From Hands to Minds: Gestures Promote Action Understanding

Seokmin Kang (sk2587@columbia.edu)
Teachers College, Columbia University
New York, NY 10027 USA

Barbara Tversky (btversky@stanford.edu)
Columbia Teachers College
New York, NY 10027 USA
Stanford University
Stanford, CA 94305 USA

John B. Black (jbb21@columbia.edu)
Teachers College, Columbia University
New York, NY 10027 USA

Abstract

Understanding dynamic concepts is more difficult than understanding static ones. The present study showed that understanding dynamic concepts can be enhanced by gestures that convey action. Participants learned how an engine worked from one of two videos, with identical verbal scripts and identical diagrams. One video was accompanied by gestures showing the structure of the system; the other was accompanied by gestures showing the actions of the system. Both groups learned the basics of the system. Participants who saw the action gestures depicted more dynamic information in their visual explanations of the system and included more dynamic information in their verbal explanations of the system. Because they are inherently dynamic, gestures appear to be especially suited for conveying dynamic information.

Keywords: gesture; diagram; complex systems; knowledge construction

Knowledge in the hands

When people explain something, they typically use gestures as well as speech. Gestures can carry information that is redundant with speech, reinforcing the message by presenting information in two modalities. Importantly, gestures sometimes carry information that is not carried in speech (e. g, Bavelas, 1994; Church & Goldin-Meadow, 1986; Perry, Church, & Goldin-Meadow, 1988). In some cases, speech refers listeners to gesture, as in "turn this way," but in other cases, there is no cuing of the gestures. Nevertheless, the information carried solely in gesture can reveal the thought of speakers and affect the thought of both those who make gestures and those who watch them (e.g., Beattie & Shovelton, 1999; Chu & Kita, 2011; Goldin-Meadow, Cook, & Mitchell, 2009; Goldin-Meadow, Kim, & Singer, 1999; Hegarty, Mayer, Kriz, & Keehner, 2005; Kessell & Tversky, 2006; Singer & Goldin-Meadow, 2005; Mcgregor, Rohlfing, Bean, & Marschner, 2009; Ping & Goldin-Meadow, 2008; Schwartz & Black. 1996; Thompson, Driscoll, & Markson, 1998; Valenzeno, Alibali, & Klatzky, 2003).

It is primarily iconic and deictic gestures that reveal the thought of those who make them and affect the thought of those who make them or observe them. Deictic gestures point to places or things in the world or in a virtual world. Iconic gestures show what something looks like or acts like (e.g., McNeill, 1992; Goldin-Meadow, 2003). Together, these kinds of gestures can carry rich semantic content. A train of integrated deictic and iconic gestures can be used on virtual stages to create detailed models of situations, such as environments (e. g., Emmorey, Tversky, & Taylor, 2000) and actions, such as how a lock works (e. g., Engle, 1998). Are such gestures successful in communicating knowledge as well as in representing it?

Knowledge on the page

As such, sequences of organized gestures can serve much like diagrams. In fact, many kinds of gestures can be mapped to kinds of diagrammatic features; that is, they carry the same meanings (Tversky, Heiser, Lee, & Daniel, 2009). Diagrams have some advantages over gestures as a means of representing knowledge. Diagrams have permanence, so they can be inspected and reinspected. Because they are external and persist, they do not need to be kept in mind, so the mind is free to use the diagram as a basis for reorganization, for inference, and for discovery. Diagrams use elements and spatial relations on a page to represent elements and relations that are actually spatial, as in maps or architectural plans, or that are metaphorically spatial, as in the periodic table or organization charts (e. g., Tversky, 2011; Tversky, et al., 2009). Gestures, like language, are external, but lack permanence. A series of gestures used to create a model of a situation requires working memory to create, understand, and remember, and can tax working memory. On the other hand, diagrams are static, so it can be challenging to convey action, change, and process in diagrams. Typically, arrows are used, but they can be ambiguous (e. g., Heiser & Tversky 2006; Tversky, 2011; Tversky, Heiser, MacKenzie, Lozano, & Morrison, 2007). Gestures are by nature dynamic, so they can portray action, if schematically (e. g., Kita & Özyürek, 2003; Rizzolatti & Arbib, 1998). In fact, when gestures are used with diagrams in explanations, diagrams are often used to convey structure, and gestures to portray action (e. g., Engle, 1998).

Complex systems

Many explanations, in conversational as well as learning situations, are of complex systems, scientific, mechanical, social, athletic, or political. Complex systems typically have elements--actors or agents or object--that have properties and structure, social or geographic or other relations. They also have action or behavior: the actors or agents or objects act or are acted on in some sort of systematic ways usually associated with their properties and their relationships or structure. Many complex systems, from traffic patterns to election procedures, from spread of disease to workings of the nervous system, from the operation of an engine to a court of law, can be explained, especially when accompanied by deictic and iconic gestures. They can also be diagrammed, and, as noted, diagrams readily portray the structural relations of agents, actors, and objects, but do not easily portray the action or behavior of systems. Yet, it is the action of a system and its outcomes that is hardest for novices to comprehend (e. g., Hmelo-Silver and Pfeffer, 2004). Making inferences about action or function separates novices and experts across domains (e.g., Suwa & Tversky, 1997). Here we ask whether gestures showing action can promote understanding of the behavior of complex systems.

To ask whether iconic gestures that convey action can promote understanding of explanations of complex systems, we compared explanations that were identical except for gesture. One explanation was accompanied by gestures that portrayed action, and a control explanation used gestures to convey the form and structure of the parts of the system. Students viewed one of two videos of explanations of the operation of a four-stroke engine, the typical engine in an automobile. The language of the explanations was identical, and each explanation was based on a diagram of the structure of the engine superimposed to the front and side of the explainer. Because enactive gestures can convey action directly and information about action is more difficult, we were especially interested to know if gestures conveying action help students comprehend action information.

Performance was assessed in several ways: by questions about structure and action, by diagrams, by visual explanations, and by live explanations of the systems by the students. The questions could be answered solely on the basis of the language of the explanations and served partly as a manipulations check. Hence, if students who view action gestures have a better understanding of the action of the system than those who viewed structure gestures, they should be more likely to include action information, in their diagrams, and they should be more likely to deliver action information and use action gestures themselves in their later explanations to new learners.

Method

Participants 59 (15 male) university students ranging in age from 20 to 36 with an average age of 26 ($SD = 3.50$), participated in the study. They were all native English speakers and did not have prior knowledge of the system to be learned.

Materials We created two videos explaining how a four-stroke engine works. The videos were identical in language and number of gestures but differed in kinds of gesture. A diagram typical of those in science and engineering showing the labeled parts and configuration of the system was superimposed in front and to the side of the explainer. The explanations began with an introduction overviewing the structure using deictic gestures. The core portion of the explanation was a step-by-step explanation of the processes comprising the workings of the system. The final portion of the explanation explained how the process caused the car's wheels to rotate. Because the diagram showing the structure was always in view and because the introduction to both explanations overviewed the system structure, the gestures emphasizing structure served as a control and were not expected to affect performance on the questions.

For the core portion of the explanation, in the action video, the explainer used only gestures that portrayed the action of each part, always in the same location, so no structural information was provided. In the control structure video, the explainer used only gestures that pointed to the location of the parts of the system and showed the shape of each part as the process was explained. The accompanying verbal script explained both the locations of the parts and the actions of the parts identically. Figure 1 shows snapshots of two instructional videos.
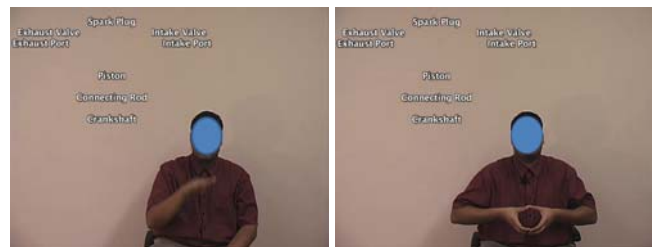

Figure 1. Still shots from the *action* (left) and *structure* (right) videos showing the superimposed diagrams.

The information in the script was categorized as structure or action, and gestures appropriate for each were devised. For the action gesture video, the explainer showed the rotational motion of the crankshaft, the direction of the piston's movement, the flow of fuel and air, the movement of the intake and exhaust valves, and so on with his hands. The action gestures were performed in the same place off the diagram, avoiding any positional information.

For the structure gesture video, the explainer used his hand(s) successively to show the shapes of the crankshaft, piston and cylinder, and showed the positions of the piston, crankshaft, spark plug, intake port, intake value, exhaust port, exhaust valve, and mixture of fuel and air.

To eliminate any biasing effects of lexical stress (Heuven, 1988; Field, 2005), the speaker practiced the script several times, making sure to stress the actions and the parts for both videos.

Posttests The verbal posttest was based on the information in the script with 20 recognition questions, 16 True/False, and 4 multiple-choice questions. Of the 16 True/False

questions, 8 queried action and 8 queried structure. Action questions referred to movement, or causal relations of the parts and their consequences. Structure questions referred to shapes and positions of the parts of the system. Four multiple-choice questions queried general knowledge. The questions were presented in random order. Because the verbal posttest was based entirely on the verbal script, differences dependent on viewed gesture were not expected. The test served as a manipulation check.

The second posttest was a diagraming task. Participants were asked to diagram a visual explanation of how a four stroke engine works based on what they learned from the video. Finally, participants made a video to explain the workings of the four-stroke engine to a peer. It was expected that participants who viewed the videos with action gestures would include more action information in the latter two less-constrained measures.

Procedure Participants were seated at a table with a laptop computer with a 15.4 in screen. They were randomly assigned to either the action gesture or the structure gesture video group. The participants were then told: "Today, your job is to watch a video of how a four stroke engine works four times[1] in a row and explain the concept in the video to a peer coming later. However, since you are not directly explaining a concept, your explanation will be videotaped and showed later either to him or her. He or she will learn about the concept from your explanation." Participants were not allowed to take notes or to pause or stop the video. The experimenter left the room while participants watched the video. After watching the video, participants were given the verbal and diagrammatic posttests, and then made a video explaining the system to a peer. The video camera was set opposite the participant 3 meters away. Participants were allowed to spend as much time as they wanted.

Results

Verbal Posttest As expected, the type of gesture viewed yielded no differences in performance on action ($p = .08$), structure ($p = .85$) or general ($p = .92$) questions, nor were there interactions between gesture viewed and question type, $F(1, 114) = 1.70$, $p = .20$. However, in within group comparisons, those who viewed action gestures performed better on action questions than on structure questions, $t(28) = 3.56$, $p < .01$, $d = 0.82$. There were no differences between action and structure questions for those who viewed gestures conveying structure ($p = .11$).

Diagram Posttest Two coders coded the diagrams for action components. The reliability for action words was $Kappa = .56$ ($p < .001$), for action arrows, $Kappa = .63$ ($p < .001$), for action effects, $Kappa = .65$ ($p < .001$), for labeling arrows, $Kappa = .60$ ($p < .001$), and for labeling lines, $Kappa = .73$ ($p < .001$). Action effects were depictions of actions, such as explosions. The means of the visual components by type of viewed gesture appear in Figure 2.

---

[1] A pilot study had revealed that two viewings were insufficient to achieve above chance performance.

The means were compared using Poisson regression analysis with the assumption that the conditional means equal the conditional variances.
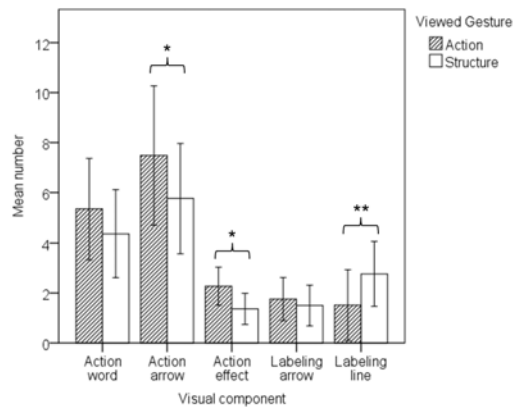

Figure 2. Mean number of visual components in diagrams. Error bars represent standard errors of the means.

Overall, those who viewed action gestures used more visual components than those who saw structure gestures ($p < .05$). In addition, those who viewed action gestures produced more action arrows ($p < .05$) and action effects ($p < .05$) and labeled fewer lines ($p < .01$) than those who saw structure gestures. Labeled lines typically linked names and parts; that is, structural information. Thus, for the diagrams, those who saw action gestures included more information about action and those who saw structure gestures included more information about structure, showing that the viewed gestures affected viewers' comprehension and later production.

Explanations to a peer Recall that after learning the system, participants made videos explaining the four-stroke engine to novices. Will those who saw action gestures use more of them in their own explanations? A gesture unit was defined as "the period of time between successive rests of the limbs (McNeill, 1992)." If the hands did not return to a resting position between two gestures, the boundary was defined by a pause in motion and an obvious change in shape or trajectory. If participants used both hands simultaneously to describe one object, concept, or part, it was regarded as one gesture. If participants used one hand to describe an object, a concept, or a part and the other hand a different concept, the gestures were coded as two different gestures.

For this study, only gestures conveying action or structure were coded. Action gestures were defined as showing the action of a part or process of a system. Structure gestures were defined as showing the location or static properties, notably shape, of objects or parts of the system. Inter-rater reliability was assessed on randomly selected 240 subsets (18%) of the data by a second coder who was trained and blind to the experimental design. Agreement for identifying gestures was 87.8% and for categorizing gestures was 99.6%.

For the speech analysis, the participants' verbal descriptions were segmented into propositions (following

Heiser & Tversky, 2006). The information units were coded as *action, structure*, or *other*. Propositions that contained action such as movement of each part within a cylinder were coded as *action* information, for example, *"...that byproduct is pushed back up through the exhaust valve..."*. Propositions that contained 'is-a' or 'has-a' were coded as *structure* information unless they referred to action, for example, *"...then it has an exhaust valve"*. *Other* information included greetings, such as *"Good evening,"* introductory information such as *"I'm going to explain how a four stroke engine works,"* and meta-comments such as *"...let me tell you a little bit more about each stage..."*

Gesture analysis[2] The average explanation time was 177.14 sec (*SD* = 56.84) for the action group and 152.34 sec (*SD* = 55.94) for the structure group (ns. *p* = .10). There were a total of 1306 gestures: 754 by those who had viewed action gestures, 552 by those who had viewed structure gestures (ns. *p* = .13). The means of action and structure gestures produced by participants who viewed action and structure videos are shown in Figure 3.

There was an interaction between type of gesture viewed and type of gesture produced, $F(1,112) = 8.58$, $p = .004 < .01$, $\delta = .84$. In within group comparison by paired sample t-test, even though participants in both groups delivered more action gestures than structure gestures, the action group ($t(28) = 7.15$, $p < .0001$, $d = 1.49$, $r = .60$) reliably used more action gestures, when compared to the structure group ($t(28) = 2.88$, $p = .008 < .01$, $d = 0.58$, $r = .28$).
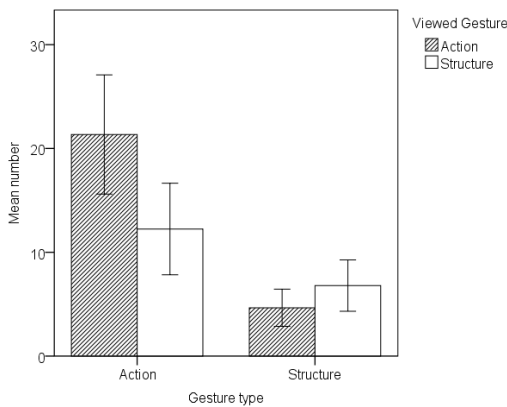

Figure 3. Mean number of type of produced gesture by type of viewed gesture. Error bars represent standard errors of the means.

It is possible that the number and the pattern of gestures differed by length of explanation. Although there were no significant differences in the overall gesture use, explanations in the action group were longer. Consequently, the next section presents more detailed analyses of the results by gesture rate, that is, gestures per minute.

[2] One participant's explanation was not recorded because of malfunctioning of a video camera. Therefore, 58 participants' videos were analyzed.

The same pattern of gesture use was observed. The action group used 6.89 (*SD* = 4.13) action gestures per minute and 1.45 (*SD* = 1.35) structure gesture per minute. The structure group used 4.62 (*SD* = 3.49) action gestures per minute and 2.51 (*SD* = 2.38) structure gesture per minute.

In group comparison, there was an interaction such that the action group used more action gestures and the structure group used relatively more structure gestures, $F(1,112) = 8.83$, $p = .004$, $< .01$, $\delta = .84$. In within group comparison, when compared to the structure group ($t(28) = 3.08$, $p = .005 < .01$, $d = 0.71$, $r = .33$), the action group ($t(28) = 7.95$, $p < .0001$, $d = 1.77$, $r = .66$) reliably used more action gestures than structure gestures.

Speech analysis The participants delivered a total of 2550 information units in their speech. Among them, 1607 conveyed *action* information, 737 *structure* information, and 206 *other* information. Those who saw action gestures delivered a total of 1425 information units. Among them, 929 conveyed action, 387 conveyed structure, and 109 conveyed other information. Those who saw structure gestures delivered a total of 1125 information units, 678 conveying action, 350 conveying structure and, 97 conveying other information. Figure 4 shows mean number of information units delivered by two groups.

Overall, those who viewed action gestures (*M* = 49.14, *SD* = 20.81) delivered more information units than those who viewed structure gestures (*M* = 38.79, *SD* = 17.22), $F(1,56) = 4.25$, *MSE* = 364.75 , $p = .044 < .05$. In addition, those who viewed action gestures delivered more action information than the structure group, $F(1,56) = 6.87$, *MSE* = 158.03, $p = .01 < .05$. There were no differences in the quantity of structural information ($p = .52$) or other information ($p = .66$), but there was an interaction between kind of information and kind of viewed gesture ($p = .02 < .05$). Post hac tests (Tukey HSD) showed that more action information was given than structural information ($p < .001$) and more structure information than other information ($p < .001$).
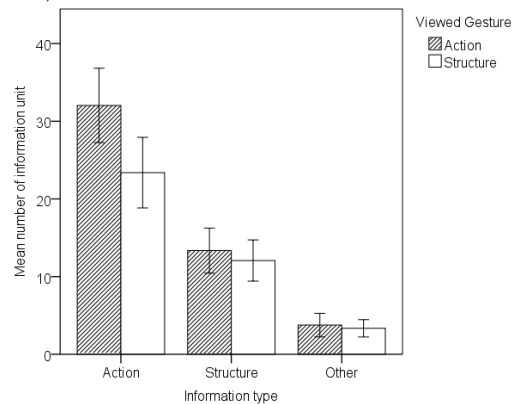

Figure 4. Mean kinds of information units by viewed gestures. Error bars represent standard errors of the means.

Proportion of information types in speech Although there were no group differences in explanation time, the group

who had viewed action gestures took more time and delivered more information units than the group who viewed structure gestures. To take that into account, percentages of information types were analyzed and appear in Figure 5.
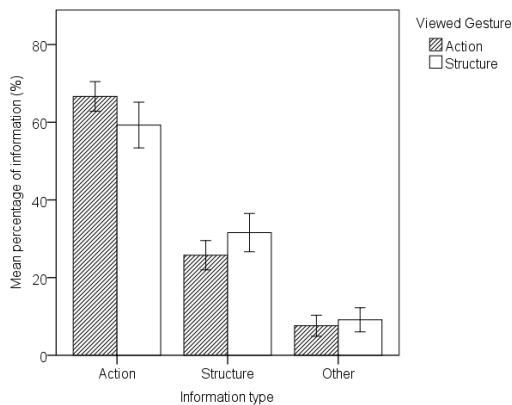

Figure 5. Mean percentage of kinds of information units by viewed gesture. Error bars represent standard errors of the means.

For those who viewed action gestures, action information accounted for an average of 66.62% (SD = 10.30), structure information accounted for an average of 25.76% (SD = 10.13), and other information accounted for an average of 7.61% (SD = 7.30). For those who viewed structure gestures, 59.28% (SD = 15.89) was action information, 31.59% (SD = 13.28) was structure information, and 9.14% (SD = 8.34) was other information. There was an interaction between group and information type, $F(2,168)$ = 5.16, MSE = 126.74, p = .007 < .01. Those who viewed action gestures delivered relatively more action information than those who viewed structure gestures and those who viewed structure delivered relatively more structure information than those who viewed action group. Thus, in their own explanations, those who had viewed action gestures produced both more verbal information about action and showed more action in their gestures. Similarly, those who had viewed structure gestures used more structure gestures and included proportionately more verbal structure information than those who had viewed action gestures.

Discussion

Understanding the behavior of complex systems is challenging (e. g., Hmelo-Silver & Pfeffer, 2004). Actions are not apparent in static diagrams, and the nature of actions often has to be imagined from purely symbolic language. Animations are typically too complex and too fleeting to be comprehended (e. g., Tversky, Morrison, & Betrancourt, 2002) and are not part of most natural settings for explanations. There is abundant evidence that gestures provide a rich source of information, including information about structure and process (e. g., Beattie et al., 1999; Becvar, Hollan, & Hutchins, 2008). Here we asked if gestures can successfully transmit dynamic information, over and above verbal and diagrammatic explanations,

simply and abstractly at a pace that allows comprehension. We taught a complex system, the operation of a four-stroke engine, to novices under two conditions. One group saw action gestures that conveyed the behaviors of the parts of the system; the other group saw structure gestures that conveyed static qualities of the parts of the system and their structure. Both groups heard exactly the same explanation and saw the same structure diagram of the parts of the system. The verbal explanation was sufficient to convey the basics of the structure and the dynamics of the engine. A number of posttests were administered: a verbal test based on the verbal explanation, a visual explanation task, and a videotaped explanation of the system to new novices.

The verbal memory test showed that both groups adequately learned the essentials of the structure and the operation of the system. However, the diagramming and the explanation tasks revealed substantial differences in the understanding of the behavior of the systems; the group who had viewed the action gestures appeared to have a deeper understanding of the behavior of the system than the group who had viewed the structure gestures. In the visual explanation task, those who had seen action gestures depicted more specific actions of the system than the group who had viewed the structure gestures. Furthermore, the group who had viewed the action gestures used more action gestures in their videoed explanations than the group who had viewed the structure gestures. Although the increase in the number of action gestures in explanations might be attributable at least in part to imitation of what they had viewed, the increase in number of depictions of specific actions cannot. The depictions of action must come from a deeper understanding of the specific chain of behaviors of the system. Moreover, many of the gestures used differed from those viewed.

The effects of the viewing the gestures that conveyed the structure of the system were weaker but evident both in diagrams and in explanations. The structure of the system was apparent from the diagram that was displayed during the viewed explanation, and the structure of the system was described in the verbal portion of the explanation. Furthermore, the structural information is easier than the behavioral because it was apparent in the diagram.

In both groups, gestures conveying action far outnumbered gestures conveying structure, suggesting that participants regarded the behavior of the system as paramount and regarded gesture as a good means for conveying action, over and above language.

Discourse in the wild, including explanations, is an integrated combination of word, gesture, and props, elements in the world (such as a diagram) or in a virtual world that can be continuously referred to during the course of the discourse. Each, word, gesture, prop, plays roles, sometimes overlapping, sometimes complementary. Understandably, actions, even miniature schematic ones as those in gestures, appear to be especially effective for conveying action, another example of cognitive congruence (e. g., Tversky, et al., 2002).

References

Bavelas, J. B. (1994). Gestures as part of speech: Methodological implications. *Research on Language and Social Interaction, 27*, 201-221.

Beattie, G., & Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica, 123,* 1-30.

Becvar, A., Hollan, J., & Hutchins, E. (2008). Representational gestures as cognitive artifacts for developing theories in a scientific laboratory. In M. S. Ackerman, C. A. Halverson, T. Erickson, & W. A. Kellogg (Eds.) *Resources, Co-Evolution and Artifacts: Theory in CSCW* (pp. 117-143). London, England: Springer-Verlag.

Chu, M. & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology, 140*, 102-116.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*, 43-71.

Emmorey, K., Tversky, B., & Taylor, H. (2000). Using space to describe space: Perspective in speech, sign, and gesture. *Journal of Spatial Cognition and Computation, 2*, 157-180.

Engle, R. A. (1998). Not channels but composite signals: Speech, gesture, diagrams, and object demonstrations are integrated in multimodal explanations. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 321-326). Mahwah, NJ: Erlbaum.

Field, J. (2005). Intelligibility and the Listener: The Role of Lexical Stress. *TESOL Quarterly, 39*, 399-423.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge: Belknap Press.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science, 20,* 267– 272.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology, 91*, 720-730.

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science, 1,* 127–138.

Hegarty, M., Mayer, S., Kriz, S., & Keehner, M. (2005). The role of gestures in mental animation. *Spatial Cognition and Computation, 5*, 333-356.

Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science, 30*, 581-592.

Heuven, V. J. van (1988). Effects of stress and accent on the human recognition of word fragments in spoken context: gating and shadowing, *Proceedings of the 7th FASE/Speech-88 Symposium*, edited by W.A. Ainsworth and J.N. Holmes, 811-818. Edinburgh: Institute of Acoustics.

Kessell, A. M. and Tversky, B. (2006). Gestures for thinking and explaining. *Proceedings of the meetings of the Cognitive Science Society*.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Cognition, 48*, 16-32.

McGregor, K. K., Rohlfing, K. J., Bean, A., & Marschner, E. (2009). *Journal of Child Language, 36*, 807-828.

McNeill, D. (1992). *Hand and mind*. Chicago: University of Chicago Press.

Perry, M., Church, R. B. & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359-400.

Ping, R., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology, 44,* 1277-1287.

Rizzolatti, G., & Arbib, M. A. (1998) Language within our grasp, *Trends in Neurosciences*, 21, 188-194

Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules. *Cognitive Science, 20*, 457-497.

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teachers' gestures and speech differ. *Psychological Science, 16*, 85-89.

Suwa, M., & Tversky, B. (1997). What architects and students perceive in their sketches: A protocol analysis. *Design Studies, 18*, 385-403.

Thompson, L. A., Driscoll, D., & Markson, L. (1998). Memory for visual-spoken language in children and adults. *Journal of Nonverbal Behavior, 22*, 167-187.

Tversky, B. (2011). Visualizations of thought. *Topics in Cognitive* Science, 3, 499-535.

Tversky, B., Heiser, J., Lee, P., & Daniel, M.P. (2009). Explanations in gesture, diagram, and word. In Coventry, K. R., Tenbrink. T., & Bateman, J. (Eds.), Spatial Language and Dialogue. Oxford: Oxford University Press.

Tversky, B., Heiser, J., MacKenzie, R., Lozano, S., and Morrison, J. B. (2007). Enriching animations. In R. Lowe and W. Schnotz, *Learning with animation: Research implications for design.* NY: Cambridge University Press.

Tversky, B., Morrison, J. B. & Betrancourt, M (2002). Animation: Can it facilitate? *International Journal of Human Computer Studies. International Journal of Human Computer Studies*, 57, 247-262.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28*, 187-204.