

# Journal of Experimental Psychology: General

## **Understanding Gesture: Is the Listener's Motor System Involved?**

Raedy M. Ping, Susan Goldin-Meadow, and Sian L. Beilock

Online First Publication, April 8, 2013. doi: 10.1037/a0032246

### CITATION

Ping, R. M., Goldin-Meadow, S., & Beilock, S. L. (2013, April 8). Understanding Gesture: Is the Listener's Motor System Involved?. *Journal of Experimental Psychology: General*. Advance online publication. doi: 10.1037/a0032246

# Understanding Gesture: Is the Listener's Motor System Involved?

Raedy M. Ping, Susan Goldin-Meadow, and Sian L. Beilock  
The University of Chicago

Listeners are able to glean information from the gestures that speakers produce, seemingly without conscious awareness. However, little is known about the mechanisms that underlie this process. Research on human action understanding shows that perceiving another's actions results in automatic activation of the motor system in the observer, which then affects the observer's understanding of the actor's goals. We ask here whether perceiving another's *gesture* can similarly result in automatic activation of the motor system in the observer. In Experiment 1, we first established a new procedure in which listener response times are used to study how gesture impacts sentence comprehension. In Experiment 2, we used this procedure, in conjunction with a secondary motor task, to investigate whether the listener's motor system is involved in this process. We showed that moving arms and hands (but not legs and feet) interferes with the listener's ability to use information conveyed in a speaker's hand gestures. Our data thus suggest that understanding gesture relies, at least in part, on the listener's own motor system.

*Keywords:* gesture, embodiment, motor simulation, gesture understanding, sentence comprehension

When we watch others act on the world, our own motor systems are activated, which, in turn, affects how we interpret the actors' goals (e.g., Buccino et al., 2001; Hamilton, Wolpert, & Frith, 2004; Sebanz, Bekkering, & Knoblich, 2006; Wilson, Collins, & Bingham, 2005). Here we ask whether this same process takes place even when the actions we observe do not have a direct effect on the world but rather impact the world indirectly through their communicative potential. When people speak, they often move their hands—they gesture—and listeners are able to glean substantive information from these gestures, although typically without being aware of doing so. Does watching a speaker who gestures activate our own motor system?

Perception and action have been shown to be linked, both within an individual and across individuals. In one line of studies, participants must perceive or produce bilateral movements that are perceptually and motorically difficult. Without training, humans

can easily perceive and produce movements that are either identical (e.g., simultaneously moving both pointer fingers left and right together—a 0° phase) or symmetric (e.g., simultaneously moving both pointer fingers in and out together—a 180° phase). Distinguishing or producing bilateral movements at any other phase is difficult, requires extensive practice, and falls apart once movements reach a threshold frequency. Within an individual, learning to perceive distinctions at difficult phases (perception) improves the ability to produce movements (action) at those phases—movements that otherwise would require extensive motor practice to master (Wilson, Snapp-Childs, & Bingham, 2010; see also Bingham, Schmidt, & Zaal, 1999; Zaal, Bingham, & Schmidt, 2000). Conversely, learning to produce movements at particular phases, without visual feedback of one's own body, improves perceptual discrimination of those phases specifically (Hecht, Vogt, & Prinz, 2001). Across individuals, there is considerable overlap between the neural circuitry activated in perceiving someone perform an action and the neural circuitry activated when we ourselves plan and produce that same action (e.g., Buccino et al., 2001; Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2005; Hamilton et al., 2004; Jacobs & Shiffrar, 2005; Maeda, Mazziotta, & Iacoboni, 2002). If the motor system is recruited when an observer attempts to understand another's action, then reducing the motor resources available to the observer should have an impact on the way that action is understood (e.g., Beilock & Holt, 2007).

In one of many experiments demonstrating this effect, Reed and McGoldrick (2007) asked observers to judge whether two sequentially presented pictures of body postures were the same or different; in some of the trials, the leg posture differed in the two pictures; in others, the arm posture differed. While observers were making their judgments about the postures, they planned and produced movements with either their arms or legs. When the interval between the two pictures was relatively short (2 s), observers were less accurate in detecting changes in arm posture while moving their arms and less accurate in detecting changes in

---

Raedy M. Ping, Susan Goldin-Meadow, and Sian L. Beilock, Department of Psychology, The University of Chicago.

This work was supported by National Science Foundation (NSF) Grant SBE-0541957 (Spatial Intelligence and Learning Center; Principal Investigator: Nora Newcombe; Co-Principal Investigator: Susan Goldin-Meadow), National Institute of Child Health and Human Development Grant R01-HD47450, and NSF Grant BCS-0925595 to Susan Goldin-Meadow; and by NSF Fostering Interdisciplinary Research on Education (FIRE) Grant DRL-1042955 to Sian L. Beilock.

We would like to thank the following people for assistance with data collection and stimuli design: Maggie Kendall Zimmerman, Ji-Sook Yim, Mary-Anne Decatur, and Hector Santana. These studies were conducted as part of the first author's doctoral dissertation, and we would like to thank Howard Nusbaum and Katherine Kinzler for comments and feedback during that process.

Correspondence concerning this article should be addressed to Raedy M. Ping, Department of Psychology, University of Chicago, 5848 S. University Avenue, Chicago, IL 60637. E-mail: raedyping@gmail.com

leg posture while moving their legs. When the interval between the two pictures was longer (5 s), the pattern changed, and observers were more accurate in detecting changes in the particular body part they were moving. The body movements the observer planned and executed while processing another's body configurations affected the way the observer judged those configurations. As long as the observer's body movements relied on the same body parts as he was judging, his understanding of another's body was affected.<sup>1</sup>

Activating the planning and execution resources of the motor system when observing another person act is thought to have a social function—it helps the observer to understand the goals and intentions of the other and to plan her own action responses accordingly (e.g., Knoblich & Sebanz, 2006; Sebanz, Bekkering, & Knoblich, 2006; Woodward, Sommerville, Gerson, Henderson, & Buresh, 2009). This social effect has been studied using a variation of the “Simon” paradigm. In one study (Sebanz, Knoblich, & Prinz, 2003), participants were instructed to respond to one color prompt (a green ring) by using their left hand to push a button on their left side and to another color prompt (a red ring) by using their right hand to push a button on their right side. On some trials, the ring was presented on a finger pointing toward the right; on others, the ring was presented on a finger pointing toward the left. Although this direction information was completely irrelevant to their instructed task, participants were slower to respond if the green ring was presented on a finger pointing toward the right (and vice versa for the red ring). This pattern presumably reflects the fact that the participant must inhibit the directional information when responding to color information on trials where ring color and finger point direction were incompatible. Another group of participants was asked to track and respond to only one of the two prompts (green ring *or* red ring), effectively participating in a go–no go task. These participants were equally fast to respond to their assigned ring color, regardless of the direction of the finger point—in other words, there was no interference. The crucial group of participants was assigned to respond to only one ring color, just like the participants in the go–no go condition; however, they were seated next to a partner who was instructed to respond to the other ring color. Participants in this condition showed interference between finger direction and ring color—they were slower to respond to their ring color when the finger was pointing toward the opposite side of the screen. These data suggest that individuals understand and represent the actions of *others* in functionally the same way as they represent their own actions (Sebanz, Knoblich, & Prinz, 2003; Sebanz, Knoblich, Prinz, & Wascher, 2006; also see Hommel, Colzato, & van den Wildenberg, 2009).

Not all human actions have direct physical effects on the surrounding world, and not all actions require a physical response from others. People spontaneously, and frequently, produce hand gestures when they speak—representational actions that do not have a direct impact on the physical world but impact the world indirectly through their communicative potential.<sup>2</sup> Studies have found that listeners encode and understand the information conveyed in a speaker's gestures. For example, participants who watched and listened to learners explain how they solved math problems credited the learners with problem-solving strategies that they had produced *only* in gesture and not in speech—even though the participants were never told to attend to gesture (Alibali, Flevares, & Goldin-Meadow, 1997; Goldin-Meadow & Sandhofer,

1999). As another example, when asked to retell a narrative told to them by a speaker who gestured, listeners incorporated information that was conveyed *only* in the speaker's gestures into their own verbal retellings of the story (McNeill, Cassell, & McCullough, 1994). Interviews with the listeners suggested that they were unaware of the source of the gestured information—they typically reported that the speaker had conveyed the information in speech. These studies make it clear that listeners can interpret the information conveyed in a speaker's gestures, even when they are not explicitly told to attend to it. However, very little is known about *how* listeners incorporate information from a speaker's gestures into their understanding of the accompanying speech. Our goal in the current studies was to explore the mechanism that underlies this phenomenon.

In Experiment 1, we established a new procedure for studying gesture understanding in the context of talk by measuring listener response times immediately following a speaker's utterance. The experiments described earlier did not tap gesture understanding in real time in ecologically valid ways. The Goldin-Meadow and Sandhofer (1999) study required an initial phase where listeners were trained on various problem-solving strategies so that they could later categorize children's explanations about how they solved a specific problem. The McNeill, Cassell, and McCullough (1994) study relied on listeners' retellings of a narrative, which took place quite a while after the initial telling of the story. These paradigms, while useful for demonstrating that listeners can understand gestured information not conveyed in speech, are not adequate for studying exactly how information is gleaned from gesture. The paradigm we used in Experiment 1 relies on listener response times immediately following a speaker's utterance (within 1 s), and thus can be used to study the mechanisms underlying gesture understanding in real time.

In Experiment 2, this newly developed procedure was used in conjunction with a simultaneous arm or leg movement task (cf. Reed & Farah, 1995; Reed & McGoldrick, 2007) to investigate whether the listener's motor system is involved in processing a speaker's gestures. Gestures do not require a physical response from the listener. However, seeing a speaker's gestures could activate the listener's own motor system, which, in turn, could affect how the listener understands the representational information conveyed in gesture. If so, then asking listeners to perform a concurrent motor task while observing a speaker's gestures could have an impact on how those gestures are understood.

## Experiment 1

In our procedure, reaction time is used to investigate whether listeners use the information conveyed in a speaker's gestures in building a mental model of a speaker's message. This procedure is based on a paradigm developed by Zwaan and colleagues (e.g., Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002) in

<sup>1</sup> Reed and McGoldrick's (2007) explanation for the difference between interstimulus intervals (ISIs) of 2 s and ISIs of 5 s is that visual motor integration of self and other movements requires time to activate shared portions of the body schema. At short time scales, embodied information is still unintegrated and therefore interferes with understanding; at longer time scales, this integration has occurred, resulting in facilitation of understanding.

<sup>2</sup> We focused in this study on iconic, pointing, and tracing gestures.

which participants read a sentence and then respond to a drawing of an object. Participants are instructed to respond “yes” if the name of the pictured object was mentioned in the sentence (target trials) and “no” if it was not (filler trials, included to keep participants on task). Zwaan and colleagues found that contextual information provided by the sentence affected how quickly participants responded “yes” to pictures in the target trials. For example, after reading the sentence, “He hammered the nail into the floor,” participants were faster to say “yes” to a picture of a vertically oriented nail (congruent context–picture trial) than to a picture of a horizontally oriented nail (incongruent context–picture trial). The opposite pattern of reaction times was found for the two pictures when the sentence was, “He hammered the nail into the wall.” These findings suggest that language comprehenders infer meaning (in this case, orientation of the nail) from contextual information inherent in the sentence and incorporate it into their models of the written sentence.

Our procedure allowed us to explore whether listeners incorporate information represented in a speaker’s gestures when building mental models of spoken language. Participants saw a video clip of a woman speaking a sentence (e.g., “The woman hammered the nail into the wood”). On some trials, the woman produced a gesture that conveyed additional information about the object—for this example, a gesture demonstrating hammering a vertically oriented nail *or* a gesture demonstrating hammering a horizontally oriented nail (see Table 1<sup>3</sup>). Participants were then shown a picture (in this example, a picture of either a vertically oriented nail, or a horizontally oriented nail). As in Zwaan’s paradigm, the measure of interest was the time participants took to indicate whether the object in the picture had been named in the spoken sentence. If listeners process information from the gesture and incorporate it into their perceptual model of the spoken sentence, they should be faster to respond “yes” when gesture and picture are congruent (e.g., vertical hammering gesture followed by a picture of a vertically oriented nail) than when gesture and picture are incongruent (e.g., vertical hammering gesture followed by a picture of a horizontally oriented nail).





## Method

**Participants.** Forty-eight right-handed English-speaking undergraduate students (mean age = 20.56 years,  $SE = 0.46$  years; 63% were women) participated in Experiment 1 in exchange for course credit or a small payment.

### Procedure.

**Audio-only familiarization trials (presented with feedback).** Participants were seated at a computer and received 10 audio-only sentences. The purpose of the audio-only sentences was to familiarize participants with what was considered an appropriate “yes” versus “no” response. They were instructed to listen to the sentence and then respond to the line drawing presented afterward. Participants were told to respond “yes” if a name for the pictured object had been mentioned in the sentence and “no” if it had not. For example, participants were trained to respond “yes” to the picture following the sentence, “The light bulb was hanging from the ceiling,” even if the picture depicted a light bulb shattered on the ground (because the object name, “light bulb,” was spoken in the sentence). In other words, participants were taught to respond to the object names mentioned in the sentence, not to the

Table 1  
Examples of the Different Target Trial Types for the “Yes” Sentence “The woman hammered the nail into the wood”

	Horizontal nail gesture	Vertical nail gesture
		
Horizontal nail picture 	Congruent	Incongruent
Vertical nail picture 	Incongruent	Congruent

*Note.* During target trials in the main experimental block, some participants saw congruent gesture–picture combinations (i.e., horizontal gesture followed by horizontal picture, or vertical gesture followed by vertical picture). Others saw incongruent gesture–picture combinations (i.e., horizontal gesture followed by vertical picture, or vertical gesture followed by horizontal picture).

overall scenario described in the sentence. For each audio-only sentence, participants heard a spoken sentence (without video) and then responded to a picture presented on the computer screen using a keyboard placed in front of them. They received positive feedback from the computer for pushing the *y* key (for “yes”) if the pictured object had been named by a noun in the sentence and for pushing the *n* key (for “no”) if the pictured object had not been named; they received negative feedback for all other responses.

<sup>3</sup> In approximately half of the “yes” sentences, the gesture produced along with the sentence represented an agent’s hand acting on an object or instrument (see the examples in Table 1). In the other half, the gesture represented an object, either by tracing its outline or indicating its habitual location (e.g., an index finger traces an arc on the chest around the neck to indicate a necklace). There were no systematic differences in responses as a function of these two types of gestures, although the number of items of each type was not large enough to adequately explore this issue. It is important to note that none of the hand gestures used in the study represented movements of the foot (e.g., pressing the palm down as though putting one’s foot on the brake).

**Audio-visual test trials (presented without feedback).** Participants were informed that they would now see video clips (with audio) instead of hearing only an audio clip and that they should continue to respond to the pictures that followed the sentences as quickly and accurately as possible. They were also told that instead of keying in their response, they would make the response orally.<sup>4</sup> Participants were instructed to say “yes” if a name for the pictured object had been mentioned in the sentence and “no” if it had not. Participants spoke into a microphone; the computer recorded reaction time (latency between picture presentation and microphone trigger) while the experimenter manually recorded the yes/no response. After the participant responded, there was a 1,000-ms pause and then the next video began.

We assembled a total of 80 “yes” sentences to be presented via video clip (e.g., “The woman hammered the nail into the wood,” followed by a picture of a vertically or horizontally oriented nail); the majority of these sentences were created specifically for this study, but some were modifications of items from Stanfield and Zwaan (2001); Zwaan, Stanfield, and Yaxley (2002), or Holt and Beilock (2006). We also assembled 80 “no” sentences (fillers, e.g., “The boy played the harmonica,” followed by a picture of a saxophone).

**Counterbalancing: Main experimental block.** In the main experimental block, each participant saw one of four versions of the sentences. In each version, 40 of the “yes” sentences were presented with either a congruent (20) or incongruent (20) gesture (target items). The other 40 “yes” sentences were presented without gesture—the speaker in the video clip did not move her hands. The no-gesture sentences were included to ensure that participants saw an equal number of trials with gesture and without gesture in an effort to keep them from becoming overly aware of the speaker’s gestures and explicitly relying on them for information in a way that listeners typically do not (see Goldin-Meadow & Sandhofer, 1999; McNeill et al., 1994). Thus, each participant was presented with all 80 “yes” sentences. However, whether each particular “yes” sentence was paired with gesture was counterbalanced across participants.

In line with the procedure used by Zwaan and colleagues and to equalize the number of “yes” and “no” responses for each participant, 80 sentences designed to elicit a “no” response were also included. Like the “yes” sentences, 40 “no” sentences were presented with gesture, and 40 “no” sentences were presented without gesture. In the “no” sentences, gesture was irrelevant to the response simply because it captured aspects of the object mentioned in the sentence but not portrayed in the picture (e.g., a gesture miming playing a harmonica produced with the sentence, “The boy played the harmonica,” followed by a picture of a saxophone). All of the “no” sentences, as well as the “yes” sentences presented without gesture, served as filler items to balance the total number of “yes” and “no” sentences, and the total number of sentences with and without gesture, that participants saw. Data from these filler sentences were not analyzed. Our measure of interest was the response time for “yes” sentences when presented in congruent versus incongruent gesture–picture combinations.

**Counterbalancing: Practice phase.** Prior to the main experimental block, we included a practice phase to familiarize participants with the audio-visual procedure and the stimuli used in the study. We also used accuracy data from the practice phase to ensure that the line drawings we crafted were recognizable by

participants. The practice phase trials were analogous to those in the main experimental block—each participant saw 80 “yes” sentences—half with gesture and half without—and 80 “no” sentences—again, half with gesture and half without. The gestures used during the main experimental block were *not* seen during the practice phase. For a given participant, the “yes” sentences that were accompanied by gesture during the experimental block were presented without gesture during the practice phase. We can therefore use responses to these practice phase trials to ensure that participants did not show systematic biases for particular sentence–picture pairs, when gesture is not a factor. That is, during the practice phase, participants should be equally fast to respond to sentences that will later be presented in incongruent or congruent gesture–picture combination trials, since at this point each subset of trials should be functionally identical.

**Procedure summary.** In total, each participant took part in 320 trials spread across the practice phase and main experimental block. In the main experimental block, participants saw 40 “yes” sentences with gesture and 40 “yes” sentences without gesture. They also saw 40 “no” sentences with gesture and 40 “no” sentences without gesture, for a total of 160 trials. These trials were presented in a random order within each participant and were counterbalanced across participants such that each “yes” and each “no” sentence was seen with and without gesture by a subset of participants during the main experimental block. The practice phase had the same pattern of trials but with different video clips; target sentences that appeared during the practice phase without gesture were presented with gesture during the main experimental block. Following the sentence task, participants completed a demographic questionnaire and wrote down their best guess about the purpose of the study before being debriefed. The entire study took approximately 30 min.

## Results and Discussion

On nine target sentences presented without gesture during the practice phrase, 10% or more of participants failed to realize that the object displayed in the picture had been mentioned in the preceding sentence (i.e., they failed to say “yes”). On the basis of these low accuracy levels, these nine target sentences were removed from all analyses.

We found that accuracy was high for the remaining 71 “yes” sentences ( $M = 96.8\%$ ,  $SE = 0.7\%$ ) during the main experimental block. We analyzed reaction times (RTs) for these sentences when they were paired with gesture during this block, focusing on the sentence–picture pairs to which participants accurately responded “yes” and removing outlier RTs (scores that were beyond two standard deviations from the mean for a participant). Overall, 95% of “yes” sentence data were included in the analyses. Figure 1 presents the mean RTs.

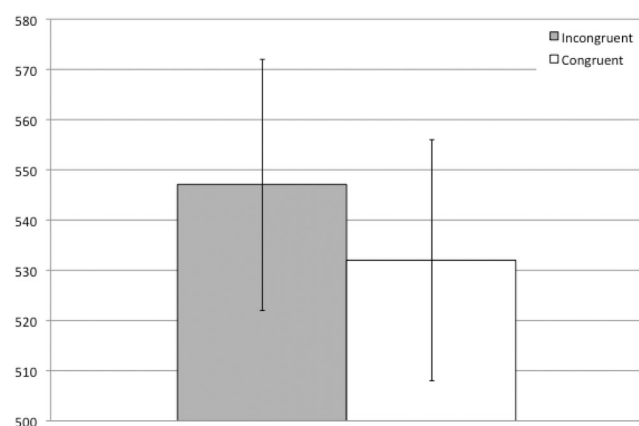
RTs were analyzed in a linear mixed model with participant and sentence as random effects and gesture–picture congruence as a fixed effect (number of observations = 1,622). Gesture–picture congruence was a significant predictor of RT (coefficient estimate =  $-17.18$ ,

<sup>4</sup> Key presses were used in the audio-only sentences so that participants could receive feedback from the computer. The procedure changed to oral responses at this point in the study so that participants kept their arms and hands still.

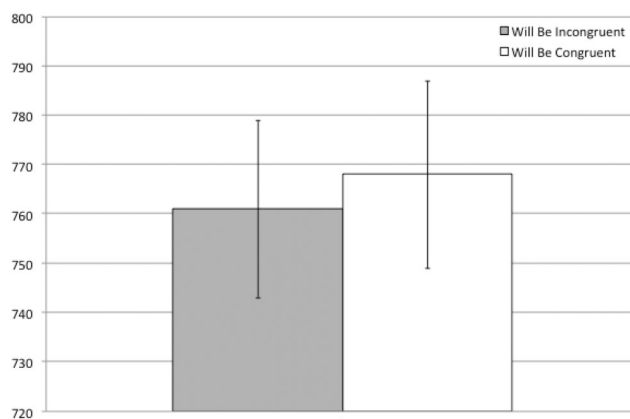
$SE = 7.84, t = -2.19, p = .0285$ )—with shorter RTs for congruent gesture–picture combinations. This model predicted the data better than a model with only random participant (variance = 27,936.35;  $SD = 167.14$ ) and sentence (variance = 905.48,  $SD = 30.09$ ) effects (Bayesian information criteria [BIC] for three-term model = 21,258; BIC for two-term model = 21,260),  $\chi^2(1) = 4.80, p = .0285$ . RTs were faster for congruent gesture–picture combinations (e.g., vertical hammering + vertical nail) than for incongruent gesture–picture combinations (e.g., vertical hammering + horizontal nail).

If these results were driven by the congruence between the speaker's *gestures* and the pictures, and not by some other relationship between the sentences and the pictures, then we should find no differences in RTs when the same sentence–picture pairs were presented without gesture. Our counterbalancing protocol allowed us to address this issue because each participant heard the same “yes” sentences presented *without gesture* during the practice phase that she heard presented *with gesture* during the main experimental block. RTs for these trials were analyzed in a linear mixed model with participant and sentence as random effects and later gesture–picture congruence (i.e., whether the sentence would later be paired in an incongruent or a congruent combination for that particular participant during the main experimental block) as a fixed effect (number of observations = 1,626). Mean RTs are presented in Figure 2. As predicted, later gesture–picture congruence was not a significant predictor (coefficient estimate = 5.30,  $SE = 8.33, t \text{ value} = 0.64, p = .5222$ ) of RT during the practice phase when sentences were presented without gesture. Further, the model including the later gesture–picture congruence term did not improve data fit beyond a model including only the random participant (variance = 15,341,  $SD = 123.86$ ) and sentence (variance = 3,419,  $SD = 58.47$ ) effects (BIC for three-term model = 21,524; BIC for two-term model = 21,517),  $\chi^2(1) = 0.40, p = .5246$ .<sup>5</sup> Although we must be cautious in interpreting data from the practice phase before participants had been thoroughly familiarized with the task, this analysis does suggest that the congruence effect in the main experimental block is likely to be due to the information presented in gesture.<sup>6</sup>

In sum, when the information conveyed in gesture matched the information conveyed in the picture that followed, listeners were



**Figure 1.** Mean reaction times (RTs), tallied by-participant, for the main experimental block in Experiment 1. RTs were shorter for pictures following sentences with congruent gesture than for pictures following sentences with incongruent gesture. Error bars are standard errors of the mean.



**Figure 2.** Mean reaction times (RTs), tallied by-participant, for the practice phase in Experiment 1 when sentences were presented without gesture. RTs were no shorter for pictures following sentences that would later be paired with congruent gesture during the main experimental block than they were for pictures following sentences that would later be paired with incongruent gesture. Error bars are standard errors of the mean.

relatively quick to respond that the pictured object had been named in the sentence. Participants were never instructed to attend to gesture and, in the majority of sentences that contained gesture, gesture was either irrelevant (“no” sentences) or misleading (incongruent trials). These data demonstrate that listeners incorporate information from a speaker’s hand gestures into the mental representations they form of the speaker’s message. Experiment 2 explores a potential mechanism underlying this phenomenon.

## Experiment 2

Listeners responded faster to congruent than to incongruent gesture–picture combinations in Experiment 1, suggesting that our picture judgment task is an effective paradigm for investigating gesture perception on a relatively short time scale (less than 1 s after the speaker’s utterance). In Experiment 2, listeners completed the same picture judgment task as in Experiment 1 (the primary task) while simultaneously performing a motor task (the secondary

<sup>5</sup> A two-way interaction between block (practice phase, main experimental block) and gesture–picture congruence during the main experimental block (congruent, incongruent) was a significant predictor (coefficient estimate =  $-29.06, SE = 12.72, t \text{ value} = -2.28, p = .0226$ ) in a model of the RT data from both blocks of trials, warranting the separate analysis of RT data for the main experimental block and practice phase.

<sup>6</sup> Note that all participants heard each sentence, presented without gesture, and saw all pictures during the practice phase. The participants’ responses during the main experimental block therefore may have been influenced by the fact that they had already heard the sentences and seen the pictures once during the practice phase. We counterbalanced the stimuli in this way so that we could be certain that it was pairing the sentence/picture with the gestures, rather than some other aspect of the sentence/picture, that determined the pattern of responses. Indeed, we found that RTs during the experimental block differed as a function of the congruence between sentence/picture (Figure 1), but RTs to the same sentence/pictures presented without gestures during the practice phase did not (Figure 2). Taken together, the findings make it clear that the participants processed the gestures they saw and used the information to construct mental models of the sentences.

task). If listeners call upon motor processes that are involved in executing gesture when watching a speaker gesture, then busying those motor processes with the planning and execution of movements should interfere with the congruence effect in the primary task. This interference should be specific to motor resources controlling the effectors used in producing gesture—in this case, the arms and hands. Consuming motor resources that control other parts of the body—for example, the legs and feet—should not interfere with the congruence effect (see Beilock & Holt, 2007; Yang, Gallo, & Beilock, 2009).

To test this prediction, we asked half of the listeners in Experiment 2 to plan and execute movements with their arms and hands, the effectors used by the speaker gesturing in the video clips (arm movements condition), while watching the videos and responding to the pictures. If understanding gesture involves motor activation in the listener, then listeners who move their arms and hands should experience interference on the picture judgment task. This interference should eliminate the congruence effect found in Experiment 1; that is, there should be no difference in RTs to congruent versus incongruent gesture-picture combinations.<sup>7</sup>

As a control, the other half of listeners planned and executed movements with their legs and feet, different effectors from those used by the speaker to gesture (leg movements condition), while watching the videos and judging the pictures. Planning and executing leg movements should *not* interfere with gesture understanding simply because the effectors in the two acts are not the same (arms in the speaker, legs in the listener). Listeners in the leg movements condition should therefore show the congruence effect found in Experiment 1; that is, they should show faster RTs for congruent than for incongruent gesture-picture combinations. Overall, our hypothesis was that planning and producing arm movements would interfere with the gesture-picture congruence effect, while planning and producing leg movements would not.

## Method

**Participants.** Ninety-six right-handed English-speaking undergraduates participated in Experiment 2 (mean age = 22.15 years;  $SE = 0.46$  years, 69% were women). One participant in each condition did not follow the motor task instructions, leaving data from 94 participants for analysis.

**Procedure.** The stimuli, design, counterbalancing, and procedure of the picture judgment task used in Experiment 2 were identical to Experiment 1. After completing the audio-only feedback trials for the primary picture judgment task, participants were introduced to the secondary motor task. Participants in the arm movements condition were asked to move their arms and hands continuously while watching the video clips and judging the pictures. They were told that they could make any movements they wanted as long as they were not repetitive. Movements needed to be nonrepetitive so that participants would be continuously planning novel movements to execute. This planning is thought to engage premotor resources (resources important for decoding others' actions; see Reed & Farah, 1995, and Reed & McGoldrick, 2007). Participants in the leg movements condition were given the same instructions with respect to their legs and feet. After receiving instructions in the secondary motor task, participants completed a few trials and were given feedback on their movements.

Participants were typically able to follow instructions, with an infrequent reminder that movements could not be repetitive.

## Results and Discussion

For consistency across experiments, the nine target sentences that had been eliminated from the analyses in Experiment 1 because of low accuracy (i.e., participants did not consistently report that the object displayed in the picture had been mentioned in the preceding sentence) were also eliminated from the analyses in Experiment 2. Accuracy levels for the remaining 71 sentences were high for “yes” sentences in the main experimental block ( $M = 96.8\%$ ,  $SE = 0.4\%$ ). RTs for non-outlying target trials where participants correctly responded “yes” were analyzed—data from 93% of “yes” sentence data are included in the following analyses.

As in Experiment 1, RTs for the main experimental block were analyzed in a linear mixed model with participant and sentence as random effects. Condition (arm movements, leg movements), gesture-picture congruence (congruent, incongruent), and the interaction between them were entered as fixed effects (number of observations = 3,122, see Table 2 for model summary). Mean RTs are displayed in Figure 3. Gesture-picture congruence did not significantly predict RT (coefficient estimate = 8.49,  $SE = 10.80$ ,  $t$  value = 0.79,  $p = .4295$ ) across the two conditions. Condition was a significant predictor of RT (coefficient estimate = 65.00,  $SE = 10.84$ ,  $t$  value = 6.00,  $p < .0001$ )—RTs in the leg movements condition were longer overall than those in the arm movements condition. One possible explanation for this main effect is that planning and producing nonrepetitive leg and foot movements might require more attention, which takes time to deploy, compared with the attention required for planning and producing nonrepetitive arm and hand movements. Another possibility is that—counter to our hypothesis—planning and producing leg and foot movements could interfere with simulating the speaker's gestures regardless of meaning, leading to longer RTs for both congruent and incongruent gesture-picture trials. We return to these two possibilities later.

Our research question centers on the relationship between information conveyed in gesture and body movement so the key term of interest in our analyses is the interaction between gesture-picture congruence (congruent vs. incongruent) and condition (arm vs. leg movements). We hypothesized that gesture-picture congruence would not predict RT for the arm movements condition but would for the leg movements condition. This interaction term was, in fact, a significant predictor of RT (coefficient estimate =  $-40.20$ ,  $SE = 15.23$ ,  $t$  value =  $-2.64$ ,  $p = .0083$ ), suggesting that gesture-picture congruence predicts RT differentially in each of the two conditions. The model including the interaction term fit the data better than a model with only the two main effects and two random effects (BIC for five-term model = 42,539; BIC for four-term model = 42,540),  $\chi^2(1) = 6.97$ ,  $p = .0083$ . Taken

<sup>7</sup> We predicted interference, not facilitation, in processing for participants in this condition because of the relatively short interval between seeing the gesture and responding—in this study, less than 3 s. Reed and McGoldrick (2007) found interference in processing body posture changes due to simultaneous effector-specific movement at 2-s intervals and facilitation in processing changes at 5-s intervals.

Table 2  
Summary of the Random and Fixed Effects Predicting Reaction Time During the Main Experimental Block in Experiment 2  
( $N = 3,122$ )

Variable	Variance	Standard deviation	Coefficient estimate	Standard error	$t$	$p$ (estimate)
Random effect						
Participant	11,209	105.88				
Sentence	586	24.20				
Fixed effect						
(Intercept)			591.54	17.33	34.13	<.0001
Condition (Leg)			65.00	10.84	6.00	<.0001
Gesture–picture relationship (Congruent)			8.49	10.80	0.79	.4295
Condition $\times$ Gesture–Picture interaction			–40.20	15.23	–2.64	.0083

together, these findings warrant analysis of RT data for each condition separately.

For the leg movements condition (number of observations = 1,567), the fixed effect of gesture–picture congruence was a significant predictor of RT (coefficient estimate =  $-32.81$ ,  $SE = 9.12$ ,  $t$  value =  $-3.60$ ,  $p = .0003$ ), and the model with gesture–picture congruence fit the data better than a model with only the random participant (variance = 19,268,  $SD = 138.81$ ) and sentence (variance = 1,264,  $SD = 35.56$ ) effects (BIC for three-term model = 20,933; BIC for two-term model = 20,939),  $\chi^2(1) = 12.89$ ,  $p = .0003$ . For participants who were planning and producing movements with their legs and feet RTs were shorter for congruent than for incongruent gesture–picture combinations—replicating the gesture–picture congruence effect found in Experiment 1.

For the arm movements condition (number of observations = 1,555), the fixed effect of gesture–picture congruence was *not* a significant predictor of RT (coefficient estimate =  $8.52$ ,  $SE = 9.84$ ,  $t$  value =  $0.87$ ,  $p = .3898$ ), and the model including the congruence term did not fit the data any better than a model including only the random participant (variance = 23,223,  $SD = 152.39$ ) and sentence (variance = 735,  $SD = 27.11$ ) effects (BIC for three-term model = 20,985; BIC for two-term model = 20,978),  $\chi^2(1) = 0.75$ ,  $p = .3862$ . Planning and producing arm and

hand movements eliminated the gesture–picture congruence effect found in Experiment 1 and in the leg movements condition.

To be certain that the findings of Experiment 2 stem from the interaction between the participants' movements and the gestures they saw rather than from the specific sentences and pictures used, we again analyzed RTs for correct responses to the same "yes" sentences presented without gesture during the practice phase. These data were analyzed in a linear mixed model with participant and sentence as random effects and condition, later gesture–picture congruence, and the interaction between the two as fixed effects (number of observations = 3,055; see Table 3 for model summary). Mean RTs are displayed in Figure 4. The only significant predictor was condition (coefficient estimate =  $23.44$ ,  $SE = 10.55$ ,  $t$  value =  $2.22$ ,  $p = .0264$ )—just as in the main experimental block, RTs for the leg movements condition were longer than those for the arm movements condition. The fact that there was a significant condition effect (i.e., that RT was higher in the leg movement condition than the arm movement condition) not only in the main experimental block (Figure 3) but also in the initial nongesture block (Figure 4) suggests that this effect was not because the leg and foot movements interfered with simulating arm and hand movements (since the effect was present even on the nongesture trials). Rather, the pattern lends support to the hypothesis that the leg and foot movements require more attention, which takes time to deploy, than the arm and hand movements.

Later gesture–picture congruence was not a significant predictor (coefficient estimate =  $-6.22$ ,  $SE = 10.59$ ,  $t$  value =  $-0.59$ ,  $p = .5552$ ). Critically, neither was the interaction term (coefficient estimate =  $11.56$ ,  $SE = 14.87$ ,  $t$  value =  $0.78$ ,  $p = .4354$ ). This model did not predict the data any better than a model including the two main fixed effects and two random effects (BIC for 5-term model = 41,474; BIC for 4-term model = 41,467,  $\chi^2(1) = 0.61$ ,  $p = .4364$ ).<sup>8</sup> As in Experiment 1, we must be cautious in interpreting data from the practice phase, before participants were thoroughly familiarized with the task. Nonetheless, it is worth noting that we found different patterns of RTs for the two move-

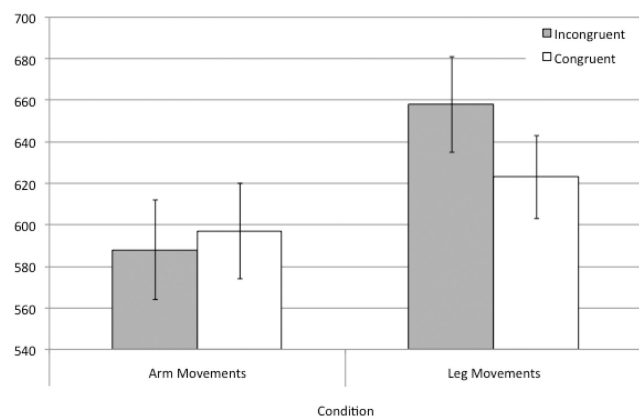


Figure 3. Mean reaction times (RTs), tallied by-participant, for the main experimental block in Experiment 2. As predicted, RTs in the leg movements condition were shorter for congruent than for incongruent trials, whereas RTs in the arm movements condition did not show a congruence effect. Error bars are standard errors of the mean.

<sup>8</sup> The separate analysis of the practice phase and main experimental block was warranted—a series of models predicting RTs from both blocks showed that the three-way interaction (Block  $\times$  Condition  $\times$  Later Gesture–Picture Congruence) term was a significant predictor (coefficient estimate =  $-49.57$ ,  $SE = 21.83$ ,  $t$  value =  $-2.27$ ,  $p = .0232$ ) and improved the fit of the model compared with a model including every term besides the three-way interaction term (BIC for nine-term model = 84,164; BIC for eight-term model = 84,167),  $\chi^2(1) = 5.16$ ,  $p = .0231$ ).



Table 3

Summary of the Random and Fixed Effects Predicting Reaction Time During the Practice Phase in Experiment 2 Before Sentence/Picture Combinations Were Paired With Gestures (N = 3,055)

Variable	Variance	Standard deviation	Coefficient estimate	Standard error	<i>t</i>	<i>p</i> (estimate)
Random effect						
Participant	9,591	97.93				
Sentence	3,159	56.21				
Fixed effect						
(Intercept)			783.37	17.31	45.26	<.0001
Condition (Leg)			23.44	10.55	2.22	.0264
Later gesture–picture relationship (Will be congruent)			−6.22	10.59	−0.59	.5552
Condition × Gesture–Picture interaction			11.56	14.87	0.78	.4354

ment conditions only when gesture was involved. When sentences were presented without gesture, there were no significant main effects or interactions as a function of sentences, pictures, and the particular movement task participants performed.

### General Discussion

Speakers often convey information in their gestures that is not conveyed in their speech (Church & Goldin-Meadow, 1986; Goldin-Meadow, 2003; Perry, Church, & Goldin-Meadow, 1988; Pine, Lufkin, & Messer, 2004). Listeners have been shown to understand and subsequently use this information, usually without being aware of its source (Goldin-Meadow & Sandhofer, 1999; McNeill et al., 1994). Even children at the early stages of language learning (Kelly, 2001; Morford & Goldin-Meadow, 1992) are able to glean information from a speaker's gestures.

In Experiment 1, we built on these studies, introducing a new experimental paradigm for exploring gesture perception. We used RT to show that listeners incorporate information that a speaker conveys in gesture into their mental models of the message conveyed in speech—participants were faster to indicate that a pic-

ture had been named in a spoken sentence when that picture was congruent with information previously conveyed in the speaker's gestures (and not speech).

In Experiment 2, we used this paradigm to show that the listener's own motor system is involved in processing the information conveyed in another's gesture. Planning and producing arm and hand movements while watching a speaker gesture interfered with the listener's ability to interpret gesture—listeners who were moving their arms did not respond more quickly to pictures that were congruent with the speaker's gestures than to pictures that were incongruent. In contrast, planning and producing leg and foot movements while watching a speaker gesture did not interfere with the listener's ability to use gestured information—listeners who were moving their legs responded more quickly to pictures that were congruent with the speaker's gestures than to pictures that were incongruent. Motor simulation of another's actions is thought to be automatic and involved in action and goal understanding, action planning, and action coordination. We suggest that motor simulation also plays a role in gesture perception and understanding.

A perception–action link has recently been established for gesture, as it has for human actions that do not serve to represent meaning (Bingham et al., 1999; Hecht et al., 2001; Wilson, Snapp-Childs, & Bingham, 2010; Zaal et al., 2000). Beilock and Goldin-Meadow (2010; see also Goldin-Meadow & Beilock, 2010) showed that a speaker's gestures can impact his or her own later actions; that is, there is a perception–action link for gesture *within* an individual. A comparable link has also been found *across* individuals—Cook and Tanenhaus (2009) found that a speaker's gestures can affect his or her *listener's* later actions. Our findings speak to the mechanism underlying these phenomena and suggest that motor simulation may play a role in the process, as it does in the perception of nonrepresentational human actions (e.g., Calvo-Merino et al., 2005; Hamilton et al., 2004; Jacobs & Shiffrar, 2005). We suggest that when listeners observe gesture, they exploit the motor simulations that automatically occur when observing human action and that this simulation has an impact on how they glean information from those gestures. Gesture is, after all, a type of human action. But it is action that does not have a direct impact on the physical world. Instead, gesture affects the world through the information it conveys—through representation. Our data show that the perception–action links found for actions

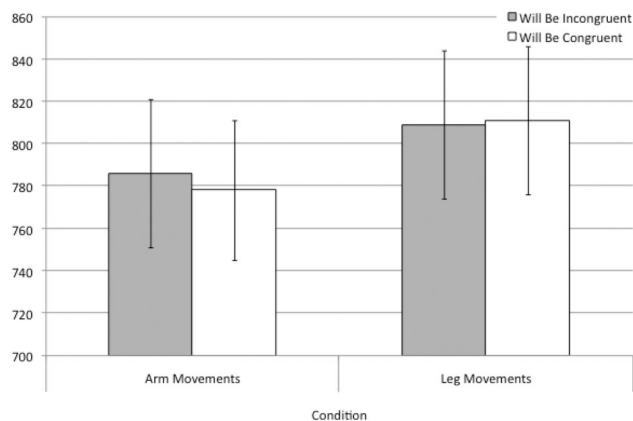


Figure 4. Mean reaction times (RTs), tallied by-participant, for the practice phase in Experiment 2. RTs were no shorter for pictures following sentences that would later be paired with congruent gesture during the main experimental block than they were for pictures following sentences that would later be paired with incongruent gesture in either condition. Error bars are standard errors of the mean.

that have a physical effect on the world extend to actions that are representational and exert their force in conversation.

Although our procedure does not tell us precisely *how* the listener's motor system is involved in the simulation process, our data do suggest that the listener's motor system *is* involved in the process. If not, then planning and producing arm and hand movements would have had no impact on the congruence effect—participants who moved their arms would show the same congruence effect as those who made no movements and as those who made leg and foot movements. Future work is needed to determine whether listeners are simulating the movements that the speaker actually makes when gesturing, or the movements that those gestures represent. One way to address this question is to examine hand gestures that represent movements of the foot; for example, wiggling the two fingers in an upside-down-V hand shape as the hand moves forward, a gesture that represents a figure walking. If listeners are simulating the movements that the gesture *represents*, then being told to plan and execute movements of their feet should disrupt processing of this particular gesture, which should, in turn, lead to the elimination of the gesture–picture congruence effect (as in the arm movement condition in Figure 3). The paradigm we have developed thus has the potential to be used to pin down the role that the listener's motor system plays when processing a speaker's gestures.

Previous work has shown that the motor system is at least partially responsible for how we comprehend information represented by language (Glenberg & Kaschak, 2002; Tucker & Ellis, 2004; Zwaan & Taylor, 2006). Motor simulation, particularly in the premotor cortex, has been found to be involved in language comprehension (e.g., Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008). In other words, understanding action concepts represented by language appears to draw on resources that one would use to plan those actions. Our findings suggest that gesture can modulate this process—seeing a representational action, that is, a gesture, while listening to speech leads to simulation of that action, which, in turn, impacts the way that the accompanying speech is processed. Hostetter and Alibali (2008; see also Hostetter & Alibali, 2010) suggested that motor simulation is at work in language *production* as well. Given the right circumstances, actual actions—in the form of representational hand gestures—emerge as a product of this simulation. Gesture as simulated action (GSA), as this account is known, posits a causal relationship between motor simulation and the production of representational actions (gestures) within a single individual. Here, we suggest a similar relationship between motor simulation and the perception of representational actions (gestures) *across* individuals—gesture perception involves online links between the speaker's gestures and activation in the listener's motor system. Automatic simulation of representational actions (gestures) may thus have a role to play in decoding the speech that accompanies those gestures.

In sum, we began by introducing a new paradigm for studying gesture perception online. We then used this paradigm to show that listeners automatically activate their own motor systems when watching a speaker's gestures, which, in turn, has an impact on their ability to use the information conveyed in gesture. Gleaning information from another's gestures seems to call upon the listener's own motor system.

## References

- Alibali, M., Flevares, L., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology, 89*, 183–193. doi:10.1037/0022-0663.89.1.183
- Beilock, S. L., & Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *Psychological Science, 21*, 1605–1610. doi:10.1177/0956797610385353
- Beilock, S. L., & Holt, L. E. (2007). Embodied preference judgments: Can likeability be driven by the motor system? *Psychological Science, 18*, 51–57. doi:10.1111/j.1467-9280.2007.01848.x
- Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., & Small, S. L. (2008). Sports experience changes the neural processing of action language. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 105*, 13269–13273. doi:10.1073/pnas.0803424105
- Bingham, G. P., Schmidt, R. C., & Zaal, F. T. J. M. (1999). Visual perception of the relative phasing of human limb movements. *Perception & Psychophysics, 61*, 246–258. doi:10.3758/BF03206886
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., . . . Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13*, 400–404.
- Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*, 1243–1249. doi:10.1093/cercor/bhi007
- Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*, 43–71. doi:10.1016/0010-0277(86)90053-3
- Cook, S. W., & Tanenhaus, M. K. (2009). Embodied understanding: Speakers' gestures affect listeners' actions. *Cognition, 113*, 98–104. doi:10.1016/j.cognition.2009.06.006
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558–565. doi:10.3758/BF03196313
- Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S., & Beilock, S. L. (2010). Action's influence on thought: The case of gesture. *Perspectives on Psychological Science, 5*, 664–674. doi:10.1177/1745691610388764
- Goldin-Meadow, S., & Sandhofer, C. M. (1999). Gesture conveys substantive information about a child's thoughts to ordinary listeners. *Developmental Science, 2*, 67–74. doi:10.1111/1467-7687.00056
- Hamilton, A., Wolpert, D. M., & Frith, U. (2004). Your own action influences how you perceive another person's action. *Current Biology, 14*, 493–498. doi:10.1016/j.cub.2004.03.007
- Hecht, H., Vogt, S., & Prinz, W. (2001). Motor learning enhances perceptual judgment: A case for action-perception transfer. *Psychological Research, 65*, 3–14. doi:10.1007/s004260000043
- Holt, L. E., & Beilock, S. L. (2006). Expertise and its embodiment: Examining the impact of sensorimotor skill expertise on the representation of action-related text. *Psychonomic Bulletin & Review, 13*, 694–701. doi:10.3758/BF03193983
- Hommel, B., Colzato, L. S., & van den Wildenberg, W. P. M. (2009). How social are task representations? *Psychological Science, 20*, 794–798. doi:10.1111/j.1467-9280.2009.02367.x
- Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review, 15*, 495–514. doi:10.3758/PBR.15.3.495
- Hostetter, A. B., & Alibali, M. W. (2010). Language, gesture, action! A test of the gesture as simulated action framework. *Journal of Memory and Language, 63*, 245–257. doi:10.1016/j.jml.2010.04.003
- Jacobs, A., & Shiffrar, M. (2005). Walking perception by walking observers. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 157–169. doi:10.1037/0096-1523.31.1.157

- Kelly, S. D. (2001). Broadening the units of analysis in communication: Speech and nonverbal behaviours in pragmatic comprehension. *Journal of Child Language*, *28*, 325–349. doi:10.1017/S0305000901004664
- Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, *15*, 99–104. doi:10.1111/j.0963-7214.2006.00415.x
- Maeda, F., Mazziotta, J., & Iacoboni, M. (2002). Transcranial magnetic stimulation studies of the human mirror neuron system. *International Congress Series*, *1232*, 889–894. doi:10.1016/S0531-5131(01)00729-4
- McNeill, D., Cassell, J., & McCullough, K.-E. (1994). Communicative effects of speech-mismatched gestures. *Research on Language and Social Interaction*, *27*, 223–237. doi:10.1207/s15327973rlsi2703\_4
- Morford, M., & Goldin-Meadow, S. (1992). Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language*, *19*, 559–580. doi:10.1017/S0305000900011569
- Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, *3*, 359–400. doi:10.1016/0885-2014(88)90021-4
- Pine, K. J., Lufkin, N., & Messer, D. (2004). More gestures than answers: Children learning about balance. *Developmental Psychology*, *40*, 1059–1067. doi:10.1037/0012-1649.40.6.1059
- Reed, C. L., & Farah, M. J. (1995). The psychological reality of the body schema: A test with normal participants. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 334–343. doi:10.1037/0096-1523.21.2.334
- Reed, C. L., & McGoldrick, J. E. (2007). Action during body perception: Processing time affects self–other correspondences. *Social Neuroscience*, *2*, 134–149. doi:10.1080/17470910701376811
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*, 70–76. doi:10.1016/j.tics.2005.12.009
- Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition*, *88*, B11–B21. doi:10.1016/S0010-0277(03)00043-X
- Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006). Twin peaks: An ERP study of action planning and control in co-acting individuals. *Journal of Cognitive Neuroscience*, *18*, 859–870. doi:10.1162/jocn.2006.18.5.859
- Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*, 153–156. doi:10.1111/1467-9280.00326
- Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta Psychologica*, *116*, 185–203. doi:10.1016/j.actpsy.2004.01.004
- Wilson, A. D., Collins, D. R., & Bingham, G. P. (2005). Perceptual coupling in rhythmic movement coordination: Stable perception leads to stable action. *Experimental Brain Research*, *164*, 517–528. doi:10.1007/s00221-005-2272-3
- Wilson, A. D., Snapp-Childs, W., & Bingham, G. P. (2010). Perceptual learning immediately yields new stable motor coordination. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 1508–1514. doi:10.1037/a0020412
- Woodward, A., Sommerville, J. A., Gerson, S., Henderson, A. M. E., & Buresh, J. (2009). The emergence of intention attribution in infancy. In B. Ross (Ed.), *The psychology of learning and motivation: Vol. 51* (pp. 187–222). Burlington, MA: Academic Press. doi:10.1016/S0079-7421(09)51006-7
- Yang, S.-J., Gallo, D. A., & Beilock, S. L. (2009). Embodied memory judgments: A case of motor fluency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1359–1365. doi:10.1037/a0016547
- Zaal, F. T. J. M., Bingham, G. P., & Schmidt, R. C. (2000). Visual perception of mean relative phase and phase variability. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1209–1220. doi:10.1037/0096-1523.26.3.1209
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*, 168–171. doi:10.1111/1467-9280.00430
- Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*, 1–11. doi:10.1037/0096-3445.135.1.1

Received March 15, 2012

Revision received November 28, 2012

Accepted January 15, 2013 ■