et al. (1998) found that a rebound effect in precentral cortex that is known to be suppressed during object manipulation was also somewhat suppressed when participants observed object manipulation by another person.

Further, activation of premotor cortex has been found using positron-emission tomography (PET) imaging when participants are asked to observe human actions with the intent to imitate at a later time (Decety et al., 1997; Grezes, Costes, & Decety, 1998). This could, of course, reflect motor imagery due to planning movements for the future and may not be directly tied to the observation of the actions. However, even when participants are not instructed to prepare for imitation, some activation in motor

and premotor areas can still be observed, although exact locations are not always consistent (Decety et al., 1997; Grafton, Arbib, Fadiga, & Rizzolatti, 1996; Grezes, Costes, & Decety, 1998; Rizzolatti, Fadiga, Matelli et al., 1996). Furthermore, Iacoboni et al. (1999) reported functional magnetic resonance imaging data showing that left prefrontal cortex is more active during imitation of finger movements than during performance of finger movements in response to symbolic cues, suggesting that this area plays a particular role in the use of perceptual information to drive the movement task in an imitative fashion. Although some authors are understandably cautious in their interpretations of these various data (e.g., Decety & Grezes, 1999), these rapidly accumulating findings seem to suggest mutual involvement of perception and imitative behavior in a very specific neurophysiological sense.

### *Innateness of the Connection*

Further evidence suggests that the isomorphism between perception of human movement and production of corresponding movement is innate. Imitation of adult models' actions and vocalizations is clearly an important learning device for infants who are starting to acquire language and other skills. But how do infants know, particularly for actions such as facial expressions where they cannot see themselves, that their own motor commands and the resulting kinesthetic feedback correspond to the perceived behavior of the adult? Although attempts have been made to account for infant imitation in terms of reinforcement (Gewirtz & Stingle, 1968) or in terms of self-observation (Piaget, 1962), such accounts have trouble explaining facial imitation in neonates (e.g., Meltzoff & Moore, 1977, 1995). Imitation of simple facial actions, such as tongue protrusion and mouth opening, have been observed in infants only a few hours old. Infants will also imitate a nonreflex gesture, such as tongue protrusion to the side, and will make successive approximations, correcting their own mistakes. Furthermore, infants with a medical condition preventing tongue protrusion show signs of effort and frustration, indicating that the cognitive system is monitoring its success or failure at imitation. (See Meltzoff & Moore, 1995, for a review.)

The precise mechanisms of how this mapping is specified without the benefit of perceptual learning remain unclear. Meltzoff and Moore (1997) proposed a model of how such mapping may occur, based on functions such as *organ identification*, and *body babbling* (a kind of exploration of body movement possibilities analogous to linguistic babbling), which may form part of the infant's innate repertoire. Regardless of the details, though, connections between perceptual and motor representations of the human body appear to be innately specified. (For further discussion, see Shiffrar & Pinto, in press.)

It appears, then, that connections between perception and imitative motor commands are indeed privileged—neurophysiologically instantiated, innately specified, and highly automatized. This privileged status may make the act of imitation a relatively likely response in social contexts. It has been proposed, for example, that imitation results from priming of a mental representation of a particular type of behavior, defined at some level of abstraction (Byrne & Russon, 1998). According to this view, sensory input signaling the behavior of others increases the activation of mental representations of those behaviors, increasing the likelihood that the same behaviors will be performed. The consequences of the

privileged status of imitation, however, go well beyond ease and automaticity of imitation, as can be seen in the next section.

### Further Consequences of Imitability

#### *Motoric Influences on Perceptual Representations of Human Figures*

A variety of findings suggest that various perceptual judgments about human figures rely in part on the activation of one's own body representations. One such finding involves people's ability to judge the weight of objects being lifted or carried by another person. It has been found that participants can extract weight information from videotapes (Valenti & Costall, 1997), point-light displays (Runeson & Frykholm, 1983), or static photographs (Valenti & Costall, 1997) of objects being lifted or carried. Although perceptual learning cannot be ruled out as an explanation for these data, it is perhaps equally plausible that the perception of the muscular, postural, and movement cues involved in lifting and carrying recruit stored knowledge of how one's own body responds to lifting and carrying. This recruited knowledge may then assist in the perceptual judgment.

A related example is the recognition of one's own gait. Beardsworth and Buckner (1981) found that participants were better at identifying themselves from a point-light display of a figure walking than they were at recognizing friends and colleagues. It appears that participants were "recognizing" something that they were already familiar with, namely, how their own gait would look. This implies that the production of movement can generate a corresponding quasi-perceptual representation.

Of course it is also possible that the causality may run in the other direction. Rather than recognizing the "look" of their gait from previously generated perceptual representations, participants may be using the perceptual input to generate a motor representation of how that particular figure moves, and noting that, for one figure more than the others, the movement feels familiar. In either case, a perceptual learning account seems implausible, since gaits that are more visually familiar (the gaits of friends and colleagues) are not recognized as well as one's own gait, which is rarely seen.

More tellingly, Parsons and colleagues have shown that judgments of whether a viewed hand is a left or right hand involve the activation of somatosensory and motor representations (see Parsons, 1994, for a review). First, it has been shown that when participants perform mental rotation of their own bodies, reaction time is a function of the size of the angle through which the rotation must occur, just as in other mental rotation tasks. Further, when participants are asked to make a left-right judgment about a viewed human hand, reaction time is a function of the angle between the orientation of the stimulus and the orientation of the participant's own hand. Furthermore, in a related study, Gentilucci, Benuzzi, Bertolani, Daprati, and Gangitano (2000) asked participants to make left-right judgments on hands holding small, medium, or large spheres. They found that the greater the motoric precision required by the posture of the hand, the longer the reaction time. These findings strongly suggest that participants are performing mental transformations of their own body in order to map it onto the stimulus, and this mapping process is used to assist in the ostensibly visual judgments required by the tasks. This conclusion is strengthened by findings from PET imaging, which

has shown that during the left-right judgment task, somatosensory and motor areas of the brain are active (Parsons & Fox, 1998; Parsons et al., 1995).

There is also evidence that motoric knowledge can influence the *content* of perceptual representations. Consider again the biomechanical effects observed by Shiffrar and Freyd (1990, 1993) for apparent motion of the human body, wherein perceived apparent motion violates the shortest path principle in favor of physically possible paths such as the arm moving around the torso. Can these effects be due to perceptual learning—that is, sheer familiarity with the visual patterns that result from physically possible paths of movement? Crucially, Shiffrar et al. (1997) demonstrated that these effects are limited to biological stimuli. When the two alternated photographs consist of an inanimate object first in front of and then behind another object (e.g., a stick in front of and behind a chair), the shortest-path principle holds and participants perceive the one object passing through the other. The fact that the longer-path effects are restricted to human body motion implies that perceptual learning is not the cause. This suggests that, instead, biomechanical knowledge derived from one's own body may be responsible for the effect.

Further evidence supporting this conclusion comes from PET imaging. Stevens, Fonlupt, Shiffrar, and Decety (2000) presented human body apparent motion stimuli and asked participants to indicate on each trial whether they had perceived the biomechanically possible or impossible path of motion. On trials where participants reported the possible path, selective activation was found in primary motor cortex, relative to trials where participants reported the impossible path.

Recent evidence also indicates that the motoric expertise involved in being right- or left-handed can influence stored perceptual representations for stimuli that imply directionality. Several studies report that right-handers tend to recall familiar human figures (e.g., on road signs or coins) as facing left, whereas left-handers tend to recall the figures as facing right (Jones, 1990; Martin & Jones, 1998, 1999; Rubin & Kontis, 1983). The effect has also been observed for memory for the direction of the Hale-Bopp comet, indicating that the stimuli need not be human (Martin & Jones, 1999). The authors suggested that the handedness effect is mediated by motoric processes that would be used for drawing. It is possible, though, that the effect is mediated by some other form of motoric representation, such as a preferred direction for turning one's own body or the use of the dominant hand to gesture the direction that a person or object is moving. Given that these kinds of processes are almost certainly far more automatized than drawing, it seems plausible that motoric imitation, rather than motoric routines for producing drawings, plays a role in memory for visually perceived directionality.

#### *Off-Line Uses of Imitability*

The impact of imitability can also be seen in cognitive processes that take place "off-line," that is, in the absence of ongoing task-relevant input and output. A prime example is working memory. The process of "rehearsing" verbal material within working memory has been the focus of a considerable body of research over the past three decades. This is in part because *maintenance rehearsal*, a kind of covert repetition of linguistic stimuli, is an astonishingly natural and ubiquitous strategy for temporary stor-

age. Indeed, a large and ongoing body of research is feasible precisely because untrained participants can be counted on to use this strategy.

The historically standard model of verbal rehearsal since the mid-1970s has been the *phonological loop* proposed by Baddeley (Baddeley, 1986; Baddeley & Hitch, 1974). According to this model, storage of verbal material in working memory makes use of two components: a passive storage buffer, which holds materials in phonological or quasi-auditory form but which decays rapidly, and an active rehearsal process, which is based on articulatory or motor-planning processes and is used to repeatedly refresh the material in the buffer. In other words, participants are saying the words to themselves over and over, just as introspection suggests.

Although there have been a number of challenges to the Baddeley (1986; Baddeley & Hitch, 1974) model, some kind of sensorimotor-based model appears to be the best way to account for a wide range of data on working memory. This evidence includes effects of sound similarity, articulation time, competing articulatory activity, and other manipulations, as well as neuropsychological data implicating speech-processing regions of the brain such as Broca's and Wernicke's areas (see Wilson, 2001, for a review). One striking finding, which is difficult to explain in any other terms than sensorimotor rehearsal, is that children who consistently mispronounce phonemes (e.g., /w/ for /r/) make those phonemic confusions in working memory. They have no difficulty distinguishing spoken words such as *ring* and *wing*, but when asked to remember a list of words (including *ring*) and then answer by pointing to pictures, they tend to point to a picture of a wing, rather than a picture of a ring (Locke & Kutz, 1975). Apparently, the participant's own pronunciation abilities underlie the maintenance of words in working memory.

If we accept a sensorimotor model of working memory, notice what it implies. The model postulates a rapid cycling of information between perceptual and motoric forms of coding. The ease and automaticity of this cycling shows in the fact that it is not only an overwhelmingly preferred temporary memory strategy but a highly effective one as well, far superior to alternatives such as visualization. In addition, there is evidence that this form of rehearsal of verbal items, for set sizes that do not exceed the participant's memory span, requires relatively little oversight by attentional or "executive" functions (e.g., Hitch & Baddeley, 1976). This automaticity of translation between the two forms of coding suggests that isomorphism between perceptual and motoric representations confers processing advantages, even in the absence of overt perceptual and motoric activity.

Furthermore, the structure of this rehearsal process is not due to the unique properties of speech but instead is due to some broader principle. It has been demonstrated that deaf participants who use American Sign Language show the same patterns of sensorimotor effects but in a different sensory modality (Hanson, 1982; Poizner, Bellugi, & Tweney, 1981; Wilson, Bettger, Niculae, & Klima, 1997; Wilson & Emmorey, 1997a, 1997b, 1998a, 1998b; Wilson, Iverson, & Emmorey, 2001). These findings indicate, at a minimum, that sensorimotor rehearsal in working memory is not restricted to spoken language. Additional evidence suggests that it extends even beyond language to other forms of imitable action. Working memory for sequences of body postures and for sequences of spatial locations displays some of the same effects that characterize immediate memory for language, again suggesting the

presence of rehearsal (Awh, Jonides, & Reuter-Lorenz, 1998; Smyth, Pearson, & Pendleton, 1988; Smyth & Scholey, 1994a, 1994b). In contrast, however, nonimitable visual stimuli, such as nonnameable shapes or matrices with random cells filled, do not lend themselves to a rehearsal process of the kind described above.

Finally, this sensorimotor structure of working memory has uses beyond maintenance rehearsal. For example, a wide variety of auditory imagery tasks rely on motoric representations of speech output (Reisberg, Smith, Baxter, & Sonenshine, 1989; Smith, Reisberg, & Wilson, 1992; Smith, Wilson, & Reisberg, 1995). Thus, the special processing abilities conferred by imitability may contribute even to off-line cognitive tasks.

### Concluding Remarks

In our daily activities, a lively commerce between perception and production can be observed. Prinz (1987) gave a compelling example of unbidden muscle activity when watching a movie hero nearly fall off a precipice. Similar examples can be found in Lotze, writing in the 19th century, and in William James (both as cited in Prinz, 1987). (I have my own anecdote along these lines. During a recent bout of office cleaning, which found me on the floor passing computer cables back and forth behind table legs, I noted that my actions felt like those of a weaver at a loom, even though I have never done any weaving. Then I recalled that I had recently watched *Silas Marner*, whose title character is a weaver. I was *feeling* what the character's actions had *looked* like.) If interactions between representations of perceived and produced movement are indeed this ubiquitous, the potential implications for understanding our cognitive and social lives are profound.

Copying the behavior of conspecifics, on which all social learning rests, may depend on the relative automaticity of imitation. Without this kind of copying, fundamental developmental processes such as language acquisition and acculturation would not occur. Even the later and more deliberate learning of complex skills would be severely impaired. Although humans are certainly capable of learning from abstract explanation, this ability occurs relatively late in development, and even for adults learning by example remains far easier. Further, spontaneous copying can facilitate not only social learning but also social communication, as in the unconscious mimicking of facial expressions (Dimberg, Thunberg, & Elmehed, 2000).

A second function of the ability to imitate may be understanding the behavior of conspecifics. Even when there is no intention to imitate, being able to process the behavior of others in terms of one's own behavior may facilitate understanding others' intentions (cf. Miklósi, 1999; Rizzolatti & Arbib, 1998, 1999). Even such early and basic abilities as shared attention and interpreting the eye gaze of others—the absence of which, as in autism, has severe consequences (e.g., Baron-Cohen, 1995)—may depend on an ability to form empathetic representations of the actions of others.

Yet a further use may be to facilitate prediction. In general, perceivers try to predict the immediately upcoming behavior of the thing perceived, possibly by running a rapid simulation of the perceived event (e.g., Clark & Grush, 1999). This process may be aided by "constraints" provided by perceptual experience or evolutionary hard-wiring, causing these mental simulations to track the expected behavior of objects obeying physical laws (cf. Shepard, 1994). In the case of human or humanlike movement, per-

ceivers may have the option of using motoric coding to assist in these predictions (cf. Arbib & Rizzolatti, 1996). Just as in the perceptual case, motoric experience would constrain the predictions to follow certain principles, in this case biomechanical principles of how joints and muscles behave. Because motoric as well as perceptual constraints would then be brought to bear on the task, predictions might be expected to be more robust and accurate and to maintain their accuracy over a longer time frame. This could be particularly important for predicting the behavior of self-powered, jointed entities that appear to move in a non-Newtonian fashion, such as animals and human beings. Because the movement in these cases is so much more complex than that of objects obeying basic Newtonian laws of velocity and force, the need for additional predictive constraints, such as those that could be provided by motoric simulation, may be essential. (This may help to explain why people tend to anthropomorphize not only animals but also complex machinery.)

Another consequence of isomorphism between perception and production may be the ability to monitor one's own behavior—in effect, to see it through the eyes of others. This may be an important function for regulating the social appropriateness of one's behavior, including body language and facial expressions. In addition, it may underlie the ability of actors and comedians, as well as ordinary people telling stories and jokes, to know in a rapid, on-line fashion when they are performing well.

Finally, we can speculate on how perceptual–motor interactions may affect individuals who possess various forms of complex, motoric expertise. Mastery of complex physical skills such as ice skating, martial arts, or dance forms may have not only motoric consequences but perceptual ones as well. To a dancer, whose body knows how to perform the observed movements, a dance performance may literally look different than it does to a non-dancer. To return to a theme with which this article began, the proximal stimulus will be the same for both observers, but, because of the expert observer's capacity for imitation, the ability to robustly represent the distal stimulus will differ.

## References

- Arbib, M. A., & Rizzolatti, G. (1996). Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition*, 29, 393–424.
- Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 780–790.
- Ayres, T. J., Jonides, J., Reitman, J. S., Egan, J. C., & Howard, D. A. (1979). Differing suffix effects for the same physical suffix. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 315–321.
- Baddeley, A. (1986). *Working memory*. Oxford, England: Oxford University Press.
- Baddeley, A., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 647–667). Hillsdale, NJ: Erlbaum.
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.
- Beardsworth, T., & Buckner, T. (1981). The ability to recognize oneself from a video recording of one's movements without seeing one's body. *Bulletin of the Psychonomic Society*, 18, 19–22.
- Berlucchi, G., & Aglioti, S. (1997). The body in the brain: Neural bases of corporeal awareness. *Trends in Neurosciences*, 20, 560–564.
- Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. *Behavioral & Brain Sciences*, 21, 667–721.
- Cabe, P. A., & Pittenger, J. B. (2000). Human sensitivity to acoustic information from vessel filling. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 313–324.
- Campbell, R., & Dodd, B. (1980). Hearing by eye. *Quarterly Journal of Experimental Psychology*, 32, 85–99.
- Clark, A., & Grush, R. (1999). Towards a cognitive robotics. *Adaptive Behavior*, 7, 5–16.
- Craighero, L., Fadiga, L., Umiltà, C. A., & Rizzolatti, G. (1996). Evidence for visuomotor priming effect. *NeuroReport*, 8, 347–349.
- Crowder, R. G., & Surprenant, A. M. (1995). On the linguistic module in auditory memory. In B. de Gelder & J. Morais (Eds.), *Speech and reading: A comparative approach* (pp. 49–64). Hillsdale, NJ: Erlbaum.
- Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences*, 3, 172–178.
- Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., & Fazio, F. (1997). Brain activity during observation of actions: Influence of action content and subject's strategy. *Brain*, 120, 1763–1777.
- Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, 11, 86–89.
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91, 176–180.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action: A magnetic stimulation study. *Journal of Neurophysiology*, 73, 2608–2611.
- Foreit, K. G. (1976). Short-lived auditory memory for pitch. *Perception & Psychophysics*, 19, 268–370.
- Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realistic perspective. *Journal of Phonetics*, 14, 3–28.
- Fowler, C. A. (1994). Auditory “objects”: The role of motor activity in auditory perception and speech perception. In K. H. Pribram (Ed.), *Brain and self organization* (pp. 593–603). Hillsdale, NJ: Erlbaum.
- Fowler, C. A., & Rosenblum, L. D. (1990). Duplex perception: A comparison of monosyllables and slamming doors. *Journal of Experimental Psychology*, 16, 742–754.
- Gallagher, S., & Meltzoff, A. N. (1996). The earliest sense of self and others: Merleau-Ponty and recent developmental studies. *Philosophical Psychology*, 9, 211–233.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593–609.
- Gaver, W. W. (1993). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Psychology*, 5, 1–29.
- Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., & Gangitano, M. (2000). Recognising a hand by grasp. *Cognitive Brain Research*, 9, 125–135.
- Gewirtz, J. L., & Stingle, K. G. (1968). Learning of generalized imitation as the basis for identification. *Psychological Review*, 75, 374–397.
- Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography: II. Observation compared with imagination. *Experimental Brain Research*, 112, 103–111.
- Grafton, S. T., Fadiga, L., Arbib, M. A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *NeuroImage*, 6, 231–236.
- Greene, R. L., & Crowder, R. G. (1984). Modality and suffix effects in the absence of auditory stimulation. *Journal of Verbal Learning and Verbal Behavior*, 23, 371–382.
- Greene, R. L., & Samuel, A. G. (1986). Recency and suffix effects in serial recall of musical stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 517–524.

- Grezes, J., Costes, N., & Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A PET investigation. *Cognitive Neuropsychology*, *15*, 553–582.
- Hanson, V. L. (1982). Short-term recall by deaf signers: Phonetic coding in temporal order recall. *Memory & Cognition*, *18*, 604–610.
- Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 15061–15065.
- Hitch, G. J., & Baddeley, A. D. (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology*, *28*, 603–621.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999, December 24). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211.
- Johansson, G. (1977). Studies on visual perception of locomotion. *Perception*, *6*, 365–376.
- Jones, G. V. (1990). Misremembering a common object: When left is not right. *Memory & Cognition*, *18*, 174–182.
- Kandel, S., Orliaguet, J.-P., & Viviani, P. (2000). Perceptual anticipation in handwriting: The role of implicit motor competence. *Perception & Psychophysics*, *62*, 706–716.
- Klapp, S. T., Porter-Graham, K. A., & Hoifjeld, A. R. (1991). The relation of perception and motor action: Ideomotor compatibility and interference in divided attention. *Journal of Motor Behavior*, *23*, 155–162.
- Klima, E. S., Tzeng, O., Fok, A., Bellugi, U., & Corina, D. (1996). *From sign to script: Effects of linguistic experience on perceptual categorization* (Tech. Rep. No. INC-9604). San Diego: University of California, San Diego, Institute for Neural Computation.
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus–response compatibility—A model and taxonomy. *Psychological Review*, *2*, 253–270.
- Kourtzi, Z., & Shiffrar, M. (1999). Dynamic representations of human body movement. *Perception*, *28*, 49–62.
- Kunkler-Peck, A. J., & Turvey, M. T. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 279–294.
- Liberman, A. M. (1996). *Speech: A special code*. Cambridge, MA: MIT Press.
- Liberman, A. M., & Mattingly, I. G. (1989, January 27). A specialization for speech perception. *Science*, *243*, 489–494.
- Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, *4*, 187–196.
- Locke, J. L., & Kutz, K. J. (1975). Memory for speech and speech for memory. *Journal of Speech and Hearing Research*, *18*, 176–191.
- Mann, V. A., & Liberman, A. M. (1983). Some differences between phonetic and auditory modes of perception. *Cognition*, *14*, 211–235.
- Mann, V. A., & Repp, B. H. (1980). Influence of vocalic context on perception of the [f]–[s] distinction. *Perception & Psychophysics*, *28*, 213–228.
- Martin, M., & Jones, G. V. (1998). Generalizing everyday memory: Signs and handedness. *Memory & Cognition*, *26*, 193–200.
- Martin, M., & Jones, G. V. (1999). Hale–Bopp and handedness: Individual differences in memory for orientation. *Psychological Science*, *10*, 267–270.
- Massaro, D. W., & Cohen, M. M. (1983). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 753–771.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.
- McLeod, P., & Posner, M. I. (1984). Privileged loops from percept to act. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance* (Vol. 10, pp. 55–66). Hillsdale, NJ: Erlbaum.
- Meltzoff, A. N., & Moore, M. K. (1977, October 6). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75–78.
- Meltzoff, A. N., & Moore, M. K. (1995). Infants' understanding of people and things: From body imitation to folk psychology. In J. L. Bermudez, A. Marcel, & N. Eilan (Eds.), *The body and the self* (pp. 43–69). Cambridge, MA: MIT Press.
- Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting*, *6*, 179–192.
- Miklósi, A. (1999). From grasping to speech: Imitation might provide a missing link. *Trends in Neuroscience*, *22*, 151–152.
- Miller, J. L. (1990). Speech perception. In D. N. Osherson & H. Lasnik (Eds.), *An invitation to cognitive science: Language* (pp. 69–93). Cambridge, MA: MIT Press.
- Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raso, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (Area F5) of the monkey. *Journal of Neurophysiology*, *78*, 2226–2230.
- Ohala, J. (1996). Speech perception is hearing sounds, not tongues. *Journal of the Acoustical Society of America*, *99*, 1718–1725.
- Parsons, L. M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 709–730.
- Parson, L. M., & Fox, P. T. (1998). The neural basis of implicit movements used in recognising hand shape. *Cognitive Neuropsychology*, *15*, 583–615.
- Parsons, L. M., Fox, P. T., Downs, J. H., Glass, T., Hirsch, T. B., Martin, C. C., Jerabek, P. A., & Lancaster, J. L. (1995). Use of implicit motor imagery for visual shape discrimination as revealed by PET. *Nature*, *375*, 54–58.
- Piaget, J. (1962). *Play, dreams, and imitation in childhood*. New York: Norton.
- Poizner, H., Bellugi, U., & Tweney, R. (1981). Processing of formational, semantic, and iconic information in American Sign Language. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 1146–1159.
- Prinz, W. (1987). Ideo-motor action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 47–76). Hillsdale, NJ: Erlbaum.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*, 129–154.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000, April 14). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*, 349–351.
- Reed, C. L., & Farah, M. J. (1995). The psychological reality of the body schema: A test with normal participants. *Journal of Experimental Psychology*, *21*, 334–343.
- Reid, T. (1965). Essays on the intellectual powers of man. In R. J. Herrnstein & E. G. Boring (Eds.), *A source book in the history of psychology* (pp. 172–178). Cambridge, MA: Harvard University Press. (Original work published 1785)
- Reisberg, D., Smith, J. D., Baxter, D. A., & Sonenshine, M. (1989). “Enacted” auditory images are ambiguous; “pure” auditory images are not. *Quarterly Journal of Experimental Psychology*, *41A*, 619–641.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience*, *21*, 188–194.
- Rizzolatti, G., & Arbib, M. A. (1999). Reply. *Trends in Neuroscience*, *22*, 152.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., & Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research*, *111*, 246–252.

- Roberts, L. A. (1986). Modality and suffix effects in memory for melodic and harmonic musical materials. *Cognitive Psychology*, 18, 123–157.
- Romanski, L. M., Tian, B., Fritz, J. B., Mishkin, M., Goldman-Rakic, P. S., & Rauschecker, J. P. (2000). *Nature Neuroscience*, 3, 965–966.
- Rubin, D. C., & Kontis, T. C. (1983). A schema for common cents. *Memory & Cognition*, 11, 335–341.
- Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, 112, 585–615.
- Scheerer, E. (1984). Motor theories of cognitive structure: A historical review. In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 77–97). Berlin, Germany: Springer-Verlag.
- Shepard, R. N. (1994). Perceptual–cognitive universals as reflections of the world. *Psychonomic Bulletin & Review*, 1, 2–28.
- Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science*, 1, 257–264.
- Shiffrar, M., & Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science*, 4, 379–384.
- Shiffrar, M., Lichtey, L., & Chatterjee, S. H. (1997). The perception of biological motion across apertures. *Perception & Psychophysics*, 59, 51–59.
- Shiffrar, M., & Pinto, J. (in press). The visual analysis of bodily motion. In W. Prinz & B. Hommel (Eds.), *Attention and performance: Vol. 19. Common mechanisms in perception and action*. Oxford, England: Oxford University Press.
- Smith, J. D., Reisberg, D., & Wilson, M. (1992). Subvocalization and auditory imagery: Interactions between the inner ear and inner voice. In D. Reisberg (Ed.), *Auditory imagery* (pp. 95–119). Hillsdale, NJ: Erlbaum.
- Smith, J. D., Wilson, M., & Reisberg, D. (1995). The role of inner speech in auditory imagery. *Neuropsychologia*, 33, 1433–1454.
- Smyth, M. M., Pearson, N. A., & Pendleton, L. R. (1988). Movement and working memory: Patterns and positions in space. *Quarterly Journal of Experimental Psychology*, 40A, 497–514.
- Smyth, M. M., & Scholey, K. A. (1994a). Characteristics of spatial memory span: Is there an analogy to the word length effect, based on movement time? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 47A, 91–117.
- Smyth, M. M., & Scholey, K. A. (1994b). Interference in immediate spatial memory. *Memory & Cognition*, 22, 1–13.
- Spoehr, K. T., & Corin, W. J. (1978). The stimulus suffix effect as a memory coding phenomenon. *Memory & Cognition*, 6, 583–589.
- Stevens, J. A., Fonlupt, P., Shiffrar, M., & Decety, J. (2000). New aspects of motion perception: Selective neural encoding of apparent human movements. *NeuroReport*, 11, 109–115.
- Surprenant, A. M., Pitt, M. A., & Crowder, R. G. (1993). Auditory recency in immediate memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 46A, 193–223.
- Thornton, I. M., Pinto, J., & Shiffrar, M. (1998). The visual perception of human locomotion. *Cognitive Neuropsychology*, 15, 535–552.
- Valenti, S. S., & Costall, A. (1997). Visual perception of lifted weight from kinematic and static (photographic) displays. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 181–198.
- Viviani, P. (1990). Motor–perceptual interactions: The evolution of an idea. In M. P. Palmari (Ed.), *Cognitive science in Europe: Issues and trends* (pp. 11–39). Boulder, CO: Golem.
- Viviani, P., Baud-Bovy, G., & Redolfi, M. (1997). Perceiving and tracking kinesthetic stimuli: Further evidence of motor–perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1232–1252.
- Viviani, P., & Mounoud, P. (1990). Perceptual–motor compatibility in pursuit tracking of two-dimensional movements. *Journal of Motor Behavior*, 22, 407–443.
- Viviani, P., & Stucchi, N. (1989). The effect of movement velocity on form perception: Geometric illusions in dynamic displays. *Perception & Psychophysics*, 46, 266–274.
- Viviani, P., & Stucchi, N. (1992a). Biological movements look uniform: Evidence of motor–perceptual interactions. *Journal of Experimental Psychology*, 18, 603–623.
- Viviani, P., & Stucchi, N. (1992b). Motor–perceptual interactions. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (Vol. 2, pp. 229–248). Amsterdam: Elsevier Science.
- Whalen, D. H., & Liberman, A. M. (1987, July 10). Speech perception takes precedence over nonspeech perception. *Science*, 237, 169–171.
- Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin & Review*, 8, 44–57.
- Wilson, M., Bettger, J. G., Niculae, I., & Klima, E. S. (1997). Modality of language shapes working memory: Evidence from digit span and spatial span in ASL signers. *Journal of Deaf Studies and Deaf Education*, 2, 150–160.
- Wilson, M., & Emmorey, K. (1997a). A “phonological loop” in visuospatial working memory: Evidence from American Sign Language. *Memory & Cognition*, 25, 313–320.
- Wilson, M., & Emmorey, K. (1997b). Working memory for sign language: A window into the architecture of working memory. *Journal of Deaf Studies and Deaf Education*, 2, 123–132.
- Wilson, M., & Emmorey, K. (1998a, November). *Spatial coding in working memory for signs*. Paper presented at the Sixth International Conference on Theoretical Issues in Sign Language Research, Washington, DC.
- Wilson, M., & Emmorey, K. (1998b). A “word length effect” for sign language: Further evidence on the role of language in structuring working memory. *Memory & Cognition*, 26, 584–590.
- Wilson, M., Iverson, A., & Emmorey, K. (2001). *Further investigation of the phonological similarity effect for sign language: Two effects of spatial similarity*. Manuscript submitted for publication.

Received June 1, 2000

Revision received November 27, 2000

Accepted December 19, 2000 ■