CHAPTER 17

Understanding gesture as representational action
A functional account of how action and gesture differ with respect to thinking and learning

Miriam A. Novack and Susan Goldin-Meadow
Northwestern University / University of Chicago

A great deal of attention has recently been paid to gesture and its effects on thinking and learning. This chapter sets forth a theoretical framework for exploring why gesture serves the functions that it does. The framework distinguishes gestures, which are representational actions, from instrumental actions, which interact directly with objects and cause physical changes to the world. The theory proposes that gesture’s status as representational action is what best explains its functions with respect to thinking and learning. Most notably, because gestures are abstracted representations and are not actions tied to particular events and objects, they can play a powerful role in thinking and learning beyond the particular, specifically, in supporting generalization and transfer of knowledge.

Decades of work has shown that hand gestures affect the minds of both the people who see them and the people who produce them (Goldin-Meadow, 2003). In this chapter, we provide a framework for understanding the widespread functions of gestures on cognition, both for producers and perceivers of these hand movements. Specifically, we propose that gestures produce effects on thinking and learning because they are
gestures, including representational gestures (iconics, metaphorics), deictic gestures (points), and even beat gestures (rhythmic movements closely coordinated with speech).

Gesture is
But not all movements off objects meet the criteria for being a gesture. Some empty-handed movements are produced for the purpose of movement itself—movements like dancing or exercising. According to Schachner and Carey (2013), adults view the goal of an action to be the movement itself (i.e., movement for the sake of movement) if the movement is irrational (e.g., moving toward an object and then away from it without explanation) or if it is produced in the absence of objects (e.g., making the same to-and-fro movements but without any objects present). Using the same example, the goal of moving one’s hands up and down in the air when no hammer and no nail are present might be to just exercise the wrist and elbow.

So where does gesture fit in? Gestures can
described the event in terms of external goals (e.g., “the person placed balls in boxes”). In contrast, participants never gave this type of response for either of the empty-handed videos (i.e., videos in which the actor did not touch the objects), revealing a stark difference between how adults see actions on and off objects. But the two types of empty-handed movements were also distinguished. Both the presence of objects and the presence of filtered speech affected whether an empty-handed movement was viewed as a meaningless movement or as a meaningful representations of something else – as a gesture. More specifically, participants were likely to describe the scene as movement for its own sake when the hand movements were produced in the absence of objects or without speech (e.g., “she waved her hands back and forth above some boxes”). In contrast, they were likely to describe the scene as representational movement when these contextual cues were present (e.g., “she showed how to sort objects”). In other words, when the objects were present but not touched, and when the movements were produced along with speech-like sounds, participants were likely to see beyond the low-level features of the movement and view it as a gesture.

Observers thus made a clear distinction between the instrumental object-directed action and the two empty-handed movements (movements in the presence of objects and movements in the absence of objects), indicating that actions on objects have clear external goals, and actions off objects do not. Importantly, however, they also made a distinction between empty-handed movements that are produced for their own sake and empty-handed movements that have another function. If the conditions are right, observers go beyond the movements they see to make rich inferences about what those movements can
temporally synchronized (Loehr, 2007; McNeill, 1992) and seamlessly integrated with speech in both production (e.g., Bernardis & Gentilucci, 2006; Kendon, 1980; Kita & Ozyurek, 2003) and comprehension (e.g., Kelly, Ozyurek, & Maris, 2010), supporting the claim that speech and gesture form a single integrated system (McNeill, 1992). Gestures and speech also mutually influence each other’s meaning. For example, a circular tracing gesture might refer to the movement of a toy train when accompanied by the sentence, “the train moved along the track,” but to an army closing in on its enemy when accompanied by the sentence, “the troops surrounded the fortress.” Conversely, the gestures that accompany speech can influence the meaning taken from speech. For example, the sentence, “the train traveled along the track” is likely to describe a circular track when accompanied by a round tracing gesture, but a straight track when accompanied by a straight moving point.

**Doing gesture**

In production, gesture is spontaneous and temporally linked to speech (Loehr, 2007; McNeill, 1992) – a relation not found between speech and instrumental action. For example, adults’ movements will be more tightly bound to their speech if they explain how to sweep the floor using just their hands with no object (a gesture), than if they are asked to explain how to sweep the floor using a broom in front of them (an instrumental action) (Church, Kelly, & Holcombe, 2014). The act of producing representational gesture along with speech has been found to have an effect on speakers themselves. For example, gesturing while explaining a concept reduces a speaker’s cognitive load (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Wagner, Nusbaum, & Goldin-Meadow, 2004; Ping & Goldin-Meadow, 2010). But producing movements that are not gestures, such as meaningless hand movements, does not have the same load-lightening effects on the speaker as gestures do (Cook, Yip, & Goldin-Meadow, 2012).

**Seeing gesture**

Listeners, too are influenced by the gestures a speaker produces in a communicative context. Seeing gestures increases listeners’ comprehension (Hostetter, 2011) and improves listeners’ mental imagery (Driskell & Radtke, 2003). Gesture can be particularly beneficial for helping young children understand the content of complex spoken information (McNeil, Alibali, & Evans, 2000). Moreover, the effects of seeing gesture are not the same as the effects of seeing instrumental action. For example, adults can easily ignore actions that are incongruent with the speech with which they are produced, but they have difficulty ignoring gestures that are incongruent with the speech they accompany, suggesting a difference in the relative strength of speech-gesture integration vs. speech-action integration.
(Kelly, Healy, Özyürek, & Holler, 2015). Thus, gesture has a different relationship to speech than instrumental action does and, in turn, has a different effect on communication than instrumental action.

Problem solving

Gesture not only has an impact on communication, but it also plays a role in complex cognitive processes, such as conceptualization and problem solving. For example, both doing and seeing gestures have been found to influence how people solve the Tower of Hanoi task.

Doing gesture

In the Tower of Hanoi task, participants are asked to move a number of disks, stacked from largest to smallest, from one peg to another peg; the goal is to re-create the stacked arrangement without ever placing a larger disk on top of a smaller disk, and moving only one disk at a time. Solving the task involves actions (i.e., moving the disks), and the gestures that participants use to later explain their solution represent elements of the actions they produced while solving the task in the first place. That is, participants produce grasping gestures with curved trajectories if they solved the problem with physical objects, but flatter gestures if they solved the task using a computer program in which disk icons could be dragged across the screen using a mouse curser (Cook & Tanenhaus, 2009). Gestures thus reflect a speaker’s action experiences in the world by re-presenting traces of those actions.

More than just reflecting a speaker’s actions, producing gesture can also affect a speaker’s internal representations. For example, in another version of the Tower of Hanoi task, after explaining how they solved the task and gesturing while doing so, participants were surreptitiously given a new stack of disks that looked like the original stack but had different weights – the smallest disk was made so heavy that it could no longer be lifted with a single hand (Beilock & Goldin-Meadow, 2010; Goldin-Meadow & Beilock, 2010). Participants who had gestured while explaining their solution, and incidentally used one-handed gestures to describe how they moved the smallest disk, were adversely affected by the switch in weights – it took them longer to solve the task after the switch, when the smallest disk was now too heavy to be lifted with a single hand. By gesturing about the smallest disk with one hand, participants set themselves up to think of the disk as light – the unanticipated switch in disk weights violated this expectation, leading to relatively poor performance after the switch. Importantly, if participants did not gesture before the switch, or were asked to physically move the disks again while explaining their solution (instead of gesturing), they also do not show the switch effect (Trofatter,
Kontra, Beilock, & Goldin-Meadow, 2015). Doing gesture can thus have an effect (in this case, a detrimental effect) on thinking, and it can have a more powerful effect on thinking than action does.

**Seeing gesture**
The Tower of Hanoi task has also been used to demonstrate the impact that seeing gesture has on the listener’s conceptualizations. As just mentioned, participants gesture differently based on how they solved the Tower of Hanoi task, producing flatter arches to represent the movement of the disks if they solved the task on a computer than if they solved the task with actual disks (Cook & Tanenhaus, 2009). It turns out that listeners’ conceptualization of the problem are influenced by seeing those gestures. Participants who watched someone explain how to solve the Tower of Hanoi task using gestures with high arches were more likely to produce higher arching movements themselves on the computer (even though it is not necessary to arch the movement at all on the computer) than participants who saw someone use gestures with smaller arches. The gestures we see can influence our own actions.

**Learning**
Gesture can lead learners to new ideas or concepts, both when learners see gesture in instruction and when they do gesture themselves. Learners are more likely to profit from a lesson in which the teacher gestures than from a lesson in which the teacher does not gesture (Cook, Duffy, & Fenn, 2013; Ping & Goldin-Meadow, 2008; Singer & Goldin-Meadow, 2005; Valenzano, Alibali, & Klatzky, 2003). And when children gesture themselves, they are particularly likely to discover new ideas (Goldin-Meadow, Cook, & Mitchell, 2009), retain those ideas (Cook, Mitchell, & Goldin-Meadow, 2008), and generalize the ideas to novel problem types (Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014). We argue that gesture can play this type of role in learning because it is an action and thus engages the motor system, but also because it represents information.

**Doing gesture**
Producing gesture influences learning throughout development and begins quite early. Infants begin to point between 9 and 12 months, even before they say their first words (Bates, Camaioni, & Volterra, 1975). Importantly, these first gesture forms can provide an early window onto language learning. For example, lexical items referring to particular objects first appear in a child’s verbal repertoire several months after the child produces pointing gestures for those objects (Iverson &
Goldin-Meadow, 2005). The act of pointing may even play a causal role in the development of word-learning: 18-month-old children given pointing training (i.e., they were taught to point to pictures of objects as an experimenter named them) increased their own pointing in spontaneous interactions with their caregivers, which, in turn, led to increases in their spoken vocabulary, relative to children who did not receive pointing training (LeBarton, Goldin-Meadow, & Raudenbush, 2015). Importantly, these language-learning effects are unique to pointing gestures, and do not arise in response to similar-looking instrumental actions like reaches: 18-month-old children are likely to learn a novel label for an object if an experimenter says the label while the child is pointing at the object, but not if the experimenter says the label while the child is reaching toward the object (Lucca & Wilbourn, 2016). Thus, as early as 18 months, we see that the representational status of the pointing gesture can have a unique effect on learning (here, language learning), an effect not found for a comparable instrumental act (reaching).

The effects of producing gesture on learning continue into the school-aged years. For example, when students are encouraged to gesture as they explain their answers to math problems before receiving a lesson on those problems, they are more likely to convey new ideas in gesture (ideas that are not conveyed in speech), which subsequently makes them more likely to learn from the lesson (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007). Asking children to produce gestures that represent specific problem solving strategies makes children likely to integrate those ideas into their understanding of how to solve a problem (Goldin-Meadow, et al., 2009) and retain information weeks after instruction (Cook et al., 2008).

Recent work suggests that learning via producing gesture may function by engaging a similar motor network as learning via producing action. Children who were taught how to solve mathematical equivalence problems while producing gesture strategies later showed greater activation in motor regions when passively solving the types of problems they had been taught than children who learned to solve the problems without gesture (Wakefield et al., 2017). The same motor regions have been implicated in studies looking at the effect of producing action on learning (e.g., James 2010; James & Atwood, 2009; James & Swain, 2011), suggesting that gesture and action are similar in the effect they have on the brain.

But gestures also differ from actions in that they do not physically interact with objects. This physical distance between gestures and objects might actually be useful in a learning situation. Learning through actions can encourage learners to focus on the object itself rather than on its symbolic meaning (e.g., Uttal, Scudder, & DeLoache, 1997). The perceptual features of objects can be distracting (McNeil, Uttal, Jarvin, & Sternberg, 2009), and young children in particular may lose track of the fact that the objects they are manipulating not only are objects, but also stand for something else (DeLoache, 1995). Gesture has the potential to distance
learners from the concrete details of a manipulative, thus encouraging them to approach the concept at a deeper level. And gesture seems to take advantage of this potential – children who are taught to produce gestures during a math lesson are more likely to generalize the learned information than children who are taught to produce actions on objects during the lesson (Novack et al., 2014).

Seeing gesture

Seeing gestures also supports learning in unique ways, and starts early in development. Children begin to understand other’s pointing gestures around the same age as they themselves begin to point. At 12 months, infants see points as goal directed (Woodward & Guajardo, 2002) and recognize the communicative function of other’s points (Behne, Liszkowski, Carpenter, & Tomasello, 2012). Infants even understand that pointing hands, but not non-pointing fists, communicate information to those who can see them (Krehm, Onishi, & Vouloumanos, 2014).

Importantly, seeing pointing gestures also results in effects that are not found for similar-looking instrumental actions. For example, Yoon, Johnson, and Csibra, (2008) found that 9-month-old children are likely to remember the identity of an object if someone points to the object, but not if someone reaches to it (they remember the location of the object that is reached toward). Thus, as soon as children begin to understand pointing gestures, they seem to understand them as representational actions rather than as instrumental actions.

Around the second year of life children’s comprehension of gesture grows to include iconic gestures. For example, seeing iconic gestures helps 2-year-old children learn labels for objects (Capone & McGregor, 2005), novel spatial terms (McGregor, Rohlfing, Bean, & Marschiner, 2009) and novel verbs (Goodrich & Hudson Kam, 2009). Children can even learn about the function of a novel toy by watching a gesture demonstration at age 2 (Novack, Goldin-Meadow, & Woodward, 2015). However, at this young age, children’s ability to learn from gesture is not yet robust, and lags behind their ability to learn from other kinds of non-representational actions. That is, 2-year-old children are much more likely to learn from an action demonstration (even an incomplete action) than a gesture demonstration, a learning difference that disappears by age 3. This finding suggests that learning from gestures requires a set of processing skills above and beyond what it takes to learn from actions.

As children develop, seeing gesture continues to serve an important learning function in more formal educational contexts (e.g., Cook et al., 2013; Ping & Goldin-Meadow, 2008; Singer & Goldin-Meadow, 2005; Valenzeno et al., 2003). Seeing gestures can help students connect abstract ideas, often presented in speech, to the concrete physical environment (Valenzeno et al., 2003). Seeing gesture can also support learning by encouraging listeners to engage their own motor systems.
(Ping, Goldin-Meadow, & Beilock, 2014). Students learn more foreign words if they are taught those words while seeing someone produce meaningful iconic gestures, compared to seeing someone produce meaningless movements (Macedonia, Muller, & Friederici, 2011). Those students then activate areas of their premotor cortex when later recognizing words initially learned while seeing gesture, implicating the motor cortex in learning from seeing gesture.

Importantly, seeing gesture in instruction does not increase learning simply through attentional mechanisms alone (Novack, Wakefield, Congdon, Franconeri, & Goldin-Meadow, 2016). Rather, seeing gesture seems to help students better integrate multimodal information. For example, students are more likely to learn if their teacher produces different yet complementary problem-solving strategies, one in speech and the other in gesture, compared to the same two strategies produced entirely in speech (Singer & Goldin-Meadow, 2005). It may be gesture’s ability to be produced simultaneously with speech that promotes learning. Children are more likely to learn from a math lesson if the gesture strategy and the speech strategy occur at the same time (S1+G2) than if the speech strategy occurs first, followed by the gesture strategy (S1→G2). In other words, the benefit of speech+gesture instruction disappears when the two strategies are presented sequentially rather than simultaneously in time (Congdon et al., 2017). A question for future work is whether learning through action will also be affected by timing – that is, will learning differ when an action problem-solving strategy is presented simultaneously with speech, compared to when the same action strategy is presented sequentially with speech? We suspect that this is yet another area where learning via gesture will differ from learning via action.

Part 3. What’s next?

In this chapter, we have discussed how people identify movements as actions or gestures, and how actions and gestures play out in cognitive processes, specifically with respect to communication, problem solving, and learning. We now turn to open questions for future research, questions that revolve around understanding the similarities and differences between gestures and actions on objects, and their effect on cognitive functions.

First, we began by showing that adults can use minimal cues to distinguish between actions and gestures, and that contextual features, such as the contact between a hand and object or the form of the hand itself, can lead observers to see a movement as an action, gesture, or movement for its own sake. However, it would be useful to better understand the actual aspects of the motor act that lead observers to draw these conclusions. Do gesturing hands make the same sized
movements, or movements of the same shape, as acting hands or dancing hands? There is reason to think that these three different types of movement may have distinct movement signatures. For example, actions depend on the affordances of the objects with which they interact. We have to use a precision grip to grab a cheerio; without such a grip, the cheerio isn’t getting picked up (which is why cheerios are such good motor practice for babies). Gestures, on the other hand, do not need to be exact replicas of the actions to which they refer. We can gesture about grabbing a cheerio using a precision grip, but we can also just show the trajectory of the moving cheerio without specifying the grip. Or we could highlight the small size of the cheerio with the gestures we produce as we talk about picking it up. Gestures may have fundamentally different features from actions and, perhaps as a result, different functional effects on cognitive processes. New methodologies, such as continuous measures that allow us to quantify movement in three-dimensional space (see Hilliard & Cook, 2015, for an explanation of this method), offer new ways to classify and understand how actions, gestures, and movement for its own sake differ.

Second, we reviewed evidence showing that, from a very young age, children respond differently to actions and gestures. These developmental findings raise a number of questions about how children arrive at a mature state of gesture understanding. How do children learn that hands provide important information? Fausey, Jayaraman, and Smith (2016) recently showed that infants’ visual fields are dominated by faces first and, only later in development, by hands. Once children do begin to notice gesturing hands, what do they think of them? Do they initially see gestures as movements for their own sake – as mere handwaving (cf. Schachner & Carey, 2013)? Or, even before infants can correctly understand what a gesture means, are they able to categorize the gesture as a representational act? Knowing what infants think about gesture (and whether they categorize it as a unique form) would contribute to our understanding of the development of gesture processing.

Finally, our chapter has explicitly discussed the functions of gesture for both producers and perceivers. We have provided evidence suggesting that gesture’s functions arise from its status as representational action both for gesture doers and for gesture seers. Our framework can thus be applied to both situations. Nevertheless, there may be differences in how gesture works for doers and seers. For example, when children are shown a gesture that follows speech and is thus produced on its own, they do no better after instruction than when the same information is displayed entirely in speech (Congdon et al., 2017). In other words, learning from a seen gesture may depend on its being produced simultaneously with speech. In contrast, when children are told to produce a gesture, they profit from that instruction (Brooks & Goldin-Meadow, 2016) and retain what they have learned (Cook et al., 2008) even when the gesture is produced on its own without
speech. Learning from a *produced* gesture does not seem to depend on its being produced along with speech. Doing vs. seeing gesture might then function through distinct mechanisms, although we suggest that gesture’s status as a representational form is still essential to both. Additional studies that directly compare learning from doing vs. seeing gesture are needed to determine whether the mechanisms that underlie these two processes are the same or different, and whether seeing vs. doing gesture interacts with interpreting gesture as representational action.

**Conclusions**

In this chapter, we present a framework for understanding gesture’s *function*. We propose that gesture has unique effects on thinking and learning because of its status as *representational action*. More specifically, the fact that gesture is representational action, and not instrumental action, is critical to its ability to support generalization beyond the specific and to support retention over a period of time. Our proposal is agnostic as to whether gesture’s role in learning depends on its being embodied – it might or might not. It is also agnostic as to how gesture is produced, that is, its mechanism. The proposal is designed to account for why gesture is produced – for the functions it serves, particularly in a learning context. Our proposal is thus not inconsistent with the mechanistic account of gesture production proposed in the Gesture-as-Simulated-Action framework (Hostetter & Alibali, 2008). But it does offer another perspective, a functional perspective, that highlights the differences between gestures and other types of actions. By expanding the investigation of gesture to include a framework built around its functions, we hope to arrive at a more complete understanding of how and why we move our hands when we talk.

**Acknowledgements**

This work was supported by NIH grant number R01-HD047450 and NSF grant number BCS-0925595 to Goldin-Meadow, NSF grant number SBE-0541957 (Spatial Intelligence and Learning Center, Goldin-Meadow is a co-PI), and a grant from the Institute of Education Sciences (R305 B090025) to S. Raudenbush in support of Novack. SGM thanks Bill Wimsatt and Martha McClintock for introducing her to the distinction between mechanism and function, and for convincing her of its importance in understanding scientific explanation back in 1978 when we taught our first Mind course together in the Social Sciences Division at the University of Chicago.
References


Bernardis, Paolo and Gentilucci, Maurizio 2006. “Speech and gesture share the same communication system.” *Neuropsychologia* 44 (2): 178–190. doi:10.1016/j.neuropsychologia.2005.05.007


Schachner, Adena and Carey, Susan 2013. “Reasoning about ‘irrational’ actions: When intentional movements cannot be explained, the movements themselves are seen as the goal.” *Cognition* 129 (2): 309–327. doi:10.1016/j.cognition.2013.07.006


Trofatter, Caroline, Kontra, Carly, Belicko, Sian and Goldin-Meadow, Susan 2015. “Gesturing has a larger impact on problem-solving than action, even when action is accompanied by words.” *Language, Cognition and Neuroscience* 30 (3): 251–260. doi:10.1080/23273798.2014.905692


