THE UNIVERSITY OF CHICAGO

PERCEIVING MOVEMENT AS REPRESENTATIONAL ACROSS THE LIFESPAN

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY

MIRIAM ALANA NOVACK

CHICAGO, ILLINOIS

AUGUST 2016

To my Oma, Ilsa Roller, who taught me that education is a gift. I learn and I teach, for her.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank all the people who have helped me with this work, and with my academic journey to this point. First and foremost, I would like to thank Susan Goldin-Meadow for welcoming me into the wonderful world of gesture research, for embracing and guiding my ideas, for giving me countless opportunities and for teaching me how to write clearly, speak competently, and teach enthusiastically. You are an amazing woman, and a true academic role model. I would like to thank Amanda Woodward for 8 years of incredible mentorship, particularly for believing in me as an undergraduate at the University of Maryland. I am truly grateful for everything you have done for me. I would also like to thank the other members of my committee, Susan Levine, and Ed Awh for their helpful advice, and guidance.

Second, I would like to thank the amazing research assistants who helped me with this work, particularly Naureen Hemani-Lopez, Kristin Pltath, Julia Villarreal and Sarah McKissock. I am grateful to the Museum of Science and Industry for allowing us to test in their space, and to the hundreds of adults, children, and infants who participated in these studies and made all this work possible.

I would like to thank my parents for their constant love and support. For raising me with the value of finding a job that I love, and for keeping me on the family cell phone plan while I do that job.

I would like to thank my friend-collaborators – Elizabeth Wakefield and Eliza Congdon, for the hours of emotional support throughout successes and challenges, joys and heartbreaks. Liz – this dissertation is as much yours as it is mine. None of these ideas would exist without you or the revolution that started this all. Eliza – you are my graduate-school soul mate. You have single-handedly defined the past 6 years of my life, and I shutter to think of the next 6 years without you by my side. May this be just the beginning of years of friendship and collaborations between the three of us.

Finally to my fiancé Alexander Itkin - for keeping me sane, calming my anxieties, reading my documents, and being my best friend. I love you. Thank you, in advance, for continuing to do all those things.

ABSTRACT

It is well established that gestures, the hand movements that accompany speech are an integral part of communication, ubiquitous across cultures, and a unique feature of human behavior (e.g., Goldin-Meadow & Brentari, in press). However, until now, no one has asked how listeners *know* when hand movements are supposed to be gestures, and not some other class of movement, such as instrumental actions, or meaningless movements produced for the sake of movement (e.g., Schachner & Carey, 2013). In this dissertation, I will explore what features of an event lead observers to see movement as gesture, and how seeing gesture changes across ontogeny. I begin by reviewing the previous literature on the role of gesture in learning and discuss a framework for understanding the functions of gesture on cognitive processes. I will then present three studies exploring how humans, across the lifespan, come to see and categorize hand movements. In Chapter 1, I will explore the features of movement that lead adults to interpret it as representational (as opposed to meaningless movement in the air, or instrumental actions-on-objects). Chapter 2 will investigate the development of movement categorization in childhood (ages 4-9). Chapter 3 will ask how infants first begin to categorize movements, and whether a system for categorizing hand movements is in place in infancy. Finally, I will integrate the findings from these studies into the broader question of how seeing movement as gesture affects learning from gesture.

Introduction

When we speak, we move our hands – we gesture. Gestures are hand movements that provide spatial, iconic, and imagistic content to our speech. They can either emphasize what we say in speech or contribute complementary ideas not present in our spoken message (McNeill, 1992). Importantly, the gestures that we produce while speaking affect our thoughts, as well as the thoughts of others (Goldin-Meadow, 2003). For this reason, gestures have been shown to be highly influential in learning and instructional contexts (Novack, & Goldin-Meadow, 2015). Previous research has shown that learners are more likely to learn new ideas and remember what they learned if their instructor gestures during a lesson than if the instructor does not gesture (e.g., Condgon et al., 2016; Cook, Duffy & Fenn, 2013; Ping & Goldin-Meadow, 2008; Singer & Goldin-Meadow, 2005; Valenzeno, Alibali & Klatzky, 2003). Further, learners who *produce* gestures themselves are more likely to discover novel insights (Goldin-Meadow, Cook & Mitchell, 2009), retain what they learned (Cook, Mitchell, & Goldin-Meadow, 2008), and generalize their understanding to novel contexts (Novack, Congdon, Hemani-Lopez & Goldin-Meadow, 2014) than learners who do not produce gestures.

Although there is substantial research demonstrating that gestures are a powerful learning tool, no one has asked how a learner *recognizes* that an instructor's hand movements are supposed to be gestures. For example, if a teacher motions toward numbers written on the chalkboard during a math lesson, those hand movements may be intended to be gestures. But what if the teacher waves her hand near the board in an attempt to clear away a cloud of chalk dust? How does the learner know to ignore those irrelevant hand movements, but pay attention to the meaningful ones? And further, how important is it for a learner to recognize movements as gestures in order to be able to learn from them?

In this dissertation I will explore how humans across the lifespan come to process movement as representational (i.e., as gestures). The ultimate purpose of the current studies is to understand how seeing movement as meaningful changes across development, and to better understand what influences the process of learning from gesture. I begin by reviewing the previous literature on the role of gesture in learning, and discuss a framework for understanding the functions of gesture on cognitive processes. I will then present three studies exploring how humans, across the lifespan, come to see and categorize human movement. In Chapter 1, I will explore what features of movement lead adults to interpret it as representational (as opposed to meaningless movement in the air, or instrumental actions-on-objects). Chapter 2 will explore the development of explicit movement analysis in childhood (ages 4-9). Chapter 3 will ask how infants first begin to categorize movements, and whether a system for categorizing movements is in place in infancy. Finally, I will integrate the findings from these studies into the broader question of how seeing movement as gesture affects learning from gesture. Overall, the goal of this dissertation is to establish where gesture fits into framework of movement analysis, and to explore how categorization of movement as gesture affects the cognitive functions that gesture serves – specifically for learning.

Learning from Gesture

Imagine you are a science teacher trying to explain what makes a molecule a stereoisomer. You need to explain to your students how to mentally rotate the molecule comparing the locations of the various atoms. Describing this complicated mental transformation in words would likely be confusing and cumbersome. However, using your hands to *show* the rotation might be more natural, and make more sense to the students. Research has found that teachers spontaneously gesture while linking new concepts in a classroom setting (Alibali et al., 2014; Roth, 2007; Roth & Lawless, 2002), and when instructing children individually on math problems (Goldin-Meadow, Kim, & Singer, 1999). In fact, it is rare for teachers *not* to gesture while explaining new concepts (Goldin-Meadow et al., 1999). Students, similarly, are likely to gesture while trying to learn new ideas. Students spontaneously gesture while explaining their reasoning on Piagetian conservation problems (Church & Goldin-Meadow, 1986), mathematical equivalence problems (Perry, Church, & Goldin-Meadow, 1988), geological phenomena (Atit et al., 2013; Kastens, Agrawal, & Liben, 2007), and even chemistry structures (Stieff, 2011).

Importantly, both the gestures that students see and the gestures that students produce support learning outcomes. With respect to *seeing* gesture, children are more likely to profit from instruction if their teacher gestures while she explains a concept, than if she does not (Congdon, et al., 2016; Ping & Goldin-Meadow, 2008; Signer & Goldin-Meadow, 2005; Valenzeno, et al., 2003). For example, the pointing and tracing gestures that teachers use to indicate the symmetry of shapes helps preschoolers learn the concept of bilateral symmetry (Valenzeno, et al., 2003). Additionally, including gesture in instruction allows teachers to provide students with multiple strategies *at the same time*. Singer and Goldin-Meadow (2005) found that children learned more from a math lesson in which a teacher simultaneously presented two correct strategies, one in speech and another in gesture (speech+gesture), compared to a lesson in which the teacher presented the same two strategies entirely in speech, which, of course, had to be produced sequentially (speech→speech).

Children's own gestures also influence their learning. Children who spontaneously gesture as they explain their ideas, at times, convey *different* information across speech and gesture – a phenomenon known as *mismatching*. Children who produce mismatches when explaining their explanations of a task are more likely to profit from instruction on the task than children who do not produce mismatches (e.g., Church & Goldin-Meadow, 1986; Perry et al., 1988). Gesture-speech mismatch is thought to occur when a child is on the brink of conceptual change. If a child is entertaining novel ideas, the child may not be able to express her underdeveloped knowledge through speech. Instead, the act of gesturing may allow these implicit ideas to come to the surface, bringing them into child's awareness, and priming the child to benefit from subsequent instruction (Alibali & Goldin-Meadow, 1993).

Since gesture provides an avenue through which learners can consider new ideas, what happens to children who do not gesture? Might they be missing out on important learning opportunities? It turns out that simply encouraging learners to gesture can allow these implicit ideas to surface. Broaders, Cook, Mitchell and Goldin-Meadow (2007) asked children to explain their solutions to incorrectly solved mathematical problems; they then asked them to solve a new set of comparable problems and encouraged half the children to gesture as they explained their solutions. Broaders et al. found that children told to gesture added novel strategies to their repertoires, but those strategies were found only in gesture; children who were not encouraged to gesture did not add strategies to their repertoires in either gesture or speech. Importantly, children who added novel strategies uniquely in gesture were also more likely to profit from instruction in math—after the lesson (when they were no longer gesturing), the children were able to solve math problems on a paper-and-pencil test that they could not solve before the lesson. Encouraging children to move their hands can activate implicit, and correct, ideas that then prime children for learning.

The same phenomenon has recently been replicated in the domain of moral reasoning. Children who were encouraged to gesture while explaining their reasoning about a moral dilemma (debating contractual obligations against obedience to a higher authority) were more likely to express multiple perspectives in gesture (reflecting a greater understanding of the multiple elements involved in the dilemma) than children who were not encouraged to gesture. Moreover, after a lesson in moral understanding (when children were no longer gesturing at high rates), the children who had gestured during the lesson increased the number of perspectives they mentioned in speech; in fact, the more multiple-perspectives they produced in gesture prior to the lesson, the more multiple-perspectives they produced *in speech* after the lesson (Beaudoin-Ryan & Goldin-Meadow, 2014). Encouraging children to move their hands not only activates implicit ideas that prime children for learning in spatial domains, like math, but also in inherently non-spatial domains, like morality

If spontaneously producing correct strategies uniquely in gesture leads to learning, what happens if children are explicitly *taught* to produce correct strategies in gesture? Goldin-Meadow, Cook and Mitchell (2009) gave third-graders instruction in how to solve a missing addend mathematical problem, such as *3+4+9=__+9*, by teaching them to produce gestures that represented a correct strategy for solving the problem. All of the children were taught to say the words, "I want to make one side equal to the other side" (*equivalence,* a correct strategy for solving the problem). One group was taught to produce another correct strategy (*grouping*) in gesture—they produced a *V-point* gesture with their index and middle fingers to the first two numbers in a math problem (i.e., the 3 and 4 in the *3+4+9=__+9* problem) and then pointed at the blank on the other side of the equation; the *V-point* gesture, which represents the idea that the problem can be solved by *grouping* and then adding the two addends, is one that children spontaneously produce when explaining their correct solutions to these problems (Perry et al., 1998). Another group of children was taught to produce a partially correct *grouping* strategy in gesture—they produced their *V-point* to the second two numbers in the problem (i.e., the *4* and *9* in the example) and then pointed at the blank; this gesture highlighted grouping but focused children's attention on the wrong numbers. A final group of children was taught no gestures at all and just learned the equivalence strategy in speech. Encouraging children to produce the *grouping* gesture during the lesson led them to produce the strategy explicitly in speech after the lesson, and to solve more problems correctly on a written posttest (when they were no longer gesturing) than they had solved prior to the lesson. Interestingly, even children in the *partially correct gesture* condition, whose attention was directed to the wrong numbers, improved relative to children in the *speech alone* condition (although not as often as children in the *fully correct gesture* condition), suggesting that gesture was doing more than just directing the child's attention. If all gesture did was direct attention, then children in the *partially correct gesture* condition should have performed worse than children in the *speech alone* condition. Teaching children to produce correct strategies uniquely in gesture can lead to learning.

How Gesture's Status as Representational Action Supports Learning

Thus far, I have presented evidence showing that gestures support learning. Given this phenomenon, we can begin to think about *why* gesture supports learning. Gesture is a unique tool because it naturally encompasses many features that have been found to support learning. For example, in my work I have found that gesture can influence visual attention (Novack et al., 2016), engage the motor system (Wakefield, et al., 2016), and provide multiple simultaneous ideas (Congdon et al., 2016), all factors that are partially responsible for gesture's learning effects. However, rather than exploring *how* gesture works, it is also useful to think about why gesture serves the *functions* that it does.

In recent work, we have proposed that gesture produces effects on thinking and learning because it is *representational action* (Novack & Goldin-Meadow, 2016). When we say here that gestures are *representational actions*, we mean that gestures are meaningful substitutions and analogical stand-ins for ideas, objects, actions, relations, etc. This use of the term *representational* should not be confused with *representational gestures*—a category of gestures that look like the ideas and items to which they refer (e.g., iconic and metaphoric gestures). Our proposal that gestures are representational is meant to apply to all types of nonconventional gestures, including representational gestures (iconics, metaphorics), deictic gestures and even beat gestures. For example, iconic gestures can represent actions or objects, deictic gestures stand in for the entities to which they refer; beat gestures reflect discourse structure. Most of this dissertation explores the functions of iconic gestures, but we believe that our framework can be applied to all (non-conventional) gestures.

Gestures are *representational* in that they represent something other than themselves, and they are *actions* in that they involve movements of the body. Most importantly, the fact that gestures are representational actions differentiates them from full-blown instrumental actions, whose purpose is to affect the world by directly interacting with it (e.g., grabbing a fork, opening a canister). In other words, gesture is a type of action, but a special kind of action - one that represents the world rather than directly impacting it. For example, producing a twisting gesture in the air near, but not on, a jar will not open the jar; only performing the twisting action on the jar itself will do that. Yet even though gesture does not accomplish anything physical, it can change our cognition in ways that action does not. Additionally, gestures are movements that occur in the air, physically *off* objects. This feature means that gestures are not beholden to the affordances of any single object or set of objects, but instead can provide an abstract representation that highlights only the most crucial features of a concept or idea. This physical and metaphorical ‘distance’ from physical objects and real-world actions might allow learners to form a more flexible representation of a new idea that is broadly applicable to novel contexts and situations.

Gestures have many similarities to actions simply because they are a type of action. Theories rooted in embodied cognition maintain that action experiences have profound effects on our interactions with objects (James, 2010), perceptions of other’s actions (Casile & Giese, 2006), and even language processing (Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008). The Gesture as Simulated Action (GSA) framework, a recent mechanistic account of gesture production, grew out of the embodied cognition literature. The GSA proposes that gestures are the manifestation of action programs, which are simulated (but not actually carried out) when an action is imagined (Hostetter & Alibali, 2008). Following at least some accounts of embodied cognition (see Wilson, 2002, for review), the GSA suggests that when we think of an action (or an object that can be acted upon), we activate components of the motor network responsible for carrying out that action, in essence, simulating the action. If this simulation surpasses the “gesture threshold,” it will spill over and become a true motor expression—an overt gesture. The root of gesture, then, according to this framework, is *simulation*—partial motor activation without completion.

The GSA framework offers a useful explanation of how gesturing comes about (its mechanism) and the framework highlights gesture’s tight tie to action. However, this framework is primarily useful for understanding how gestures are produced, not for how they are understood. I suggest that viewing gestures as simulated actions places too much emphasis on the action side of gesture and, in so doing, fails to explain the ways in which gesture’s functions differ from those of instrumental actions. My hypothesis is that the effects gesture has on thinking and learning grow not only out of the fact that gesture is itself an action, but also out of the fact that gesture is abstracted away from action—the fact that it is representational. This framework proposes that whether or not gesture is simulated action in terms of its mechanism—it is clearly *not* reducible to action in terms of its function. 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.

Learning from Gesture Versus Other Types of Movement

The proposal that gestures are representational actions implies that the effects of seeing and producing gestures differs from seeing or producing other types of movements. Specifically, I argue that the functions of gesture differ from both instrumental actions, whose purpose is to affect the world by directly interacting with it (e.g., grabbing a fork, opening a canister), as well as movements for their own sake (see Schachner & Carey, 2013), whose purpose is the movement itself (e.g., dancing, exercising).

There is existing evidence that the *functions* of these three different kinds of movements are distinct. For example, relevant gestures, but not irrelevant hand movements prime children for instruction (Brooks & Goldin-Meadow, 2015). Children who were asked to produce meaningful gestures (sweeping their hands under the two sides of an equation) before a math lesson were more likely to profit from that instruction than children who produced *irrelevant* movements, or meaningless movements, prior to the lesson. This suggests that guiding a learner's movements can have a delayed effect on learning, and can prime children for later instruction. Importantly, this effect only happens for *meaningful* movements.

Gestures have also been found to have different effects on learning outcomes than actions on objects. In a previous study, I found that producing gestures during a math lesson, but not producing actions on objects, supported generalization (Novack, et al., 2014). In this study, children were asked to either produce actions on objects, or gestures off objects while learning how to solve mathematical equivalence problems. Although all children learned how to solve the types of problems on which they were trained, only children who learned via gesture were also able to correctly transfer their understanding to novel problem forms, suggesting that learning through gesture leads to more flexible understanding than learning through action on objects. Additionally, children's verbal explanations at the conclusion of the study showed this distinction. All children had learned a correct speech strategy along with their movement. Children who had learned to produce those words while producing action would often parrot that phrase at the posttest while explaining their *incorrect* solutions. That is, if given the problem 2+8+5=__+5, they may write down the answer 15 (an incorrect add-to-equal answer) but say, "I want to make one side equal to the other side," the correct *equalizer* strategy that they were taught in training. In contrast, children in the gesture conditions (there were two), who had learned that equalizer phrase while producing gesture, were particularly likely to use the phrase if they answered the problems correctly. That is, children who learned a correct verbal strategy accompanied by gesture came to understand the meaning of the words better than children who learned that same verbal strategy accompanied by actions on objects.

Finally, studies outside the area of learning have found cognitive disparities between seeing or producing actions on objects and gestures off objects. For example, if adults are asked to explain how to throw a dart using the object in front of them (an instrumental action) or using just their hands with no object (a gesture), they display a tighter link between speech and the accompanying dart-throwing gesture than between speech and the accompanying dart-throwing action (Church, Kelly & Holcombe, 2013). Other signatures of the gesture-speech system also seem to be unique to gesture, and are not found in instrumental actions. For example, gestures are more often produced with the right hand (suggesting a link to the left-hemisphere speech system), whereas self-touching adaptors (e.g., scratching, pushing back the hair), which are instrumental actions, are produced with both hands (Kimura, 1973).

Together, the previous literature clearly shows that gestures affect thinking and learning in ways that differ from other forms of movement, specifically instrumental actions, and meaningless movements in the air. What is unclear is how learners know which movements are gestures and which movements are not. Although scholars have provided definitions of "gestures" (See McNeill, 1992; Kendon, 2004) and researchers have used various criteria to code gestures in experiments (Church, Kelly & Wakefield, 2015), no one has asked how people, specifically, children (who are most likely to be seeing gesture in instructional contexts) spontaneously distinguish gestures from other types of movements.

The goal of the following studies is to understand how humans view movements as gestures, and whether a propensity to see movement as gesture influences learning. Ultimately, these questions are part of a broader research program aimed at understanding the perceptual, contextual, and even social influences that allow us to categorize a movement as gesture.

Chapter 1: What Makes a Movement a Gesture?

Published as: Novack, M. A., Wakefield, E. M., & Goldin-Meadow, S. (2016). What makes a movement a gesture? *Cognition, 146,* 339-348. doi: 10.1016/j.cognition.2015.10.014

When a runner grabs a water bottle (or a pint of ice cream) after a long run, we quickly interpret his action as being goal-directed and intentional. We do not puzzle over the specific arm movements used to acquire the water or ice cream; we understand that his actions were performed to reach a goal: to quench his thirst or reward himself with a treat. This example illustrates a skill found in both adults and children—the ability to interpret the actions and intentions of others as goal-directed, actions that are guided by top-down, hierarchical cognitive processes (e.g., Baldwin & Baird, 2001; Bower & Rinck, 1999; Searle, 1980; Trabasso & Nickels, 1992; Zacks, Tversky, & Iyer, 2001).

The literature on action understanding typically assumes that, when an actor performs a movement, the actor's goal is not the movement itself, but rather the impact that the movement has on the surrounding world. Recently, Schachner and Carey (2013) challenged this assumption as overly simplistic by demonstrating that, under certain conditions, movements can be interpreted, not as a means to an external goal, but instead as a goal unto itself—movement for the sake of movement. Here we broaden the investigation of action understanding to include a third type of movement. We examine the conditions under which movement[1] is interpreted not as movement of an object directed toward a goal, nor as movement performed for its own sake, but rather as movement that *represents* other types of actions. This third type of movement is *gesture.*

A foundational body of research demonstrates that humans interpret actions in terms of the actor's intentions and goals (e.g., Baldwin & Baird, 2001; Bower & Rinck, 1999; Searle,

[1] Here, we use the word movement to refer only to intentional actions. Accidental or unintentional actions (e.g., knocking a glass off a table; moving up and down on a merry-go-round) are beyond the scope of this paper.

1980; Trabasso & Nickels, 1992; Zacks et al., 2001). For example, adults naturally parse continuous action into smaller, goal-directed segments (Newtson, 1976), and when asked to describe ambiguous scenes, adults and 5-year-old children will create a goal-based explanation for an actor's actions (Trabasso, Stein, Rodkin, Munger, & Baughn, 1992). The ability to interpret object-directed actions in terms of an actor's goals arises in the first year of life—infants as young as 6-months notice when an actor uses the same movements to reach for a new goal, but not when the actor uses new movements to reach the same goal (Buresh & Woodward, 2007; Woodward, 1998). In other words, infants consider a change in what is being reached for as novel and worth attending to, but do not consider a change in the movements made to achieve the same goal as noteworthy or surprising. At 10-11 months, children are surprised when an actor's reaching movement stops abruptly before an object is grasped, but not when the actor's reaching movement stops after the object has been grasped (Baldwin, Baird, Saylor, & Clark, 2001). There is thus ample evidence that both adults and children interpret movements as goal-directed. An important caveat, however, is that much of the relevant research has been done on movements directed toward objects, with an emphasis on external goals.

Recently, Schachner and Carey (2013) broadened the scope of these studies to include movement performed in the absence of objects, and showed that, under these circumstances, adults tend to interpret the movement as having its own intrinsic goal, as movement for its own sake. In the first in a set of experiments, participants were asked to interpret videos in which a character either moved objects from one location to another (*Objects Present* condition) or made the same movements without the objects present (*Objects Absent* condition). Participants described the movements in terms of external goals when the objects were present (e.g., his intention was to sort the colored balls), but in terms of movement-based goals when the objects were absent (e.g., his intention was to jump into the air and move to the left and right). In a another experiment, Schachner and Carey (2013) showed that participants also attribute movement-based goals to an actor producing inefficient movements. The actor's movements were either inefficient (i.e., moving toward the target and then away from it) or efficient (moving toward the target), and participants were more likely to provide movement-based descriptions for inefficient than for efficient movements. The traditionally held view that humans interpret actions in terms of external goals thus seems to be too narrow. In certain circumstances, adults will see the intent of an action as the completion of the action itself—movement for the sake of movement.

Into this two-dimensional perspective of how humans interpret action, we propose another dimension of intentional, goal-directed movement—gesture. Gestures are movements of the hands that accompany speech and communicate information to listeners (Kendon, 1994). Although researchers have created elaborate coding systems for identifying, describing, and interpreting the meaning of gestural movements (e.g., Church, Kelly, & Wakefield, 2015; Kendon, 2004; McNeill, 1992), and even naïve observers, who are not trained gesture coders, are able to reliably interpret gestural movements (Goldin-Meadow & Sandhofer, 1999), there has not been a systematic investigation of the circumstances under which observers interpret movement as a representation (i.e., as gesture), as opposed to goal-directed or for its own sake. Gesture is unlike object-directed movement, whose goal is to achieve some change in the world, and unlike movement for its own sake, whose goal is to produce the movement itself. In contrast to both of these goals, the purpose of gesture is to reference or represent other movements, objects, or even abstract ideas, and importantly, can include movements that represent *either* external or movement-based goals. In other words, gesture does not depict a change in the world (e.g., opening a jar by twisting it), or display patterns of movement (e.g., performing steps in a dance), but instead *represents* movement that could change the world (e.g., a gesture showing how the jar could be twisted open) or *represents* movement that stands on its own (e.g., a gesture showing how the dance should be performed). Importantly, observers respond differently to gesture than to other types of movements (Dick, Goldin-Meadow, Hasson, Skipper, & Small, 2009; Kelly, Healy, Ozyurek, & Holler, 2014), and these differences in response can have an impact on thinking and learning (Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014; Trofatter, Kontra, Beilock, & Goldin-Meadow, 2015). Our goal is to determine the conditions under which a distinction between gesture and other types of movement is made, and to explore how those conditions contribute to a top-down categorization of movement as gesture.

We begin by noting that a movement has the capacity to be seen as a gesture if it does *not* cause effects on the external environment. As a result, any empty handed-movement is a candidate for being interpreted as a gesture. Our first prediction, then, is that hand movements that interact with objects will likely be interpreted as having external goals, whereas hand movements that occur *off* objects have the potential to be seen either as movements produced for their own sake, or as movements that represent.[2]

The next question is how to determine whether an empty-handed movement is a movement meant for its own sake, or a movement meant to represent. We suggest that humans have a bias to interpret movement as more than movement for its own sake. In fact, Schachner and Carey (2013) found that almost one-third of participants who saw a character move in empty space *still* attempted to describe his movement in terms of external-goals. Seeing empty-handed

[2] In some cases, an empty-handed movement can be interpreted as having an external goal; for example, when an observer thinks an agent is trying to complete an action on an object but the agent fails to make contact with the object (i.e., the agent performs an incomplete action), perhaps because the object is out of reach or is behind a barrier.

movement as movement for its own sake may be a default that is activated only when observers are unable to see the movement as anything else. If there are contextual cues that could confer meaning to an empty-handed movement, observers are likely to exploit those cues and interpret the movement as meaningful and, as a consequence, seek something that the movement could represent. Our second prediction, then, is that empty-handed movements shift from being seen as movement for their own sake to movements that represent when there is information in the context that allows observers to interpret the movements as meaningful.

Taken together, these predictions form a framework in which hand movements are seen as having external goals, movement-based goals, or representational goals depending on (1) whether the hands interact with objects and, if not, (2) whether there is information in the context that makes the movements appear meaningful to the observer. Finally, it is important to note that this framework focuses on the *hands*. Hands may be special in their ability to both carry out goals *and* communicate or represent, and may therefore be a special cue for seeing movement as representational.

We explore this framework across two studies. In the first study, we test the prediction that movements that do not interact with objects (i.e., empty-handed movements) are candidates for having representational goals if there is information in the context that makes them seem meaningful to the observer. We extend Schachner and Carey's (2013) paradigm to include a condition that meets these requirements and thus is likely to be interpreted as gesture. Schachner and Carey gave observers two types of scenes to interpret, one in which an actor acts directly on an object (to which observers attributed external-goals), and another in which the actor performs the same movements but without the object present (to which observers attributed movement-based goals). We add a third type of scene—one in which the actor performs the same movements but over (not directly on) the objects, which are present. According to our framework, we expect observers to have a bias to see empty-handed movements as meaningful, particularly if there is information in the context (in this case, the presence of objects) that supports attributing meaning to the movement. Observers should therefore interpret movement over, but not directly on, an object as a r*epresentation* of a goal-directed action on the object, that is, as a gesture. In the second study, we extend our second prediction, expanding the context to include additional cues to meaning. We ask whether these additional cues influence the observer's inclination to interpret empty-handed movement in terms of movement-based goals or representational goals. We predict that contextual cues to meaning—the presence of objects, the presence of speech, the form of the movement itself—will increase an observer's inclination to see an empty-handed movement as a representational gesture.

Study 1

As just described, Schachner and Carey (2013) held the movements performed by an actor constant, but varied the presence or absence of objects. When objects were present and acted on, observers systematically believed that the goal of the actor was to move the objects (i.e., an external goal). When objects were absent and the actor produced the same movements in the air rather than on an object, half of the observers attributed a movement-based goal to the actor (i.e., the movement itself was the intended outcome). In Study 1, we add a third condition—the objects are present but are not acted upon—and hypothesize that movements produced in this condition are likely to be interpreted as a gesture representing action.

Study 1 thus contained three conditions, the first two modeled after Schacher and Carey (2013): (1) *Action on Objects*, in which an actor directly manipulates objects; (2) *Action with*

Objects Absent, in which an actor performs the same movements, but without the objects present; (3) *Action off Objects with Objects Present***,** in which an actor performs the same movements over (not directly on) the objects, which are present. We expect to replicate Schachner and Carey's findings in the first two conditions: Observers will describe movements in the *Action on Objects* condition in terms of external goals (i.e., "she organized objects"), but will be more likely to describe movements in the *Action with Objects Absent* condition in terms of movement-based goals (i.e., "she moved her hands to the right and left").

The more interesting question to us is how participants will interpret movements in the third condition. In the *Action off Objects with Objects Present* condition, an actor performs movements in the presence of objects, but does not use her movements to manipulate those objects. Following Schachner and Carey (2013), we might expect that a movement produced in the presence of objects would be interpreted as having a movement-based goal if the movements are seen as intentional yet inefficient. However, if the presence of objects provides context that invites richer interpretation, the movements might instead been seen as representing an external goal.

Method

***Participants*.** 120 adult English-speaking residents of the United States (63 females, 55 males, 2 unreported; 40 participants in each of the 3 conditions) participated in the experiment via Amazon Mechanical Turk (https://www.mturk.com), a website through which individuals can complete tasks for small amounts of compensation (see Crump, McDonnell, & Gureckis, 2013 for validation of AMT for experimental studies). All participants were required to have had at least 95% of their previous work on Amazon Mechanical Turk judged as acceptable, and were required to have previously completed at least 100 Mturk studies. Participants were excluded if they provided inappropriate answers, suggesting that they had not watched the video stimulus (*n* = 2), or if they completed the experiment multiple times (*n* = 16). Participants were all over 18 years of age, and were ethnically diverse: White (*n* = 92); Black (*n* = 11), Asian (*n*=8); Native American (*n* = 2); more than one race (*n* = 3); unreported (*n* = 4). The task took just under 4 minutes to complete on average, and participants were compensated $0.25 for their time.

Stimuli. Video stimuli showed the torso of a woman standing in front of a table. Her chest, arms, and hands were visible, but her face was not. All videos were 10 seconds long and displayed the actor producing movements with her hands. In two conditions (*Action on Objects; Action off Objects with Objects Present*), four balls (two orange and two blue), as well as two boxes (one orange and one blue), sat on the table in front of the woman. In the third condition (*Action with Objects Absent*), no objects were present on the table. Each video is described in detail below (see Figure 1 for still frames taken from the videos).

Action on Objects. A woman picks up each of the balls on the table in front of her and places them in the color-matched boxes (see Figure 1a). Specifically, she picks up the inner blue ball with her left hand, and places it in the blue box on her left; then she picks up the inner orange ball with her right hand and places it in the orange box on her right. These actions are repeated with the outer two balls.

Action Off Objects with Objects Present. A woman produces the same movements in the *Action on Objects* condition but over the objects, not on them (see Figure 1b). More specifically, the woman maintains the hand shape necessary to grasp the balls (i.e., a palm down C-shape) and moves her hands following the trajectory she took when actually moving the objects (i.e., from one ball to the box, from another ball to the other box). As the woman never actually touches the balls or boxes, there is no change in the location of the balls during the video.

Action with Objects Absent. A woman produces the same movements as in the other two conditions, but without any of the objects present (see Figure 1c). The woman maintains the same hand shape and trajectory as in the first two conditions.

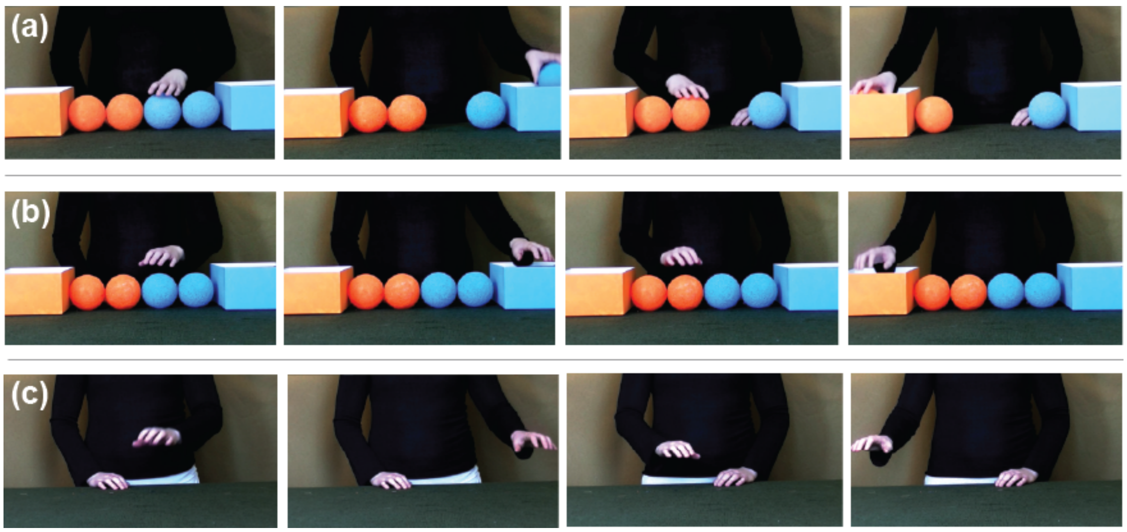

Figure 1. Stills from video stimuli in the three conditions in Study 1: (a) Action on Objects, (b) Action off Objects with Objects Present, (c) Action with Objects Absent.

To control for timing across videos, the *Action on Objects* condition was recorded first. The woman synchronized her movements to a metronome set to 80 beats per minute. After this video was filmed, she recorded a co-speech narrative ("I put one here, another in this box, the orange one in this one, and the last one in this place") that she synchronized to her movements in the *Action on Objects* video. The narrative was then played back to her as she performed the movements in the other two conditions to ensure that timing would match across videos. Audio was not included in any of the videos.

Procedure. The study was conducted through Amazon Mechanical Turk. Participants were randomly assigned to one of the three conditions. They were told that they would watch a 10-second movie and be asked questions about what they saw. Each participant saw one of the three videos and answered the free-response question, "What happened in the scene?" Finally, participants answered optional demographic questions.

Coding and Reliability. Free-responses were classified into one of 3 goal-based categories—External Goal, Movement-based Goal, Representational Goal. Although the prompt "What happened in the scene?" does not specifically probe the intentions of the agent, the pragmatics of the question, combined with the fact that the agent was producing movements of her own volition, should, for the most part, invite descriptions of goals and intentions (see Schachner & Carey, 2013 for discussion of assumptions about intentionality). A small number of responses (*n*=1, 0.01% of the total responses) were considered Uncodable. The codes are described below:

(1) *External Goal:* The movie is described in terms of actions completed on objects, with a focus on movement of objects, rather than movement of hands (e.g., "a person put blue balls in blue box and orange balls in an orange box"; "balls were placed in boxes"; "the person picked up colored balls one at a time and placed them in color coordinated boxes.").

(2) *Movement-Based Goal:* The movie is described in terms of low-level spatiotemporal movements without mentioning a higher-level goal – the description is focused on the movement of the hands themselves (e.g., "a woman waved her hand over two blue spheres and a blue box", "waving her hand, alternating sides", "a person was moving hands").

(3) *Representational Goal:* The movie is described in terms of either (a) movements representing external goals (e.g., "she demonstrated how to put balls in boxes"; "she pretended to sort objects"; "gestured about moving colored spheres"), or (b) movements

representing internal goals (e.g., “the actress was telling you how to perform the steps of a dance and where to put your feet”).

(4) *Uncodable:* The movie is described without mentioning movement at all (e.g., “there was a woman and some boxes and balls”, “red ball”).

Two researchers assigned a single code to all responses, and were blind to the condition of each participant. Coders agreed on 93% of trials, $K = 0.90$. Any disagreements were discussed between the coders and resolved.

Results

Figure 2 presents the proportion of participants whose responses fell into each of the three primary coding categories. Because the three response codes were not independent, we ran log-linear poisson models, partitioning the 3 (Condition) x 3 (Response Code) contingency table to conduct three independent analyses, with the aim of determining: (1) whether a different pattern of responses was given across conditions, and (2) whether empty-handed conditions elicited a different number of Movement vs. Representational goal responses.

To begin, we collapsed across the *Action off Objects with Objects Present* and *Actions with Objects Absent* conditions, and asked whether there was a different pattern of responses between these combined, empty-handed conditions, and the *Action on Objects* condition, creating a 2 (Condition) x 3 (Response) contingency table. Our model revealed that responses in the *Action on Objects* condition were significantly different from responses in the conditions in which empty-handed movements were observed (χ^2, 2 = 120.98, $p < .001$).

Next, we considered whether the pattern of responses was different across the two, empty-handed conditions. We directly compared responses elicited in the *Action off Objects* *with Objects Present* and *Actions with Objects Absent* conditions, creating a 2 (Condition) x 3 (Response) contingency table, which did not include data from the *Action on Objects* condition). Again, our model revealed a significant difference between conditions (χ^2, 2 = 14.73, $p < .001$).

Note that external goal responses were extremely rare in both empty-handed movement conditions (3 external goal responses in the *Action off Objects with Objects Present* and 0 in the *Actions with Objects Absent* condition). The difference between the two empty-handed conditions must then have been driven by different rates of movement-based versus representational goal responses. Indeed, if we consider the proportion of each of these response types, we find that, in the *Action off Objects with Objects Present* condition, representational goals were the predominant response (68%), whereas movement-based goals were the dominant response (63%) in the *Actions with Objects Absent* condition. A 2 (Condition) x 2 (Response) chi-square, comparing the number of movement-based and representational goal responses elicited in the *Action off Objects with Objects Present* versus *Actions with Objects Absent* conditions, confirmed this difference as significant (χ^2, 1 = 10.54, $p < .001$).

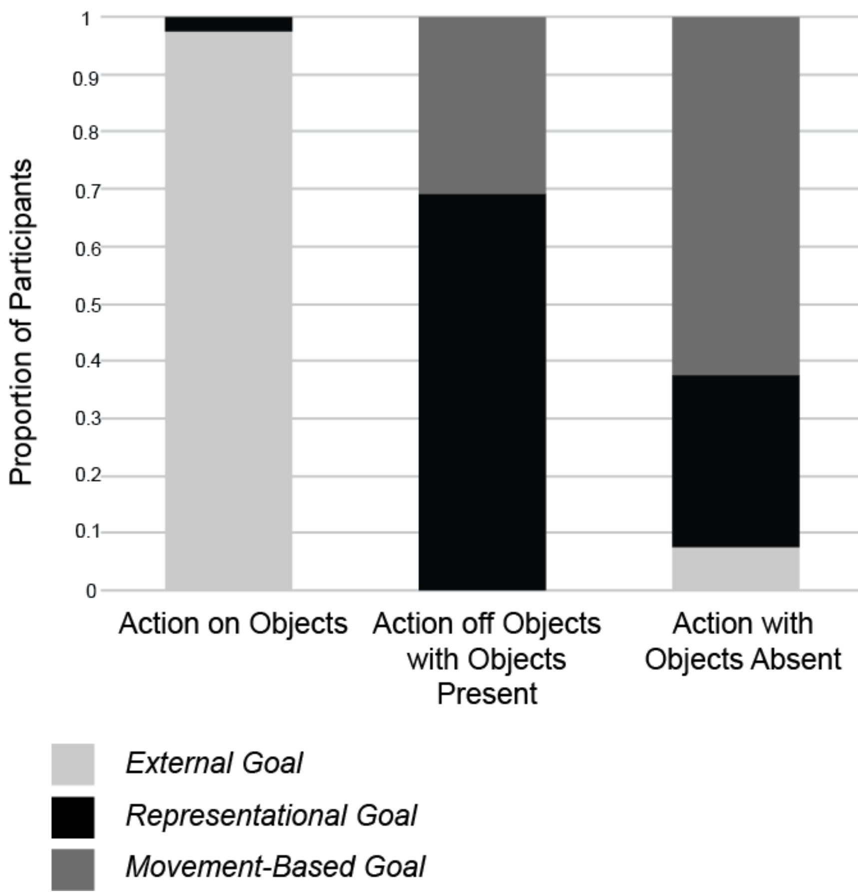

Figure 2. Proportion of participants in each condition giving External, Representational or Movement-based goal responses.

Our results provide support for a framework in which empty-handed movements have the capacity to be interpreted in terms of representational goals. Humans spontaneously describe some empty-handed movement as representational, giving credence to gesture as a unique movement category. Our results also show that observers predominately interpret empty-handed movements as representational when objects are present, but not acted on. In addition, we replicate Schachner and Carey (2013), by showing that observers predominately attribute a movement-based goal to an actor who produces movements without any objects present, although *some* participants do describe movement in this context as representational.

Study 1 therefore establishes that empty-handed movements can be seen as representational, but also as movement for the sake of movement. In Study 2, we explore contextual cues that have the potential to make the empty-handed movements appear more meaningful. We predict that these cues will then encourage observers to interpret the empty-handed movements as representational.

Study 2

In Study 2, we explore the effect of context on the interpretation of empty-handed movements. Our framework predicts that the richer the context in which an empty-handed movement occurs, the more likely it is that the movement will be interpreted as representational, rather than as movement for its own sake. This hypothesis was supported by our results from Study 1—having objects in the scene provided a richer context that elicited more representational goal responses than having no objects present. Next, we test the effect of three contextual cues that should help viewers differentiate between interpreting movement in terms of representational goals versus movement-based goals. We aim to replicate our results on the presence or absence of objects, and consider two additional cues—the form of the movement itself (in particular, the shape of the hand) and communicative intent (whether the movement is accompanied by speech).

Study 1 gave us reason to believe that the presence of objects can have a strong influence on whether empty-handed movements are seen as representational or movement-based Performing an empty-handed movement in the presence of an object makes it easier to glean meaning from that movement than if it were performed without the object present. For example, if a twisting movement is performed near a jar, an observer is more likely to interpret the movement as a gesture for jar-opening than if it is performed without the jar present (the twister might just be flexing her fingers). Thus, as in Study 1, the first factor that we vary is the presence or absence of objects, none of which is ever touched. We hypothesize that observers will be more likely to interpret an empty-handed movement as a gesture if the movement is seen as meaningful; producing the movement near a relevant object is likely to increase its meaningfulness.

A second way that the empty-handed movements may be informative is in how closely the movement resembles actual movement needed to act on the objects—in other words, the precision or specificity of the movement. A rotating motion produced with a grasping handshape resembles the act of jar-opening, making it easier to glean meaning from the movement—easier than if the rotating movement were produced with an index finger handshape, Thus, the second factor we vary is the shape of the hand used in the movement. In particular, we vary whether the hand is shaped as it would have been had the movement been performed directly on the object—a grasping handshape vs. an index finger that traces the path, a tracer handshape. We hypothesize that observers will be more likely to interpret an empty-handed movement as a gesture if it is produced with a meaningful handshape (in this case, a grasping handshape) than if it is produced with a handshape that is more difficult to interpret (a tracer handshape).

Finally, most spontaneous gestures are produced along with speech (McNeill, 1992) and are thus part of a communicative act. Typically, the speech that accompanies a gesture is relevant to that gesture; the information conveyed in gesture either complements or supplements the information conveyed in speech (Goldin-Meadow, 2003). As we were interested in whether the presence of speech (rather than its content) affects when an empty-handed movement is interpreted as a gesture, we filtered the speech to retain its prosody but render it unintelligible.

We hypothesize that observers will be more likely to interpret an empty-handed movement as representational if it is seen as part of a communicative act, that is, if it is produced along with speech (even if the speech is uninterpretable), than if it is produced in silence.

Method

Participants. An additional 320 adult English speakers participated in Study 2 (40 in each of 8 conditions; 166 females, 150 males; 4 unreported), through the Amazon Mechanical Turk website (https://www.mturk.com). Requirements for participation and compensation were identical to Study 1. Participants were excluded if they failed a prescreening audio check or had other technical difficulties viewing stimuli (*n* = 5), or if they had previously completed Study 1 (*n* = 42). As in Study 1, participants were over 18 years of age, and were ethnically diverse—White (*n* = 224); Black (*n* = 38), Asian (*n* = 22); Native American (*n* = 4); Native Pacific Islander (*n* = 2); more than one race (*n* = 21); unreported (*n* = 9). Participants took just over 4 minutes to complete the task.

Stimuli. There were eight videos used in Study 2 (see for stills taken from the videos). Videos were structured as in Study 1, showing an actor performing the same movements viewed by participants in Study 1, and maintaining the timing of the movements across videos. The videos varied along 3 dimensions: (1) *Handshape* (Grasping handshape vs. Tracer handshape): The actor produced the movements using either a grasping handshape (i.e., the same palm-down C-hand shape used in Study 1) or a tracer hand shape (i.e., a pointing hand shape). (2) *Object Presence* (+Objects vs. –Objects): The actor produced the movements near, but not on, the same objects used in Study 1 (i.e., 4 balls, 2 boxes) or without any objects present. (3) *Speech Presence* (+Speech vs. –Speech): The actor produced the movements with speech or without it. The speech described what the actor was doing with her hands as she was moving them (see

Study 1); the speech was filtered to retain prosody but be unintelligible. Filtering was achieved through the program Praat (a rectangular band filter was applied to frequencies less than 450 Hz), and individuals naïve to the original recording confirmed that the speech sounded muffled and was not comprehensible. The –Speech videos contained no sound.

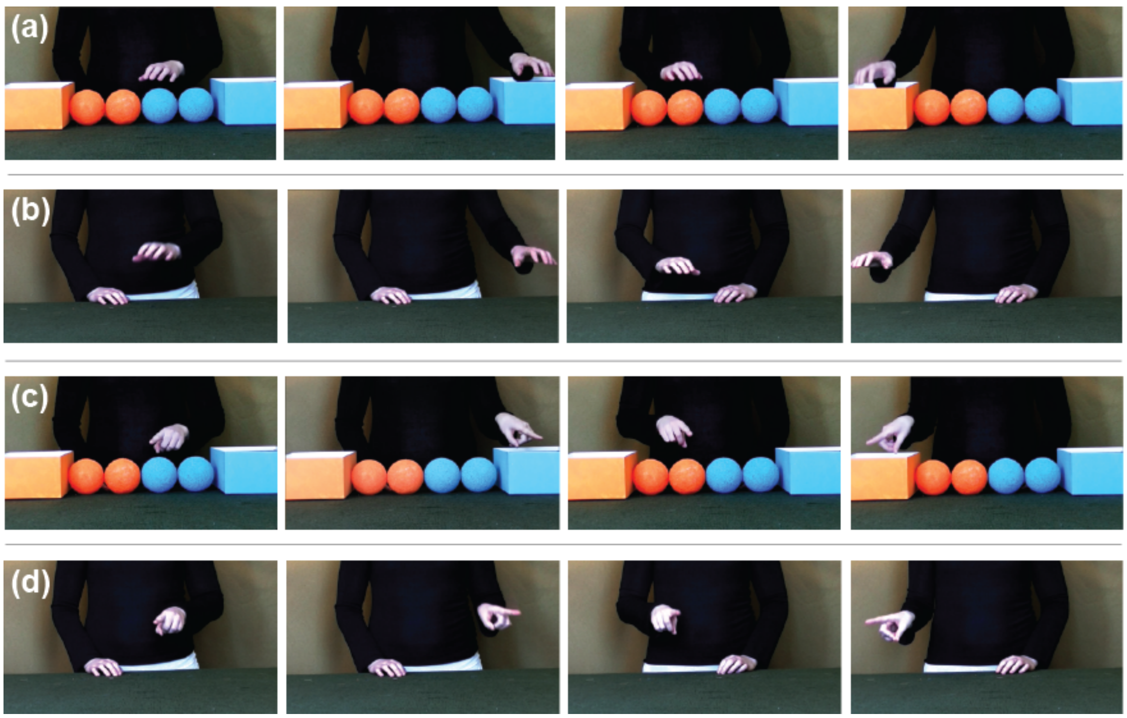

Figure 3. Stills from the video stimuli in Study 2: (a) Grasping Handshape +Objects, (b) Grasping Handshape **-**Objects, (c) Tracer Handshape +Objects, (d) Tracer Handshape **-**Objects. Two version of each video were created (+Speech, **-**Speech), resulting in 8 videos.

Procedure. The procedure was identical to Study 1 with one exception: Participants viewing +Speech videos were given an audio file and asked to transcribe what they heard to confirm that the audio was functioning properly on their computers. They were permitted to listen to the audio file, which contained a string of 6 numbers, as many times as needed. Participants who failed to report the correct numbers were excluded from analyses (n = 5).

Coding and Reliability. As in Study 1, participants' responses were coded as describing either (1) *External goals*, (2) *Representational goals*, (3) *Movement-based goals*, or (4) as *Uncodable*. Only 7 responses (2%) were uncodable. Again, two researchers who were blind to the experimental conditions of the participants coded each response, and each response was only assigned a single code, using the same method as Study 1. Coders agreed on 94% of trials, $K = .89$, and any coding disagreements were discussed and resolved by the researchers.

Results

Our goal was to determine the conditions under which observers identify a movement as a gesture. To that end, we tabulated the proportion of participants who described a *representational goal* in each of the 8 conditions (see Figure 4). As would be expected from Study 1, almost no participants described *external goals* for the empty-handed movements in Study 2. When participants did not describe *representational goals*, they primarily described *movement-based goals.* We therefore used a logistic regression to predict participants likelihood of providing a *representational goal* response with Handshape (Grasping, Tracer), Speech Presence (+Speech, –Speech) and Object Presence (+Objects, –Objects) as fixed effects. The model revealed main effects of Handshape, $\beta= 1.49$, $z=5.9$, $p < .001$, and Object Presence, $\beta = .91$, $z = 3.778$, $p < .001$, but no main effect of Speech Presence, $\beta = .181$, $z = .809$, $p =.41$. As predicted, *representational goal* responses were higher when the actor used a grasping handshape than when she used a tracer handshape, and higher when objects were present than when they were absent. However, there was a significant interaction between Handshape and Speech Presence, $\beta = 1.62$, $z = 3.01$, $p <.01$. No other 2- or 3-way interactions were significant.

To explore the interaction between Handshape and Speech Presence, we conducted two independent-sample t-tests examining the presence or absence of speech, first for all participants who viewed a Grasping Handshape, and second for all participants who viewed a Tracer Handshape. For participants who viewed an actor using a Grasping Handshape, the presence or absence of speech did not affect the proportion of participants who described *representational goals*, $t(158) = 1.27$, *ns*. Observers were likely to attribute representational goals to the agent when the movement was performed by a grasping handshape, whether or not speech was present. In contrast, for participants who viewed an actor using a Tracer Handshape, the presence or absence of speech did affect the proportion of participants who described *representational goals*, $t(158) = 2.99$, $p < .01$. Observers were likely to attribute representational goals to the agent when the movement was performed by a tracer handshape only when speech was present.

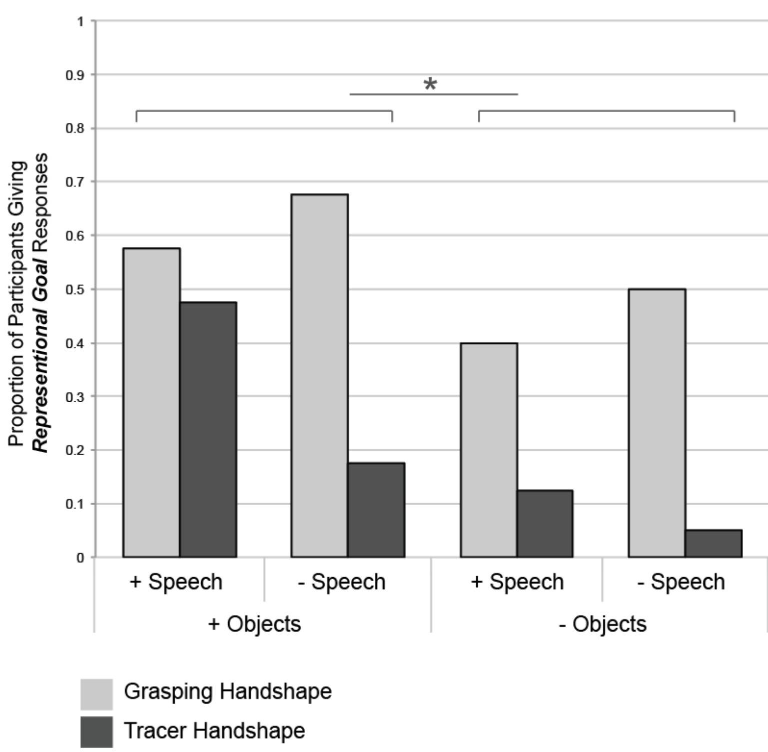

Figure 4. Proportion of participants giving Representational goal responses in each of the 8 conditions.

Together, our results confirm that when observers view movements that are not used to directly manipulate objects (empty-handed movements), multiple factors affect whether the movement is seen as a gesture. (1) The main effect of Object Presence replicates Study 1, and suggests that if empty-handed movements are performed in a meaningful context, in this case, in the presence of relevant objects, those movements are likely to be considered representations of acts. (2) The main effect of Handshape suggests that if empty-handed movements are performed with a meaningful handshape, one that resembles the handshape used when actually moving the objects, those movements are likely to be considered representations of acts. (3) The interaction between Handshape and Speech Presence suggests that if empty-handed movements are performed with a difficult-to-interpret handshape (in this case, a tracer handshape), those movements are more likely to be considered representations of acts if they are produced in the presence of speech, even if it is unintelligible. In contrast, when the handshape is meaningful (a grasping handshape), empty-handed movements tend be considered representations of acts with or without the presence of speech.

Discussion

When do adult observers interpret a movement as representation of movement—as a gesture— rather than as a movement *per se,* produced either to affect an object or for its own sake? Our findings support a framework in which observers view empty-handed movements as having representational goals if those movements can be interpreted as meaningful. In Study 1, we replicated previous work (Schachner & Carey, 2013) showing that movement produced on an object is likely to be interpreted as a goal-directed action on that object, whereas movement produced in the absence of an object is likely to be interpreted as movement for the sake of movement. Importantly, we also tested participants' interpretations of a third type of movement––movement produced off of an object (empty-handed movement), but in the presence of the object—and found that, under these circumstances, observers are likely to see the movement as a representation, that is, as a gesture for action, rather than as an action *per se.* Pushing the envelope further, in Study 2 we explored how contextual cues that make a movement appear more meaningful can work together to encourage observers to give representational responses for empty-handed movement. We found that observers are also likely to see an empty-handed movement as a gesture when an actor uses a handshape that resembles the handshape that would be used to manipulate the object as it is moved. If the actor uses a handshape that does not capture features of the hand as it manipulates the object (i.e., if it is a pointing index finger), the empty-handed movement will be interpreted as a gesture only if the movement is accompanied by speech-like sounds.

Our findings thus corroborate Schachner and Carey's (2013) claim that there are nuances to how movements are interpreted, and that observers do not inevitably see movement as having an external goal. The results further enrich the literature on action understanding by detailing a set of circumstances under which movement is likely to be interpreted as a representation of action (as gesture), rather than as action itself. The results also inform traditional view of gesture. Although gesture is often talked about as movements that accompany speech[3] (Hostetter & Alibali, 2008; Kendon, 2004; McNeill, 1992), our results emphasize that speech is not always necessary for a movement to be seen as representational. Instead, contextual factors and top-

[3] One type of gesture, called *emblems* (Ekman & Friesen, 1969, e.g., thumbs-up or the okay gesture), is often produced without speech; because emblems have standardized forms, they can be interpreted without speech. Our paper focuses on iconic gestures (*illustrators* in Ekman & Friesen's terms) that are tied to speech but, as we show here, can be recognized as a gesture even if processed without speech.

down processes can lead to the categorization of movement as representational, and therefore gesture. Our findings can be used to create a more general framework for understanding how empty-handed movements are interpreted, and have implications for education and for how the differences between gesture and other forms of movement are conceptualized.

How Context Drives Classification of Empty-Handed Movement. Using the cues to meaning from the present study as examples, we can explore a general framework for different types of cues to meaning, and consider how these cues can drive representational versus movement-based goal interpretations of empty-handed movements. More specifically, in this framework, the richer the context, the more likely an empty-handed movement will be interpreted as meaningful, and thus as representational, rather than as movement for the sake of movement. When interpreting another individual's movements, observers can attend to contextual cues arising from a variety of sources. (1) *Cues internal to the movement.* First, cues to meaning can come from properties of the movement itself (e.g., the shape of the hand as it moves, the path of the movement, its rhythm or speed). If these properties resemble properties of a movement that might actually be performed on an object, the movement is more likely to be seen as meaningful and thus representational. In the present study, handshape was this type of cue—handshape affords a way for an individual to interact with the external world and thus provides cues about that interaction. (2) *Cues external to the movement.* Second, cues to meaning can come from the environment within which a movement is produced. In the present study, the presence of objects was this type of cue. It is likely that an observer's prior knowledge of a particular object will affect her ability to use that object to interpret a movement. Adults know that balls can be picked up, and this knowledge likely drove their interpretation of the movements above the balls as representing actions to move these objects. If the agent had produced the same movements above novel objects (whose affordance was more difficult to guess from their appearance), observers might have been more inclined to attribute movement-based goals to the agent (e.g., *she's waving her hands above some strange looking items)*. (3) *Communicative cues.* Third, cues to meaning can come from features that signal that the movement is part of a communicative act. In the present study, speech was this type of cue (although participants did not see the face of the person in the videos, the speech-like stimulus was interpreted as coming from her). Other communicative cues can include the specific content of the speech, the agent's facial expressions and eye-gaze, or even just the presence of a listener. Finally, we can consider how these different types of cues can interact. In the present study, cues to meaning for the tracer hand (an ambiguous handshape cue internal to the movement) were enhanced when the movement was produced along with speech-like sounds (a communicative cue). Thus, a combination of different types of cues can make a movement appear more meaningful and, as a result, lead an observer to interpret the movement as representational.

We have focused primarily on how various contextual cues can make it more likely for an observer to interpret a movement as representing an external goal. However, the framework can also be applied to interpreting a movement as representing a movement-based goal. For example, if an individual points left and then right, and then right again (a cue internal to the movement), and this is the only cue present, it is likely that the movements would be interpreted in terms of a movement-based goal. But if the same movements were performed along with music (a cue external to the movement) and/or with the speech, "So I have to move left, then right, then right again" (a communicative cue), the movements would likely be interpreted in terms of a representational goal, in this case, representing movement for its own sake (e.g., she was demonstrating steps in a dance, cf. Kirsh, 2010).

Implications for education. Over a decade of research has shown that gesture can facilitate learning (e.g., Alibali & DiRusso, 1999; Perry, Church, & Goldin-Meadow, 1988; Ping & Goldin-Meadow, 2008; Singer & Goldin-Meadow, 2005; Valenzeno, et al., 2003; Wakefield & James, 2015), but the mechanisms underlying this powerful effect are still uncertain. One of the reasons gesture is thought to benefit learners is because it represents information about the to-be-learned concept in an easily accessible format (e.g., Goldin-Meadow, 2010). Gesture is assumed to represent concepts for the learner, as opposed to being mere hand-waving. But the learner may not always see it that way.

Our findings suggest that there are circumstances under which learners are more likely to see movement as gesture, and perhaps then profit from that movement. For example, in Studies 1 and 2, we saw that movements made near (but not directly on) objects were more likely to be interpreted as a representation than movements performed in the absence of objects. Classroom teachers naturally gesture near objects (Alibali & Nathan, 2012), and most of the gesture strategies used in experimental situations to investigate the utility of teachers' gestures have also been performed in reference to objects (Cook, Mitchell, & Goldin-Meadow, 2008; Goldin-Meadow, Cook, & Mitchell, 2009; Goldin-Meadow et al., 2012; Novack et al., 2014; Valenzeno et al., 2003). Our results suggest that part of the reason gesture may have helped children in previous studies is because the proximity of the gesture to the object it was referencing encouraged children to interpret the movement as a meaningful gesture.

Given our findings, educators should be informed that proximity to referents *does* matter when using gesture as a teaching tool. If proximity is not possible, other methods of making it clear that a movement is representing information may be able to be used. For example, Ping and Goldin-Meadow (2008) showed that children can learn from gesture when objects are absent;

however, before the children were shown the gestures, they had been familiarized with the objects about which the experimenter gestured, and thus could interpret the experimenter's movements in the context of those now-absent objects. Producing a movement in the presence of objects (or being familiar with those objects if they are not present) is likely to allow learners to glean meaning from that movement and thus interpret it as gesture. When a movement is produced in the absence of objects, learners may be more likely to see it as movement for its own sake, and thus not reap educational benefits from it.

A second cue that influenced whether a movement was interpreted as a representation was the handshape used by the actor in Study 2. Participants were significantly more likely to view an empty-handed movement as a representation if the actor used a grasping handshape during the movement than if she used a tracer handshape, whether or not objects were present. Although both grasping and tracer handshapes *can* render an empty-handed movement a representation (particularly if the tracer handshape is accompanied by speech-like sounds), the two handshapes are not interchangeable in the eyes of the participants. Paying attention to the form of an empty-handed movement (as well as its context of use) thus appears to be important in determining whether it will be interpreted as a gesture and therefore relevant to a lesson.

The relation between gesture and speech. The final factor we considered in Study 2 was the presence or absence of speech. We assume that if the speech that accompanies an empty-handed movement were to explicitly refer to that movement (e.g., "I'm moving the orange ball to the orange box"), the movement would be interpreted as a gesture. Our question was whether the presence of speech (rather than its content) would affect when an empty-handed movement is interpreted as a gesture. We therefore filtered the speech to retain its prosody but render it unintelligible. We predicted that the presence of these speech-like sounds would significantly increase the likelihood that an empty-handed movement would be interpreted as a gesture. However, we found this to be the case only when an ambiguous handshape was used in the movement. In other words, making an empty-handed movement seem more communicative in the context of speech-like sounds does not increase its likelihood of being considered gesture unless the movement is difficult to interpret in the first place.

We do know that the participants in our study saw the speech-like sounds as communicative—they thought that someone was trying to say something to them that they could not understand (e.g., “A woman’s voice was garbled, like she was underwater”). The fact that this communicative, albeit unintelligible, cue did not generally increase the likelihood that an empty-handed movement would be interpreted as a gesture (only when the gesture was ambiguous) suggests that the aspect of speech that drives individuals to interpret movement as meaningful gesture is not just the intent to communicate, but also the meaning of the spoken language itself. Neuroimaging results lend weight to this interpretation—using fMRI, researchers have found a different pattern of activation when participants process empty-handed movements that are produced with a spoken language they know, compared to the same movements produced with a language they do not know (Green et al., 2009). This finding has interesting implications for using gesture to teach a second language. We know that producing movements along with the sounds of a new language can help individuals learn that language (Kelly, McDevitt, & Esch, 2009; Macedonia, Muller, & Friederici, 2011), but it may be important to tell the learners that the movements are meaningful (i.e., are gestures) in order for them to be effective.

Conclusions

This chapter has shown that, under the right conditions, adults will view empty-handed movements as more than just movements for their own sake. Adults spontaneously view gesture-like movements as representational, suggesting that we may privilege gestures as a unique category of movement. However, these studies are only a first step in exploring the features of movement that make it likely to be interpreted as a representation. I am thus perfectly positioned to ask how the ability to see movement as representational action develops over ontogeny.

Chapter 2: How do children talk about gesture?

Results from Chapter 1 (hereafter, referred to as Novack, Wakefield, & Goldin-Meadow, 2016) indicate that people have distinct ways of processing object directed and empty-handed movements. But the main goal of this dissertation is to consider how seeing movement as gesture affects learning from gesture. This question is most applicable for school-aged children, who frequently see gesture in the classroom (e.g., Alibali & Nathan, 2012). In this chapter, we ask how children, ages 4-9, describe gesture-like movements.

The framework provided by Novack et al. (2016) can be used to understand how a mature thinker parses empty-handed movements into gestures versus movements-for-their-own sake, but there is reason to believe that this ability may look different in less mature thinkers. Based on previous work, we know that children *can* glean information from gesture (Broaders & Goldin-Meadow, 2010; Novack, et al., 2015; Sekine, Sowden & Kita, 2015; Stanfield, Williamson, & Özçalışkan, 2014), but neuroimaging literature has revealed a protracted development in children's ability to process gesture, specifically with respect to integrating gesture with contextual cues like speech. Specifically, researchers have found developmental changes in the posterior middle temporal gyrus (pMTG), a region implicated in speech-gesture integration in adults (Özyürek, Willems, Kita, & Hagoort, 2007; Willems, Özyürek, & Hagoort, 2007). Wakefield et al. (2013) showed that activation in the pMTG increases between the ages of 5 and 11 years during processing speech accompanied by iconic gesture, and Dick and colleagues (2011) also showed gesture-speech integration differences in 8-11 year olds versus adults. Taken together, the literature suggests that young children are sometimes sensitive to information in gesture, but it is not yet clear how contextual cues other than speech affect children's ability to interpret movement as representational. In the current study, we focus on another contextual cue

– the presence of related objects – to further explore the question of how children know when movements are "gestures".

Work on children's ability to process actions performed *on* objects (e.g., picking up a toy) suggests that humans have a systematic and uniform way of interpreting this form of movement. Habituation paradigms demonstrate that infants encode object-directed movements in terms of external goals (Woodward, 1998) and, at an older age, this understanding is reflected in children's verbal descriptions of vignettes (Trabasso et al., 1992). Neuroimaging work confirms that infants and adults show similar neural responses in electrodes over sensorimotor processing regions when viewing an actor reaching towards and grasping objects (de Klerk et al., 2015; Virji-Babul et al., 2012; Rotem-Kohavi et al., 2014), and that the action-observation network is predominately adult-like in childhood (Biagi et al., 2015). Thus, in contrast to the developmental changes that we might expect to see in children's interpretation of empty-handed movements, the action literature suggests that we can expect a uniform interpretation for movements that involve manipulating objects.

The purpose of the present study is to further explore how children interpret empty-handed movements across development, and whether providing a richer context affects the likelihood that children will think of movement as representational, rather than movement for its own sake. Here, we focus on children's explicit descriptions of movement, rather than using their behavioral responses or brain activation during a task as an indirect measure of their ability to interpret empty-handed movement as gesture. We manipulate whether empty-handed movements are done in the presence or absence of objects as a test of contextual effects. We expect that, like adults (Novack et al., 2016), children will be more likely to interpret movements as representational if they occur in the presence of objects, but that, in both cases, the likelihood of representational interpretations of empty-handed movement will grow across development. Given that children show a protracted development in their ability to integrate one particular contextual cue – speech – with gesture, we may find an interaction, where context has more of an effect in later childhood than it does in younger children. However, most of the previous literature has focused on spoken language as a cue, and this cue is more dynamic and complex than static objects, thus we may find no interaction, indicating that children are able to integrate this simpler contextual information equally well across development. Although it is well established that actions *on* objects will be processed uniformly across development, we include this form of movement in the present study as well. Findings that age has no effect on how actions on objects are interpreted will show that children of all ages understand the current paradigm.

Method

Participants. Usable data were collected from 339 children between the ages of 4 and 9 years (142 females, 197 males) at a large, science museum. Parents provided informed consent, and children provided verbal assent. Children were relatively evenly distributed across the age range (4 yrs: *n*=70; 5 yrs: *n*=58; 6 yrs: *n*=50; 7 yrs: *n*=59; 8 yrs: *n*=56; 9 yrs: *n*=46), although birthdates were collected and exact age was calculated and used as a continuous variable in all analyses. Children were randomly assigned to one of three conditions with a target of ~20 children of each age group in both the *Action on Objects with Objects Present* Condition and the *Action with Objects Absent* Condition and a target of ~10 children per age group in the *Action on*

Objects Condition[1]. An additional 64 children were recruited for the study, but excluded from analyses for a number of reasons: refusing to respond to the initial prompt (*n*=51), providing a response to the initial prompt that was completely unrelated (e.g., mentioning events from popular children's movies; mentioning things happening nearby in the museum) (*n*=10), saying they could not remember or did not understand the stimulus the prompt was about (*n*=2), or providing a response that was inaudible during coding (*n*=1). The task took 2-3 minutes for children to complete, and they received stickers or small erasers for participating.

Stimuli. Stimuli included the same three videos shown to adults in Study 1 of Novack et al., 2016 (see Figure 1 in Chapter 1). Specifically, this included videos of 1) *Action on Objects*, 2) *Action off Objects with Objects Present* and 3) *Action with Objects Absent*. All videos were silent.

Procedure. Children were invited to participate in the present study while they were visiting a science museum. The study was conducted at a small table located in a relatively quiet part of the museum, and children sat next to the experimenter facing a wall to decrease distractions. At the beginning of the study, children were told they would watch a very short movie and then be asked what had happened in the movie. They were also told that the movie had no sound so they had to pay very close attention. Children then watched one of three 10-second movies on a 9.7-inch display iPad, depending on the condition to which they had been randomly assigned. When the movie ended, the experimenter asked the child, "What happened in the movie?" followed by additional prompts that were designed to further probe children's interpretation of the movie. For children in the *Action on Objects* or *Action off Objects with Objects Present* conditions, the experimenter asked, "So at the end of the movie, were the balls

[1] Pilot testing showed almost no variability in responses for the *Action on Objects* condition, as such, we limited our sample size for this condition as a larger sample size would be unnecessary.

in the boxes or out of the boxes." This question was asked to determine whether children could correctly remember the end-state of the movie they watched, where the correct answer in the *Action on Objects* condition was that they were inside the boxes, whereas the correct answer in the *Action off Objects with Objects Present* condition was that they were out of the boxes, on the table. To be consistent, a similarly structured prompt was asked in the *Action with Objects Absent* condition: "So at the end of the movie, were her hands on the table or off of the table?" Having focused on the end-state of the movie, the experimenter again probed children's interpretation of the scene by asking, "So tell me one more time, what happened in the scene?" And finally, after children responded, the experimenter asked, "Did she do it for a reason?" For children who answered "yes", the experimenter asked "What reason?" Children's responses were audio recorded for coding purposes. After completing the study, children were thanked for their participation and got to choose stickers or small, animal-shaped erasers to take home.

Coding and Reliability. Children's responses to the two instances of the prompt, "What happened in the movie?" were classified into goal-based categories based on an adapted version of the coding scheme used by Novack and colleagues (2016, see chapter 1). Although the prompt does not specifically probe the intentions of the woman in the stimuli, the pragmatics of the question, combined with the fact that movements were being produced voluntarily, should invite descriptions of goals and intentions (see Schachner & Carey, 2013, for discussion of assumptions about intentionality). The codes are described below:

(1) *External Goal:* The movie is described in terms of actions completed on objects, with the description focusing on movement of objects, rather than movement of hands (e.g., "she took balls and put them in boxes"; "the blue balls goed in the blue container and the orange balls goed in the orange container"; balls were placed in boxes"). In rare

cases, this also included attempted actions on objects "it looked like he was trying to put the balls in the box").

(2) *Movement-Based Goal:* The movie is described in terms of low-level spatiotemporal movements without mentioning a higher-level goal – the description is focused on the movement of the hands themselves (e.g., "a guy had his hands going back and forth"; "she was like, moving her arms"; "someone moved their hand from a ball to the block and back to the ball").

(3) *Representational Goal:* The movie is described in terms of movements representing external goals (e.g., "the person was imitating putting the balls in the boxes"; "he was pretending to put balls in the container"; "she looked like she was playing a piano in the air").

(4) *Other:* The movie is described (a) without mentioning movement at all (e.g., "there was a hand", "It showed a rectangle that was orange and two orange balls two blue balls and a blue rectangle"), or (b) mentioning movement, but the response was too ambiguous to assign a goal-oriented code (e.g., "it did um some marbles and um I saw some box"; "he was doing one by one").

Two researchers assigned a single code to each response, and were blind to the condition of each participant. Coders agreed on 638 of 678 trials (94.1%), $K = 0.93$. Any disagreements were discussed between the coders and resolved.

A separate coding system was used for the final set of prompts, "Did she do it for a reason? What reason?" As we were especially interested in children's ability to interpret empty-handed movement as representational, we coded whether responses to these questions for children in the *Action off Objects with Objects Present* and *Action with Objects Absent* conditions suggested they interpreted the movement as representational. Responses were coded as *Representational* if children described the woman as having a goal to represent information through her actions (e.g., "to show us what he was gonna do with the balls into the blue boxes"; "To show little kids how to match colors"). Responses were coded as *Nonrepresentational* if children said they were not sure or thought the woman did not have a reason for her movements, or if their response suggested the woman had a movement-based goal (e.g., "To like practice for something she was going to get tested for") or was not goal-directed (e.g., "Um, because she can't talk".

As with coding the main prompts, two researchers assigned a single code to each response, and were blind to the condition of each participant. Coders agreed on 261 of 267 trials (97.8%, $K = 0.93$). Any disagreements were discussed between the coders and resolved.

Results

Action on Objects. An established body of literature shows that the ability to interpret actions on objects in relation to external goals develops in infancy (e.g., Woodward, 1998). Thus, it was unsurprising that pilot data showed practically no variability in responses elicited to the *Action on Objects* movie, regardless of age, and we expected children in our final sample would have a very systematized and uniform way of describing this form of movement. In line with our pilot work, we found that almost all responses to the *Action on Objects* condition were coded as external goal responses (97.2%), and a binomial logistic regression, predicting the likelihood of an external goal response by age, confirmed that there was no effect of age ($\beta = 0.39$, $SE = 0.46$, $z = 0.84$, $p = 0.40$).

Empty-handed Movement. In contrast to the responses elicited by the *Action on Objects* movie, children did *not* show a uniformed way of interpreting empty-handed movements in the *Action on Objects with Objects Present* and *Action with Objects Absent* conditions (see Figure 5). Children provided both representational goal and movement-based goal responses to the *Actions with Objects Absent* movie, although movement responses were much more common than representational responses (83.7% vs. 12.6%). Children were most varied in their responses if they were in the *Actions off Objects with Objects Present* condition: 42.4% described movement-based goals, 22.0% described representational goals, and surprisingly, 20.5% described external goals. This same movie also had a larger proportion of 'Other' responses; 15.2% of the responses, compared to 3.0% in the *Action with Objects Absent* condition.

To determine how responses to the two empty-handed movement conditions compared to each other, we ran a log-linear poisson model on a 2 (Condition) x 4 (Response Code) contingency table, which allowed us to compare across non-independent response codes. Our model revealed that responses in the *Action off Objects with Objects Present* condition were significantly different from responses in *Action with Objects Absent* condition (χ^2, 3 = 64.57, p < .001), suggesting that the two types of movies elicited distinct patterns of responses with children age 4-9, just as they do with adults (Novack et al., 2016, see also chapter 1).

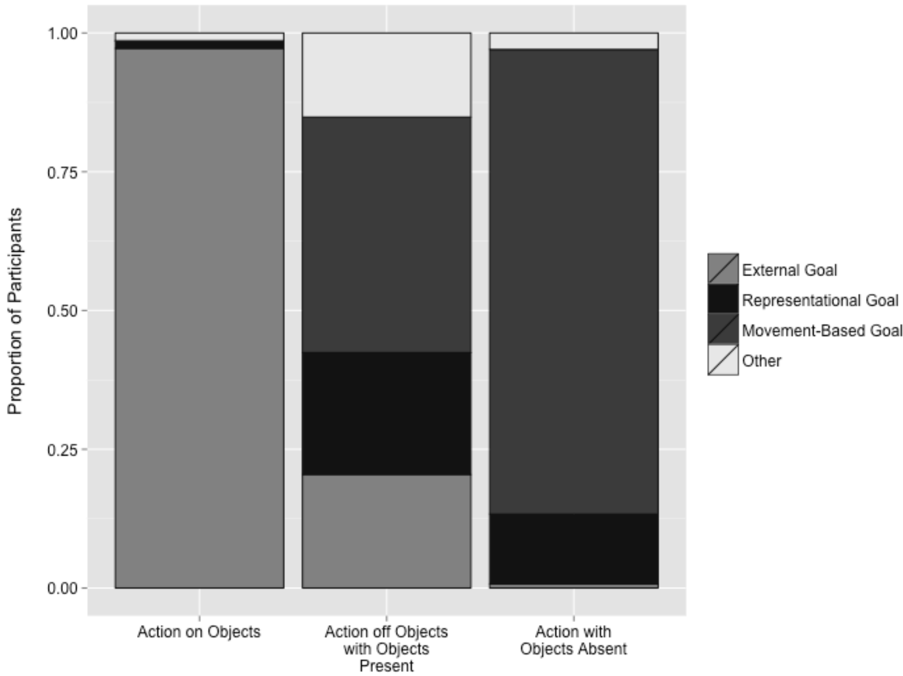

Figure 5. *Proportion of participants in each condition giving External goal, Representational goal, Movement-based goal, or 'Other' responses.*

Empty handed movement as representational. Having established differences in general patterns of movement interpretation, our second question was to ask how the ability to interpret different forms of movement might change across development. Our main objective was to understand the development of children's ability to interpret empty-handed movements as representational actions, and how their interpretations are affected by contextual cues (the presence or absence of objects). To address this question, we used a binomial logistic regression to determine whether children's likelihood of providing a representational response was predicted by age and condition (*Action off Objects with Objects Present; Action with Objects Absent*). The model revealed a positive effect of age ($\beta = 0.36$, $SE = 0.10$, $z = 3.50$, $p < .001$), whereby children became more likely to give a representational response as they got older, as well as an effect of condition ($\beta = 0.71$, $SE = 0.34$, $z = 2.16$, $p < .05$), whereby children were more likely to provide a representational response when objects were present than when they were absent. Given the possibility that children are more reliant on object cues when they are younger, we asked whether there was an interaction between condition and age, but the model did not reveal a significant interaction ($\beta = 0.07$, $SE = 0.21$, $z = 0.32$, $p = .75$).

Our findings were, overall, in line with our predictions that (1) children's ability to understand empty-handed movement as representational develops with age, and (2) a richer context supports representational goal responses. However, we were surprised at the overall low rates of these representational responses across the sample (17.2%). These findings may suggest that children between the ages of 4 and 9 rarely interpret empty-handed movement as representational. Alternatively, it is possible that children, whose language abilities are less developed than adults, simply need more opportunity to express their competency. We explored this possibility using data from the additional prompts.

Recall, after children gave their initial response to the prompt, "What happened in the scene?" they were asked about the ending-state of the movie. They were then asked a second time, "What happened in the scene?" and, "Did she do it for a reason? What reason?" Here we assess whether children ever attributed a representational goal response across these three prompts. Specifically, for this analysis, we considered children to have given a representational goal response if they provided a representational goal response to at least one of the first two prompts (as opposed to an external goal, movement-based goal, or 'Other' response), and/or if they provided a representational goal response to the final prompt (as opposed to a nonrepresentational response).

Based on data from these three prompts, we found that 31.1% of our sample provided at least one representational response across the study. This almost doubles the proportion of children describing empty-handed movements as representational, and suggests that the initial low response rates were not caused by an inability to see movement this way. Compared to their responses to the first prompt, 23.5% (31 of 132)[2] of children in the *Action off Objects with Objects Present* condition and 12.6% (17 of 135) of children in the *Action with Objects Absent* condition changed their interpretation over the remaining prompts[3]. It was much more common for children who changed their response to provide external goal, movement-based goal, or 'Other' responses initially, and subsequently provide a representational goal on the second or third prompt, than for a child to begin with a representational goal response and then deviate from that response: 87.1% of children (27 of 31) in the *Action off Objects with Objects Present* condition and 94.1% of children (16 of 17) in the *Action with Objects Absent* condition followed the predominant pattern. Interestingly, the older children were, the more likely they were to change their interpretation ($\beta = 0.26$, $SE = 0.10$, $z = 2.63$, $p < .01$), and across age, it was more common for children to alter their interpretation in the *Action off Objects with Objects Present* condition than in the *Action with Objects Absent* condition ($\beta = 0.80$, $SE = 0.34$, $z = 2.38$, $p < .05$).

[2] An additional six children changed their interpretation across prompts, but not in a meaningful way. For example, a child might provide an 'Other' response initially (e.g., "balls and boxes") and then a 'Movement' response (e.g., "She moved her hands") on the second prompt. Because the child did not change between a nonrepresentational and representational understanding of the prompt, she would not be counted in this analysis.

[3] To further show that children's interpretation of movement in the *Action on Objects* condition was stable, we also asked whether children changed how they interpreted movement in this condition. Only two children altered their interpretation, and these were the only children who had not initially provided an external goal response. Both children provided an external goal response on the second prompt, thus 100% of children ultimately interpreted action on objects in terms of an external goal.

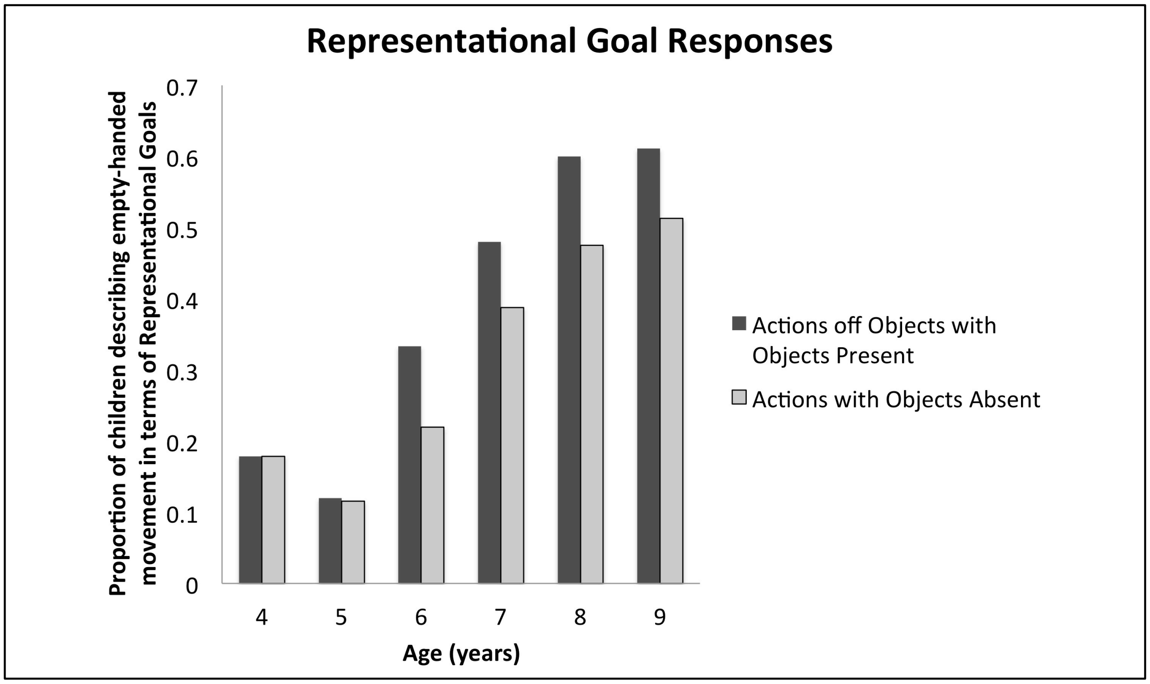

Figure 6. Proportion of participants in each age group giving at least one Representational Goal response during the experimental session, separated by condition.

We determined how age and condition predicted children's likelihood to *ever* provide a representational response across prompts (see Figure 6). As with the model based on their first responses, we found a significant effect of condition ($\beta = 0.74$, $SE = 0.28$, $z = 2.60$, $p < .01$), and age ($\beta = 0.41$, $SE = 0.09$, $z = 4.80$, $p < .001$), and no significant interaction between these predictors ($\beta = 0.17$, $SE = 0.17$, $z = 1.01$, $p = 0.31$).

Empty handed movement as action. Surprisingly, unlike adults (Novack et al., 2016), it was not uncommon for children to provide external goal responses to the *Action off Objects with Objects Present* condition. These responses were structured as if the actress in the movies had actually completed an action (e.g., "She moved balls into boxes") despite the fact that no competed actions occurred (i.e., the woman never physically touched or moved the objects). This raises two possibilities: (1) that children actually encoded movement *off* objects as movement *on* objects, or (2) that children's language abilities are not as fine-tuned as adults, and they may use an external-goal structure to describe movement.

To test the first possibility, we considered how children in the *Action off Objects with Objects Present* condition responded to our second prompt, "*So at the end of the movie, were the balls in the boxes or out of the boxes.*" Of the 27 children (20.5% of the sample) who provided an external goal response to the initial prompt, 24 (88.9%) correctly identified the balls as being outside of the boxes at the end of the movie. This response rate was comparable to the number of children who gave the correct answer across the other response types (82 of 93; 88.2%); thus, it is unlikely that children providing an external goal response encoded the movement as an action on an object.

To test the second possibility, we reasoned that language abilities improve across development; finding an effect of age on the likelihood of giving an external goal response to the *Action off Objects with Objects Present* Condition would thus suggest that children favored a response of this type because of less sophisticated language abilities. However, a binomial logistic regression, with external goal response (1,0) as the dependent measure and age as a predictor, revealed no effect of age ($\beta = 0.02$, $SE = 0.13$, $z = -0.15$, $p = .88$), suggesting that children did not become *less* likely to provide an external goal response as they became older and more sophisticated in their language abilities. Still, it is possible that it is easiest for children to structure their responses in terms of external goals. Thus, we can consider whether children decreased the number of external goal responses on the second prompt when given more time to think and explain themselves. On the initial prompt, 20.5% (27 of 132) of children provided an external goal response. On the second prompt, this percentage was reduced to 14.4% (19 of 132). It may be that asking children to consider whether balls were inside or outside the box helped them recognize how their initial response had not been an accurate reflection of the movements in the movie.

Discussion

In line with previous research, we found that children's ability to describe actions on objects is uniform across ages 4-9. In contrast, children's ability to describe actions off objects is much more varied, and undergoes significant change across this developmental period. First, children across all ages were less likely to describe empty-handed movements in the absence of objects as representational than empty-handed movements in the presence of objects. This result is similar to the pattern found in adult's responses from Novack et al. (2016, see also chapter 1), suggesting that, in both childhood and adulthood, the presence of objects that can be acted upon supports interpretation of hand movements as meaningful. Children's descriptions of empty-handed movements in the absence of objects were uniformly split between movement-based goals and representational goals. This result is also similar to the adult responses, indicating some amount of individual difference in seeing movement as meaningful over the lifespan.

The most interesting and varied set of responses were found for the videos of actions produced off objects but in the presence of objects. In addition to the two types of responses that adults had provided for these videos (movement-based goals and representational goals), children provided a significant number of uncodable responses and external-goal responses. Both of these response types, which were not found in the adults in Novack et al (2016), reflect the idea that this form of movement was challenging for children to think about and talk about. The external-goal responses are particularly interesting – in this case children were actually describing, in concrete language, an event that did not occur (the movement of balls into boxes).

Our prompting questions confirm that children did not actually *think* that the balls had moved – children were unlikely to say that, at the end of the movie, the balls were inside the boxes in response to our prompt questions. Rather, the concrete-language may reflect a challenge in *talking* about representational actions. In fact, in unpublished pilot work testing this paradigm in an English-Language-Learning classroom in a local public school third grade class, we found an even higher incidence of external-goal responses in description of these videos, suggesting that this concrete language may reflect a challenge in describing representational actions.

Importantly, describing gesture-like movements with an external-goal response indicates that the child does understand the purpose of the movements. This type of response describes what the gestured movement is *supposed* to represent, even if it does not contain language indicating that the movements are representational. As such, children who provide external-goal responses to the empty-handed movements do seem to understand the meaning of the movement. Future work may consider whether a child's ability to *talk* about representational actions and their ability to *comprehend* representational actions is related, or whether they may require distinct skills, with the later more dependent on linguistic skills, and the former more dependent on representational processing skills.

Finally, the variability in responses to gesture-like movement raises the question of whether individual differences in seeing movement as representational may be related to individual differences in individual's ability to learn from gesture. This question will be explored in greater detail in the general discussion.

Chapter 3: Categorizing Movement in Infancy

Chapters 1 and 2 found that adults, and even children as young as four think differently about actions on and off objects. Next we ask what infants think about different categories of movements whether a system for processing representational goals may already be in place in the first years of life. This question is particularly interesting with respect to infants because research indicates that before the age of two, children do not recognize the relation between iconic gestures and their referents[1]. For example, 17-month-olds fail to recognize the iconicity in gestures, and are no more likely to learn an iconic gesture-action pairing than an arbitrary gesture-action pairing (e.g., Namy, 2008; Namy, Campell, & Tomasello, 2004). In contrast, by 24 months, infants show superior learning from iconic gesture over arbitrary gesture (e.g., Namy, 2008; Namy, Campell, & Tomasello, 2004), suggesting that around age two, infants begin to develop the decoding skills necessary to interpret representational forms (see DeLoache, 1995). Results from other studies find converging evidence - around age two, children are beginning to learn how to interpret the meaning of iconic gestures, although this ability matures in the third and fourth year (e.g., Goodrich & Hudson Kam, 2007; Marentette & Nicoladis, 2011; Namy, 2008; Namy, Campell, & Tomasello, 2004; Novack, Goldin-Meadow, & Woodward, 2015; Tolar, Lederberg, Gokhale, & Tomasello, 2008).

Here, we ask - what do infants *think* gestures are, prior to developing the ability to correctly interpret and understand them. Infants are exposed to iconic gestures well before the age of two - on average parents produce about two iconic gesture during a typical 90 minute

[1] Iconic gestures are visual representations of referential meaning (McNeill, 1992). The gestures discussed in chapter 1 were all *iconic* gestures because they used movement (path) and in some cases hand shape (grasping) to represent features of the action to which they refer. The statement that infants do not understand iconic gestures before the age of 2 is specific to *iconic gestures*, and is not the case for *all* gestures. Considerations of how these questions might apply to other types of gestures (e.g., *deictic* gestures) will be explored in the discussion.

interaction with their 18-month-olds (Özçalışkan & Goldin-Meadow, 2005), a number that can be extrapolated to around 15 iconic gestures per day during an infant's waking hours. Clearly, infants are *seeing* iconic gestures produced by their parents – so what do infants think about them? Do infants recognize gestures as representational of *something* and view them as a unique category, distinct from either actions on objects or meaningless movement? Are there properties of gesture (such as being performed near, but not on, an object) that encourages young children to categorize it as a representation, even if they cannot decode the representation? At just 12-months, infants seem to know that speech can communicate even if they cannot understand that speech (Vouloumanos, Onishi & Pogue, 2012). Therefore it is possible that at 18 months, infants may know that gestures are meant to *represent* without being able to understand *what* they represent. Alternatively, prior to developing the ability to understand gestures, infants may just see gesture as mere hand-waving – similar to empty-handed movements produced for its own sake (cf., Schachner & Carey, 2013).

We can first consider what is known about infants' understanding of other's actions. As is the case with adult's understanding of other's actions, the majority of research on infants' understanding of other's actions has been conducted within the framework of external goals and actions on objects. For example, in a seminal study, Woodward (1998) demonstrated that when infants see a person reach toward an object, they interpret that event in terms of the relation between the person and the object, making inferences about the person's goals and intentions. More specifically, Woodward (1998) habituated 5-month-old infants to an event in which a person reached out and grabbed one of two objects (See Figure 5a). Infants were shown this event at least 6 times, during which their looking time to the event decreased, indicating habituation. The experimenters then switched the location of the two objects (out of view of the infant) and then showed infants one of two alternating test events. In novel-object/familiar-reach events (Figure 5b) the person reached to the same location, grasping the object that had not been touched during habituation (i.e., using a *familiar reaching path* to obtain a *novel object*). In familiar-object/novel-reach events (Figure 5c) the person reached to the same object that they had grasped during habituation trials, even though this required moving along a new reaching path (i.e., using a *novel reaching path* to obtain the *familiar object*). Thus, both types of test events involve something novel (either the goal object, or the reaching movement).

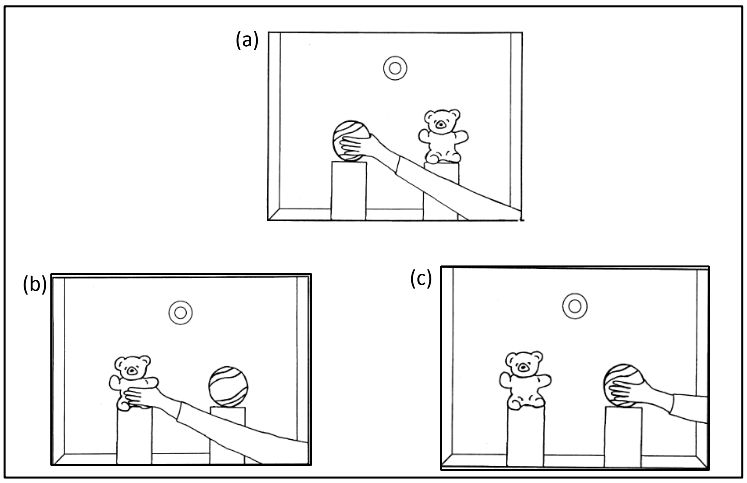

Figure 7. Example of stimuli used object-directed habituation task, taken from Woodward, 1998. Figure 7a shows a sample habituation event. 7b and 7c show test events in which the location of the two objects is switched. 7b demonstrates a novel-object/familiar-reach event, 7c demonstrates a novel-reach/familiar object event.

In this type of paradigm, infants prefer to look more to a novel-object/familiar reach test event than a familiar-object/novel-reach test event, suggesting that infants prioritize objects when interpreting other's object-directed reaches, and are more interested in a change in object, than a change in spatiotemporal patterns (Woodward, 1998). This phenomenon has been replicated numerous times in both habituation experiments (e.g., Buresh & Woodward, 2007; Henderson & Woodward, 2012; Woodward, 1999) and anticipatory eye tracking studies (Cannon &

Woodward, 2012; Krogh-Jespersen & Woodward, 2014). Thus, when infants see a person act on objects, they encode the action in terms of the relation between the agent and the objects and ignore the spatiotemporal movements used to achieve the object.

However, there are situations in which infants will prioritize the spatiotemporal movement of an action. For example, after being shown a movement in which a starfish does jumping jacks while moving over a stationary ball, 7-month-olds prefer to look at a video that shows a change in manner (starfish doing toe-touching while moving over the ball) or a change in path (starfish doing jumping jacks while moving *under* the ball) indicating that they can attend to either of these movement-based features (Pulverman, Song, Pruden, Golinkoff, & Hirsh-Pasek, 2013). Similar results can be found in behavioral studies with older children. Carpenter, Call, and Tomasello (2005) found that 12 and 18-month-olds will imitate an experimenter's movements, such as the zig-zag movements of a toy mouse hopping across a mat, suggesting that infants can attend to spatio-temporal movement patterns under some conditions. However, if the experimenter produces the same movement with a concrete end-goal (i.e., they use the zig-zag movement of the mouse to put him into a toy house), then children will ignore the movements and copy only the end goal (putting the toy mouse into the house without copying the zig zag). Taken together, these previous studies suggest that infants are able to attend to movement patterns, particularly if the movement is the only relevant feature of an event.

To summarize, the existing literature suggests that when infants see agents acting on objects, they encode the movement in terms of goals and ignore the spatial properties of movement itself. In contrast, when infants see agents producing intransitive actions off objects, or moving with no clear external goal, infants encode and attend to the spatial properties of a movement itself. An open question is how infants will interpret *gestures*. Specifically, we can ask how infants will interpret actions produced *off* objects, but in the presence of objects, the kind of movement that adults systematically view in terms of *representational goal* (Novack, et al., 2016). One possibility is that, to an infant, all actions produced in the air may be seen as meaningless movements. Infants may be less likely than adults to integrate contextual cues, such as the presence of objects that can be acted upon, into their interpretation of a movement. Since infants under the age of 2 are not yet proficient at understanding the meaning of iconic gestures, they may fail to see how a grasping gesture relates to a present object, in which case they may ignore the objects and focus exclusively on the hand movements. Another possibility is that infants may already have in place a system for categorizing movements as representational, using the same types of cues that adults use (e.g., presence of objects, speech). If that is the case, the infants should distinguish between empty-handed movements produced in the presence of objects, and the absence of objects.

To explore this question, we used eye-tracking and looking time measures to ask how infants view movements on and off objects. We created videos inspired by the stimuli used in Novack, et al. (2016) and developed a paradigm loosely based off Woodward (1998). Infants were familiarized to either to a woman acting on objects (*Action on Objects)*, moving in the presence of objects without touching them (*Action off Objects with Objects Present*) or moving in the absence of objects (*Action with Objects Absent)*. After seeing the same event a number of times, infants were shown a split screen preferential looking test with two types of events. As was the case in Woodward (1998) the location of the objects (for the two conditions that included objects) was switched in both events. For infants in the *Action on Objects* condition, one side of the screen showed the woman reaching toward the same object that she reached for during familiarization (which consequently involved moving along a novel path, since the object had changed locations) (a novel-object/familiar reach event). The other side of the screen showed the woman reaching toward the novel object, but using the same reaching movement used in familiarization (familiar-object/novel reach). The videos were similar for infants in the *Action off Objects with Objects Present* condition, except that, as in the familiarization events, the woman never made contact with any of the objects, she just *pretended* to reach for the objects. Finally, infants in the *Action with Objects Absent* condition saw a woman reaching either in the same path, or a novel path (with no objects present).

Based on previous research, we expect infants in the *Action on Objects* condition to encode the video in terms of the relation between the woman and her goal (See Woodward, 1998). Therefore, after being familiarized to an event in which a woman reaches for one of two objects, infants should prefer, during the preferential looking test event, to spend more time looking to the event of the woman reaching to a novel object (i.e., novel-object/familiar reach event) rather than to the video showing the woman reaching in a novel path (i.e., novel-reach/familiar object event).

The design of the *Action with Objects Absent* condition makes the opposite prediction about infants' looking patterns. That is, in the absence of any objects, we expect infants to encode a movement in terms of the movement patterns itself. Therefore, after being familiarized to an event in which a woman reaches along one path (e.g., to the right) infants should prefer, during the preferential looking test event, to spend more time looking to the video showing the woman reaching along a novel path (i.e., novel-reach event) rather than the video showing the woman reaching along a familiar path (i.e., familiar reach event).

The crucial question is how infants in the *Action off Objects with Objects Present* condition will respond. If infants, like adults, use contextual cues to interpret empty-handed movements, then they should know that the hand movements are not the goal of the movement, and should not attend specifically to the hand movements themselves. If they understand that the movements *represent* reaching toward one of the objects, they should, in theory, look longer when the agents starts producing movements that represent reaching toward a *novel* object. In contrast, if infants process all empty-handed movements the same, regardless of the context, then they may see the *Action off Objects with Objects Present* condition just as hand movements. If that is the case, then they should look identical to infants in the *Action with Objects Absent* condition, and are expected to look longer when the agent reaches in a novel path during the looking test.

Study 1

Method

Participants. Forty-eight 18- to 20-month-old infants participated in the current study. Infants were racially and ethnically diverse, recruited from a database in a large Midwestern city, and tested individually in a laboratory setting. Sixteen infants were randomly assigned to the *Action on Objects* condition (8 females; M_{age} = 19.17), 16 to the *Action off Objects with Objects Present* Condition (8 females; M = 19.16), 16 to the *Action with Objects Absent* Condition (8 females; M = 18.91). All infants were full term (minimum 37 weeks gestation).

Stimuli. In all videos, the event started showing the torso of a seated woman resting her right hand in the middle of a table. In two of the conditions (*Action on Objects* and *Action off Objects with Objects Present*), the two objects - a plush green cylinder, and a cardboard purple triangular prism were positioned to her right and left. In each video, the woman raised her right hand directly into the air, and then reached out, either toward the right or the left, then pulled her hand back into the center. In the *Action on Objects* condition, she would grab the object as she did this movement, bringing it toward her body (See Figure 8a). In the *Action off Objects with Objects Present* condition she produced movements *as if* she was grabbing the object, but without ever physically touching it. (See Figure 8b). Finally, in the *Action with Objects Absent* condition, she produced the same movement but just in space – there are no objects present. (See Figure 8c). The hand movements produced in each of the three videos were designed to be as identical as possible and filmed using a metronome to keep timing. Thus, the only difference between the videos is the presence of the objects, and whether or not the woman interacts with the objects.

The woman always reached with her right hand, however the side she reached, toward (right, left), the object she reached toward (green, purple), and the side that the object was on, were all counterbalanced across conditions. This resulted in four orders for each of the *Action on Objects* and *Action off Objects with Objects Present Conditions*, and two orders for the *Action with Objects Absent* condition.

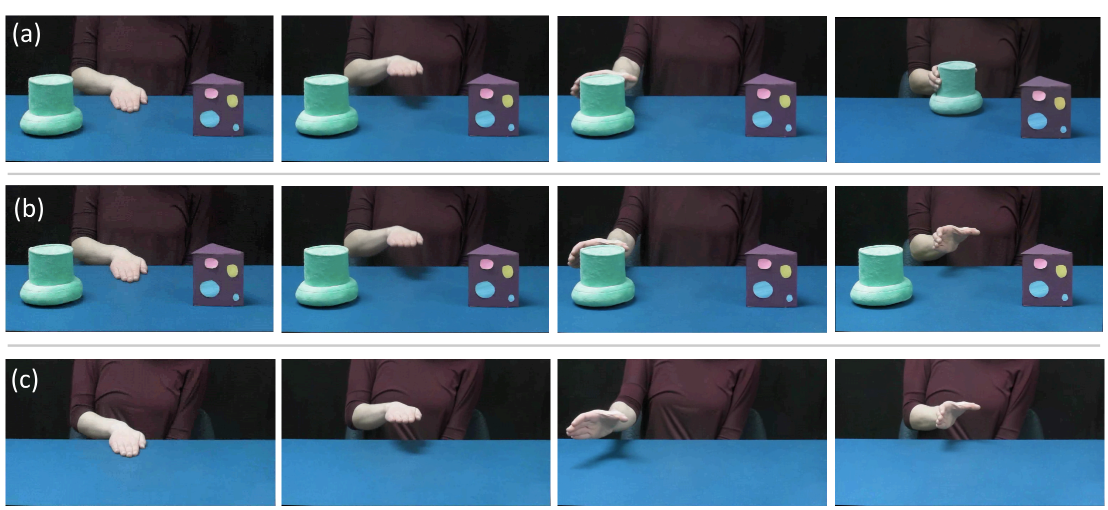

Figure 8. Stills taken from familiarization stimuli in the three conditions in Study 1: (a) A*ction on Objects*, (b) *Action off Objects with Objects Present*, (c) *Action with Objects Absent.*

Familiarization videos were used to create the preferential looking test videos (see Figure 9). During the preferential looking test, two videos were presented side-by-side, showing a novel-object/familiar-reach and a familiar-object/novel-reach event. For the *Action on Objects* and *Action off Objects with Objects present* conditions, the location of the two objects was switched from what had been shown during familiarization (i.e., if green had been on the right and purple on the left, in test events purple was on the right and green was on the left). In the *Action with Objects Absent* events these movements were just in space (i.e., novel movement, familiar movement). The preferential test showed the same reaching movement video repeated 3 times. Movements in the two videos were timed so that they were identical. The entire length of the looking time test lasted approximately 18 seconds.

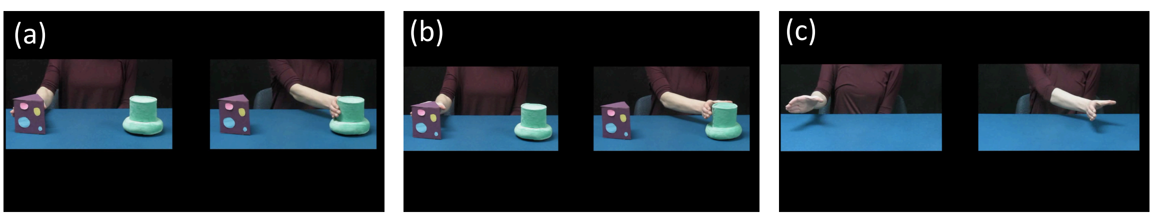

Figure 9. Stills from split-screen preferential looking test in the three conditions in Study 1: (a) *Action on Objects*, (b) *Action off Objects with Objects Present*, (c) *Action with Objects Absent*.

Procedure. Infants were randomly assigned to one of the three conditions. Infants sat on their parents' laps in front of a 24-inch monitor equipped with a Tobii T60XL corneal reflection eye-tracking system (accuracy 0.5u, sampling rate 60 Hz). Calibration was performed with a 9-point procedure using the standard animation of a duck provided by the Tobii software within the infant calibration setting. When necessary, the calibration process was repeated to improve accuracy. Data were collected using Tobii Studio (Tobii Technology, Sweden) and a video recorder positioned at the top of the eye-tracker also captured all infant looking. Infants saw 3 identical familiarization events, an anticipation test event[2], three more familiarization test events, and then the looking test.

Coding. Due to technical difficulties, we did not collect sufficient eye tracking data from most participants. Instead, we coded infants' gaze during the preferential looking test from a video recording, using the coding system ELAN (Sloetjes & Wittenburg, 2008). Two trained coders determined, on a frame-by-frame basis, whether the infant was looking to the right or left of the screen, or off screen. Coders agreed on 96% of individual frames. Differences of more than a few frames were reviewed and discussed. Data from one coder was extracted for analysis.

Results

Data were extracting during a 16 second portion of the preferential looking test event starting with the moment the actor's hand begins to reach in either direction, and continuing until the videos start to fade. Initial analyses found no effect of infant age or gender on any looking time measures; these factors were therefore dropped from subsequent analyses. Figure 10 shows the average total time infants spent looking to the screen during the test event across the three conditions. Infants in the *Action with Objects Absent* condition looked for an average of 13064 ms, significantly less than infants in the *Action off Objects with Objects present* condition, who looked for 15827ms ($\beta = 2763.6$. $t = 3.050$, $p < .01$) and less than infants in the Action off Objects with Objects Present condition, who looked for 15358 ms ($\beta = 2294.3$, $t = 2.532$, $p < .01$).

[2] In between the familiarization videos, there was a 10-second anticipation test event in which the location of the objects was switched (like the Preferential looking event), but paused when the hand was raised off the table (before reaching toward either object). This test was intended to collect anticipatory saccades to the object that infants *expected* the person to reach to. Given data collection issues, we did not collect sufficient eye tracking data from infants to look at the anticipation test event.

This difference likely stems from the fact that there were no objects in the Movement videos, making them less interesting. There was no difference in the amount of looking between infants in the *Action on Objects* and *Action off Objects with Objects Present* conditions.

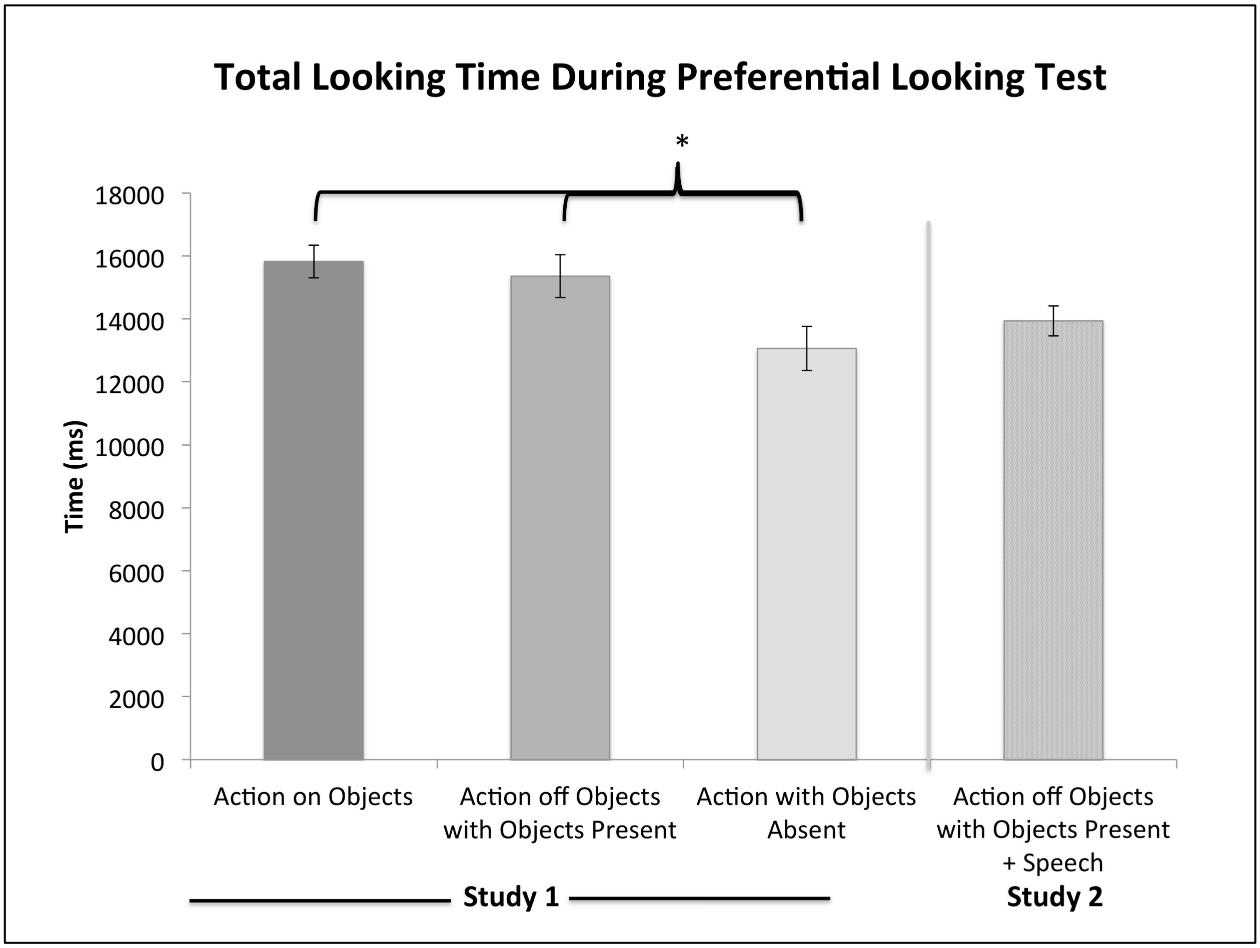

Figure 10. Mean time infants spent watching the preferential looking test event across the three conditions. Asterisk indicates $p < .001$

Next, we calculated each infant's relative preference for one of the videos during test. We calculated each infants' preference for the novel-object/familiar-reach event by dividing the time spent looking to the novel-object/familiar reach event by the total amount of time that infant watched the test portion (time novel-object/familiar-reach + novel-reach/familiar-object). Proportions for each infant (which reflect a relative preference for novel-object/familiar reach events) were then averaged within each condition. Thus, an individual, or group level average proportion of .5 indicates no preference for either video. A proportion above .5 indicates a preference for the novel-object/familiar reach video. A proportion below .5 indicates a preference for the novel-reach/familiar object-video. Figure 11 shows the mean proportion of infants' preference for the novel-object/familiar-reach event by condition. A regression using condition to predict the arcsin transformed proportions of their average preference for novel-object/familiar-reach events found that the effect of Condition did not reach significance, $F(2, 16) = 1.71$, $p = 0.19$. Individual one-tailed t-tests found that infants in the *Action on Objects* condition spent marginally more of their time looking at the novel-object/familiar reach condition $t(15) = 1.50$, $p = .07$. Infants in the *Action off Objects with Objects Present* and *Action with Objects Absent* Conditions did not systematically prefer one video over the other (p's > .4)

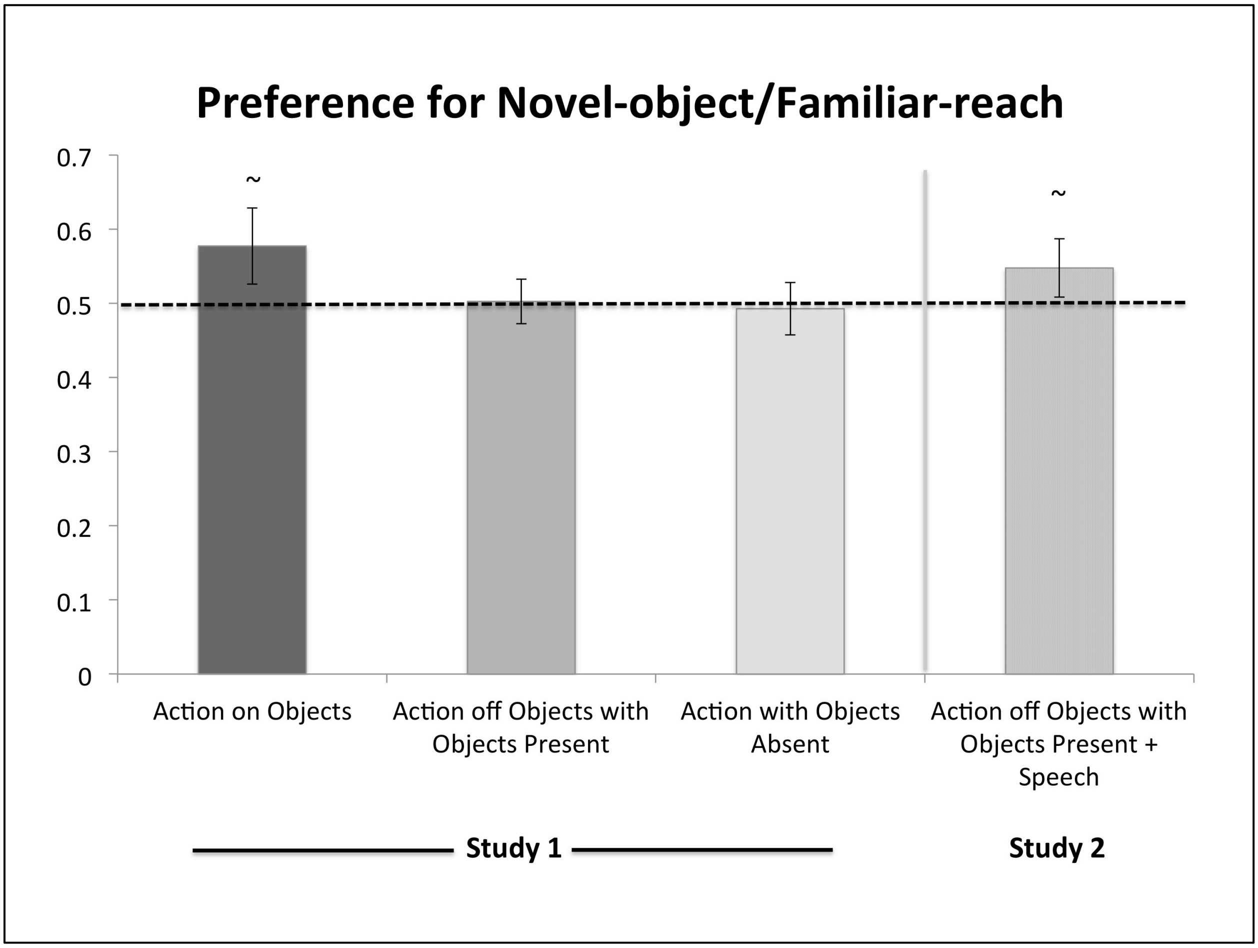

Figure 11. Average amount of time looking to the Novel-object/Familiar Reach event, divided by total looking during Preferential Looking test. The dotted line indicates equal preference for novel-object/familiar reach and familiar-object/novel reach events. Tilda indicates marginal effect, (e.g., $p < .1$)

Finally, we categorized infants' responses on an individual level, depending on whether they showed a preference for the novel object (a proportion greater than .5) or the novel movement (a proportion less than .5). Table 1 shows the distribution of infants' relative preferences during the preferential looking test across the three conditions. A one-sided binomial test confirms that significantly more infants in the Action on Objects showed a preference for the novel-object/familiar-reach event ($p < .01$), which was not the case for the other two conditions ($p = .4$ for both).

Table 1 *Number of infants exhibiting preferences for novel-object or novel-reach events during looking test phrase*

	Preference for Novel Object/Familiar Reach event	Preference for Familiar Object/Novel Reach event	Binomial *p* value
Action on Objects	13	3	$p < .01$
Action off Objects with Objects Present	9	7	*ns*
Action with Objects Absent	7	9	*ns*

Thus far, it appears that infants interpret actions *on* objects in one way (preferring to look at novel-object events at test) and interpret actions *off* objects in another way (showing no preference for novel-object or novel-movement events at test). One interpretation of these results is that infants see all actions off objects in the same way. However, we noted something unique about infants in the *Action with Objects Absent* condition. Infants in this condition appeared to look more at the contralateral reach in the test, regardless of the reach they had been shown during familiarization. Indeed, 15 of 16 infants in the *Action with Objects Absent* condition preferred the contralateral reach in the test event (binomial test: $p < .001$). Given the counterbalanced nature of our design, looking more toward the contralateral reach at test meant that half of the infants showed a novel-reach preference at test, and half showed a familiar-reach preference (relative to what they saw during familiarization), which resulted in a group-level preference of .49 (no different from chance). A preference for the contralateral reach is likely caused by the fact that the contralateral reach is a larger movement, and revels a greater amount of the actor's arm, two features which increase the salience of the movement. Yet, despite the fact that these two salient features were also present in the other two conditions, infants in the other conditions were not systemically bothered by it. Eight of 16 infants (p=.59) in the *Action on Objects* Condition, and 11 of 16 infants (p=.1) in the *Action off Objects with Objects Present*

Condition looked longer at the contralateral reach in the test event, irrespective of what they saw in familiarization. Therefore, despite the fact that, on average infants in both of the empty-handed movement conditions showed equal preference for the two events during the preferential looking test, this equal preference may stem from different sources. Infants in the *Action with Objects Absent* condition were specifically focused on the person's movements, uniformly attending to the salient movement at test, whereas infants in the *Action off Objects with Objects Present* condition were not.

Overall, these results suggest a few things. First, infants process actions-on-objects in terms of the relation between an agent and the object. They do not encode the features of the movement used to achieve an object-directed action. Rather, they notice a change in the object being reached to. Second, infants who see actions in the absence of objects find it difficult to encode anything about the movement. Instead, they look toward the more salient movement when given a preferential looking choice. Possible methodological reasons for this effect will be considered in the general discussion. Finally, infants who see actions *off* objects in the presence of objects fail to encode either the object being gestured toward, or the more salient movement at test, suggesting that they process gesture-like movements differently from instrumental actions and from meaningless-movements.

Study 2

Results from Study 1 suggest that 18-month-old infants struggled to encode anything useful about empty-handed movements in the presence of objects. They did not systematically process the event in terms of the relation between the agent and the object, nor did they attend to the specific movements. This pattern suggests that interpreting gesture-like movement is particularly challenging for young children.

One possibility, however, is that at this age, infants need additional supportive cues to be able to make sense of movements. Novack, et al., (2016) identified three features which increased the likelihood that adults would view a movement as representational – the presence of objects, a grasping handshape that could be used to carry out the action, and the presence of speech-like sounds. The videos in Study 1 contained two of those features (present objects and a grasping handshape). In study 2, we ask whether adding the third feature – speech – changes how infants watch actions off objects with objects present. In study 2, we added an audio track to the *Action off Objects with Objects Present* familiarization videos in Study 1 to determine whether the presence of speech affects the interpretation of empty-handed movements. We reasoned that 18-month-olds, who are just beginning to understand the social world, may require additional cues, in this case, meaningful speech, to make sense of empty-hand movements.

Methods

Participants. Thus far, 8 additional 18- to 20-month-old infants have participated in Study 2 (*M* = 19.04 months, range: 18.06–19.36 months, all boys). We plan to test an additional 8 subjects (8 females) to complete the sample.

Stimuli. Stimuli in Study 2 were identical to those used in the *Action off Objects with Objects Present* condition in Study 1, with the change that the familiarization videos included an audio track of a woman saying, “This one’s mine”. The timing of the audio was edited such that it synched with the reaching action. There was no audio track in the looking test videos.

Procedure and Coding. The procedure and coding were identical to study 1.

Preliminary Results

Data from infants' looking patterns during the preferential looking test indicate that infants in this *Action off Objects with Objects present* condition show a marginal preference for the novel-object/familiar reach ($M_{prop} = .54$, $t(7) = 1.20$, $p = 0.13$). This preliminary trend suggests that adding meaningful speech to actions produced off objects, but in the presence of objects, may change the way 18-month-olds watch those events (see Figure 9).

Discussion

Previous research has demonstrated that, before their first birthday, infants seem to know a lot about instrumental-actions such as object-directed reaches. If an infant sees someone reach toward on object, the infant attends to the agent's goal and ignores the specific movements used to achieve that goal (Woodward, 1998). In the current study, we broadened this line of inquiry to explore how infants process hand movement that do *not* achieve goals – empty-handed movements, either in the presence or absence of objects.

Our results indicate that when the contact between an agent's reach and an object is disrupted, infants no longer process the event in terms of external goals. Although infants in the *Action on Objects* condition showed a novelty preference for an event in which an agent reached to a new object, infants in the two empty-handed movement conditions - *Action off Objects with Objects Present,* and *Action with Objects Absent* – showed no preference for either event. Surprisingly, infants in the *Action with Objects Absent* condition systematically preferred to look at the more salient contralateral reach during test, irrespective of the event they had watched during familiarization. This unexpected finding may indicate that the reaching movements we used during familiarization were not salient enough on their own for infants to make sense of them. Previous looking time studies demonstrating that infants *can* attend to movement features used much more salient, full body movements (i.e., a starfish doing jumping jacks or bending at the waist) (e.g., Pulverman, Song, Pruden, Golinkoff, & Hirsh-Pasek, 2013; Pulverman, Golinkoff, Hirsh-Pasek, & Buresh, 2008). The hand movements we chose may have been glossed as "reaches" by the infants and not as "reach to left" or "reach to right," as we had intended.

Importantly, not all empty-handed movements lead infants to focus on salient features like the contralateral reach during the test phase. Infants who watched the *Action off Objects with Objects Present* videos did not show systematic preferences for novel object events, novel-reach events, or even the contralateral reach event. This is surpassing since the movements were identical to the empty-handed movements in the *Action with Objects Absent* condition. Thus, it appears that objects that are not acted upon are a confusing and challenging cue for 18-month-olds.

Moreover, when we added meaningful speech to the empty-handed movements (Study 2), infants' looking pattern began to shift. Infants in the *Action off Objects with Objects Present + Speech* condition (when meaningful speech was added to the familiarization event) began to look more similar to infants in the *Action on Objects* condition. Although we cannot draw solid conclusions from preliminary results, should the pattern hold, it would suggest that communicative cues that help adults see movements as meaningful also help 18-month-olds make sense of empty-handed movements. Taken together, the results from the current studies provide preliminary evidence that gesture-like movements, produced in the presence of objects and with meaningful speech, are processed more similarly to instrumental actions, than to movement-for-its own sake.

The results from the current study raise questions about whether other types of gestures might be differentially meaningful to infants. For example, around 12 months, infants understand pointing gestures as being goal-directed (Woodward & Guajardo, 2002). That is, after being habituated to an agent pointing toward one of two objects, infants show a novelty preference for an event in which an agent points to a novel object (but not if an agent points to a familiar objects in a new location). (Woodward & Guajardo, 2002). Infants even understand that pointing gestures are communicative, and expect that points should lead the observer of a point to understand something about the pointer's intentions (Krehm, Onishi, & Vouloumanos, 2014). Thus, we would predict that, if a pointing condition were added to our design, infants should look at a test event in which the actor points to a novel object. If this prediction is confirmed, we would need to explain why infants can succeed with pointing gestures before they can succeed with iconic reaching gestures. Pointing gestures are communicative and representational, much like iconic reaching gestures (see Novack & Goldin-Meadow, 2016 for discussion of the ways in which points can be considered representational). However, by 18 months, infants are proficient pointers themselves, whereas children do not reliably produce iconic gestures until 26 months (Özçalışkan & Goldin-Meadow, 2011). It may be that infants need experience producing and understanding iconic gestures before they can recognize them as having representational goals.

Finally, the results in the current study raise questions about how and when infants begin to learn from iconic gestures. My own work has found that, around 24 months, infants can learn novel ideas from iconic gesture instruction, although at this stage their ability to learn from iconic gestures is not as robust at their ability to learn from instrumental-actions (Novack, Goldin-Meadow, & Woodward, 2015). Although, on average, 2-year-olds can learn from gesture, their learning rates are low. It would therefore be interesting to determine whether individual variation in infants' looking patterns in the *Action off Objects with Objects Present* condition predict variability in the infants' learning from gesture instruction.

In general, the current chapter suggests that, at 18-months, infants are not yet expert gesture comprehenders, or even categorizers, but that they may be in the beginning stages of seeing gesture-like movements as representational Next, we turn to childhood – a range between infancy and adulthood – to ask how and when the ability to think about (and talk about) representational actions develops.

General Discussion

We know that, in general, including gesture in instruction supports learning (Novack & Goldin-Meadow, 2015). However, until now, children have been assumed to know when hand movements are gestures. This dissertation is a first step toward determining whether an inclination to see movement as gesture (i.e., as representational action) predicts learning from gesture in instruction. The studies outlined here have investigated how humans across the lifespan come to think about, and talk about, gesture-like movements.

Chapter 1 tested how adults describe actions *on* objects compared to actions *off* objects. We found that adults systematically describe movements as depicting an object-directed action when the actor moved objects, and favored describing the movements as depicting movement for its own sake when the actor produced the same movements in the absence of objects. However, participants tended to describe movements as *representations* when the actor produced the movements near, but not on, the objects. We also showed that the more context adults are given, the more likely they are to shift from seeing movement as meaningless to seeing movement as representational, suggesting that cues such as the presence of objects, the form of the movement, and the presence of speech-like sounds encourage adults to see movement as representational.

Chapter 2 explored how school-aged children develop the ability to think about and talk about different kinds of movements. We found consistency in the way that children, from age 4 to 9, talked about instrumental object-directed actions. In contrast, empty-handed movements elicited significant variability across all ages. We also found that children's ability to describe movement as representational increases across development. Children were likely to describe gestures in terms of concrete-completed actions, suggesting that children may struggle in their verbal descriptions of representational movements.

Finally, Chapter 3 tested whether an inclination to view the actions on objects and off objects as distinct is present in infancy. Results showed that infants do differentiate between actions on objects and actions off objects, privileging the status of objects as goals when an agent reaches and touches them, but not when an agent gestures a reaching movement without touching them. Importantly, 18-month-olds do not process all empty-handed movements the same. Empty-handed movements in the absence of objects draw infants' attention to low-level perceptual features of movements, while empty-handed movements in the presence of objects do not. Finally, the presence of meaningful speech may change the way infants process and understand empty-handed movements in the presence of objects, helping them see those movements as meaningfully related to objects.

Taken together, the results from the three chapters provide a stable story across the lifespan whereby humans have distinct ways of processing instrumental movements that interact with objects, compared to empty-handed movements. Gesture, which is action off objects, seems to elicit more confusion and variability than either actions on objects, or actions in the absence of objects, suggesting that its status as a representational form may present special challenges for observers. The results also suggest that an ability to make sense of gestures increases with age, supporting the proposal that gestures are representational actions and thereby require representational processing abilities to be understood.

Although the questions explored in these three chapters are interesting in their own right, they lead to future questions about how these ideas can be applied to learning contexts. Building on the current studies, future work may now be able to ask how an individual's inclination to see movement as representational affects that individual's ability to learn from gesture. Does providing a representational-goal response to one set of gesture-stimuli (for example, our *Action* *off Objects with Objects Present* videos) indicate that a child has reached the milestone of being able to see gestures as representational more generally? If that is the case, then children who describe our stimuli in terms of representational-goals should also be likely to learn from gesture in an instructional context. However, it might also be the case that responses in our task reflect an ability - in the moment - to see movement as meaningful. In other words, children's ability to learn from gesture during instruction might depend on their ability to view that *specific* instructional gesture as representational, and not just any gesture.

This possibility brings into focus the question of whether learners even need to be aware of the representational status of a gesture in order to benefit from that gesture during instruction. A learner's ability to profit from a gestural movement might be related to that learner's ability to categorize that movement as meaningful. Alternatively, learners may not need to be explicitly aware of gesture's representational properties to be able to benefit from it during instruction. For example, Thomas and Lleras (2009) suggest that body movements can support implicit learning without explicit awareness. In their work, adult participants were asked to solve a classic insight problem, Maier's two string problem, which requires the solver to figure out how to tie two strings attached to the ceiling which are too far apart to grab both at the same time. The solution requires using a regular household object (e.g., a spoon) as a weight for a pendulum, and then swinging one string to the other side. Participants who performed an arm swinging motion during a purportedly unrelated "stretch break", were more likely to solve the insight problem than participants who performed a different stretching motion. Again, the adults who were successful were not aware that the arm movements they produced were related to their problem solving success, suggesting that, at least in some cases, learners do not need to consciously see a movement as meaningful in order to learn from it.

Brooks & Goldin-Meadow (2015) found similar effects with children learning how to solve novel math problems. They found that children who were asked to produce relevant hand motions (i.e., sweeping their hands under the two sides of an equations, indicating that the two sides should be equal) but not irrelevant motions (i.e., sweeping their hands in an up-and-down motion near the problem) were more likely to learn from subsequent instruction, although these children showed no awareness of the meaning of the movements while producing them (i.e., were unlikely to recognize correct solutions to problems). The fact that children were able to integrate the ideas from the gestures later, during instruction, but not while producing them, indicates that they may not have seen those gestures as meaningful in the moment, raising the possibility that an awareness of gesture as representational action may not be necessary for children to learn novel ideas.

Another open question is whether movement is categorized as gesture in the same way for perceiving vs. producing movement. The studies in this dissertation all explored situations when the *perceivers* of a movement see the movement as representational. However, it is unclear whether the same features lead *producers* of a movement to see the movement as representational. This question is particularly relevant in tasks where learners are taught to produce movements during a lesson (e.g., Goldin-Meadow, et al., 2009). These movements are meaningless to the learner at the beginning of the lesson. The question is whether these movements become meaningful, and therefore representational, during the lesson and, if so, when? Do children think of these hand movements as "gesture" when they are initially taught them, or do they think of them first as "movement-for-its-own sake" and only gradually come to see the movements as "gesture" as their conceptual understanding of the lesson shifts? If the process is gradual, might there be markers or features within the movement itself that an observer could use to determine when a rote movement becomes a true gesture?

Another open question related to the issue of categorizing movement as gesture or instrumental action is whether there are in-between cases. For example, Clark (1996) identified a class of movements called *demonstrations*—actions produced with the intention of *showing* something to someone. For example, if a mother were to show her child how to open a jar, she could hold the jar out in front of the child, twist open the lid in an exaggerated manner, and then put the lid back on the jar, handing the jar to the child to try the action himself. This object-focused movement has elements of an instrumental action—the mother's hands directly interact with the object and cause a physical change. However, the movement also has obvious elements of representational actions—in the end, the jar is not open and the movement is clearly performed for communicative (as opposed to purely instrumental) purposes. As another example, consider "hold-ups"—gestures in which someone holds up an object to display it to someone else (e.g., a child holds up her bottle to draw it to her mother's attention). Hold-ups have some aspects of gesture—they are intended to communicate and are like deictic pointing gestures in that they indicate a particular object. But they also have aspects of instrumental actions—they are produced directly on objects. Developmentally, hold-ups tend to emerge before pointing gestures (Bates, Camaioni, & Volterra, 1975), lending credence to the idea that hold-ups may not be as representational as empty-handed gestures. One important question is whether hold-ups function like gestures for a child. It turns out that they do, in at least in one sense—they predict the onset of various aspects of spoken language. For example, hold-ups have been counted as deictic gestures in studies finding that early gesture predicts the size of a child's subsequent spoken vocabulary (Rowe & Goldin-Meadow, 2009), the introduction of particular lexical items into a child's spoken vocabulary (Iverson & Goldin-Meadow, 2005), and the developmental onset of noun phrases, Cartmill, Hunsicker & Goldin-Meadow, 2014).

Also, there is reason to ask whether an inclination to see movement as representational and to learn from gesture-like movements is a static individual trait. Adults process actions differently based on their experience producing those actions. For example, expert hockey players process hockey language differently than hockey fans (Beilock, et al., 2008), and expert ballet dancers show differential brain activation to watching ballet, compared to other dance forms (Calvo-Merino, et al., 2005). Could this mean that people who naturally gesture a great deal learn differently, or are more likely to see gesture as meaningful than those who gesture less? Research on individual variation in gesture production and its relation to categorizing movement as gesture, as well as learning from gesture, might provide answers to this question.

Finally, there are interesting open questions about the bottom-up and top-down processes involved in classifying a movement as an instrumental action, gesture, or meaningless movement. In the current study, we identified bottom-up features of movement that lead individuals to see it as gesture, such as the presence of objects that can be acted upon or the presence of speech. But there may be more fine-grained features of movement that coincide with either instrumental action or representational action that may influence how we categorize movement. Gestures have the ability to be a simpler, more abstracted version of the actions they represent. For that reason, gestures may have an identifiable set of movement characteristics that allow them to be classified. At the same time, recognizing a movement as a type of movement can also involve top-down processing. For example, the same movement (twirling one's hands) can become a gesture or a dance, depending on the accompanying language ("he ran up the stairs" versus "watch me dance!"). In the real world, we likely engage in a combination of bottom-up and top-down processes to make these rapid decisions. Future research examining how these aspects interact would be interesting and informative.

Conclusions

Research on how we understand action across the lifespan has a long and rich history, but until recently was narrow in scope, focusing predominately on actions used to manipulate objects. This dissertation expands the discussion of how humans accomplish the task of interpreting the actions of others by investigating actions that are not used to manipulate objects. I have shown that these understudied, empty-handed actions have the capacity to be interpreted either as movement for the sake of movement (Schachner & Carey, 2013) or as a representation. This distinction is not trivial—when movements are seen as representations, they have the power to influence communication, learning, and cognition in ways that movement for its own sake does not (Novack & Goldin-Meadow, 2016). By incorporating gesture into a framework for movement analysis, we take an important step towards developing a more cohesive understanding of action-interpretation.

References

Alibali, M., W., & DiRusso A. A., (1999). The Function of gesture in learning to count: More than keeping track. *Cognition, 14*, 37-56.

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology, 25*, 468-523. doi: 10.1006/cogp.1993.1012

Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Church, R. B., Jacobs, S. A., Maritinex, C. J., & Knuth, E. J. (2014). How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction, 32*, 65-100. doi: 10.1080/07370008.2013.858161

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences, 21*, 247-286. doi: 10.1080/10508406.2011.611446

Atit, K., Resnick, K., Shipley, T. F., Ormand, C. J., Manduca, C. A., Goksun, T., & Tikoff, B. (2013). Spatial gestures point the way: A broader understanding of the gestural referent. *Proceedings of the 35th Annual Cognitive Science Society*.

Baird, J. A. & Baldwin, D. A. (2001). Making sense of human behavior: Action parsing and intentional inference. In B.F. Malle, L.J. Moses, & D.A. Baldwin (Eds.). *Intentions and intentionality: Foundations of social cognition* (pp. 193-206). Cambridge, MA, US: The MIT Press.

Baldwin, D. A., & Baird, J., A. (2001). Discerning intentions in dynamic human action. *Trends in Cognitive Sciences*, *5*, 171–178. doi:10.1016/S1364-6613(00)01615-6

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development, 72*, 708-717. doi: 10.1111/1467-8624.00310

Bates, E., Camaioni, L., & Volterra, V. (1975). The Acquisition of Performatives Prior to Speech. *Merrill-palmer Quarterly of Behavior and Development, 21*, 205–226.

Beaudoin-Ryan, L., & Goldin-Meadow, S. (2014). Teaching moral reasoning through gesture. *Developmental Science, 17*, 984-990. doi: 10.1111/desc.12180

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. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 13269–73. doi: 10.1073/pnas.0803424105

Biagi, L., Cioni, G., Fogassi, L., Guzzetta, A., Sgandurra, G., & Tosetti, M. (2015). Action observation network in childhood: a comparative fMRI study with adults. *Developmental science*. doi: 10.1111/desc.12353x

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology, 136*, 539-550.

Bower, G. H., & Rinck, M. (1999). Goals as generators of activation in narrative understanding. In S. R. Goldman, A. C. Graesser & P. van den Broek (Eds.), *Narrative Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso* (pp. 111-134): Erlbaum.

Broaders, S. C., & Goldin-Meadow, S. (2010). Truth Is at Hand How Gesture Adds Information During Investigative Interviews. *Psychological Science*, *21*, 623-628. doi: 10.1177/0956797610366082

Brooks, N., & Goldin-Meadow, S. (2015). Moving to learn: How guiding the hands can set the stage for learning. *Cognitive Science, online first*, 1-19. doi: 10.1111/cogs.12292

Buresh, J. S., & Woodward, A. (2007). Infants track action goals within and across agents. *Cognition*, *104,* 287-314. doi: 10.1016/j.cognition.2006.07.001

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

Cannon, E. & Woodward, A. L. (2012). Infants generate goal-based action predictions. *Developmental Science, 15*, 292-298. doi: 10.1111/j.1467-7687.2011.01127.x

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science, 8*, F13-F20. doi: 10.1111/j.1467-7687.2004.00385.x

Cartmill, E.A., Hunsicker, D., & Goldin-Meadow, S. (2014). Pointing and naming are not redundant: Children use gesture to modify nouns before they modify nouns in speech. *Developmental Psychology, 50*, 1660-1666. doi: 10.1037/a0036003

Casile A, Giese M A (2006): Non-visual motor learning influences the recognition of biological motion, *Current Biology, 16*, 69-74.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index on transitional knowledge. *Cognition, 23*, 43-71. doi: 10.1016/0010-0277(86)90053-3

Church, R. B., Kelly, S., & Holcombe, D. (2014). Temporal synchrony between speech, action and gesture during language production. *Language, Cognition and Neuroscience*, *29*, 345-354. doi:10.1080/01690965.2013.857783

Church, R. B., Kelly, S. D., & Wakefield, E. M. (2015). Measuring gesture. In D. Matsumoto, H. C. Hwang & M. G. Frank (Eds.), *APA Handbook of Non-Verbal Communication*.

Clark, H. H. (1996). *Using language.* Cambridge, GB: Cambridge Univ. Press.

Congdon, E. L., Novack, M. A., Brooks, N. B., Hemani-Lopez, N., O'Keefe, L. & Goldin-Meadow, S. (2016). *Better Together: Simultaneous Presentation of Speech and Gesture in Math Instruction Supports Generalization and Retention.* Manuscript submitted for publication

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, *84*, 1863–1871. doi: 10.1111/cdev.12097

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*, 1047–1058. doi: 10.1016/j.cognition.2007.04.010

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS One, 8*, e57410. doi: 10.1371/journal.pone.0057410

DeLoache, J. S. (1995). Early Understanding and Use of Symbols: The Model Model. *Current Directions in Psychological Science*, *4*, 109–113. doi: 10.1111/1467-8721.ep1077240

Dick, A. S., Goldin-Meadow, S., Hasson, U., Skipper, J., & Small, S. L. (2009) Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Human Brain Mapping, 30*, 3509-3526. doi:10.1002/hbm.20774.

de Klerk, C. C., Johnson, M. H., & Southgate, V. (2015). An EEG study on the somatotopic organisation of sensorimotor cortex activation during action execution and observation in infancy. *Developmental cognitive neuroscience*, *15*, 1-10. doi:10.1016/j.dcn.2015.08.004

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

Goldin-Meadow, S. (2010). When gesture does and does not promote learning. *Language and Cognition, 2*, 1-19. doi:10.1515/LANGCOG.2010.001

Goldin-Meadow, S., & Brentari, D. (in press). Gesture, sign and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*. doi:10.1017/S0140525X15001247

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009) Gesturing gives children new ideas about math. *Psychological Science, 20*, 267-272. doi: 10.1111/j.1467-9280.2009.02297.x

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. doi: 10.1037/0022-0663.91.4.720

Goldin-Meadow, S., Levine, S. L., Zinchenko, E., Yip, T. K-Y, Hemani, N, & Factor, L. (2012). Doing gesture promotes learning a mental transformation task better than seeing gesture. *Developmental Science*, *15*, 876-884, doi:10.1111/j.1467-7687.2012.01185.x.

Goldin-Meadow, S., & Sandhofer, C. M. (1999). Gestures convey substantive information about a child's thoughts to ordinary listeners. *Developmental Science, 2*, 67-74. doi: 10.1111/1467-7687.00056

Goodrich, W., & Hudson Kam, C. L. (2009). Co-speech gesture as input in verb learning. *Developmental Science, 12*, 81-87. doi: 10.1111/j.1467-7687.2008.00735.x

Green, A., Straube, B., Weis, S., Jansen, A., Willmes, K., Konrad, K., & Kircher, T. (2009). Neural integration of iconic and unrelated coverbal gestures: A functional MRI study. *Human Brain Mapping, 30*, 3309-3324. doi: 10.1002/hbm.20753

Henderson, A.M.E., & Woodward, A. (2012). Nine-month old infants generalize object labels, but not object preferences across individuals. *Developmental Science, 15*, 641-652. doi: 10.1111/j.1467-7687.2012.01157.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

Iverson, J. M. & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*, 368-371. doi: 10.1111/j.0956-7976.2005.01542.x

James, K. H. (2010). Sensori-motor experience leads to changes in visual processing in the developing brain. *Developmental Science, 13*, 279-288. doi: 10.1111/j.1467-7687.2009.00883.x

Kastens, K. A., Agrawal, S., & Liben, L. (2007). Research methodologies in science education: The role of gestures in geoscience teaching and learning. *Journal of Geoscience Education: Research Methodologies in Science Eductaion column*.

Kelly, S. D., Healy, M., Ozyurek, A., & Holler, J. (2014). The processing of speech, gesture, and action during language comprehension. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-014-0681-7

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes, 24*, 313-334. doi: 10.1080/01690960802365567

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press.

Kimura, D. (1973). Manual activity during speaking–I. Righthanders. *Neuropsychologia, 11*, 45-50. doi:10.1016/0028-3932(73)90063-8

Kirsh, D. (2010). Thinking with the body, in (eds) S. Ohlsson R. Catrambone. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, Austin, TX: Cognitive Science Society. 2864–2869. doi: 10.1145/2037296.2037303

Krehm, M., Onishi, K. H., & Vouloumanos, A. (2014). I see your point: infants under 12 months understand that pointing is communicative. *Journal of Cognition and Development, 15*, 527-538. doi: 0.1080/15248372.2012.736112

Krogh-Jespersen, S., & Woodward, A.L. (2014). Making smart social judgments takes time: Infants' recruitment of goal information when generating action predictions. *PLoSONE,* e98085. doi:10.1371/journal.pone.0098085

Macedonia, M., Muller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping, 32*, 982-998. doi: 10.1002/hbm.21084

Marentette, P., & Nicoladis, E. (2011). Preschoolers' interpretation of gesture: Label or action associate? *Cognition, 121*, 386-399. doi: 10.1016/j.cognition.2011.08.012

McNeill, D. (1992). Hand and mind: What gestures reveal about thought. Chicago: University of Chicago Press

Namy, L. L., Campbell, A. L., & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: a U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development, 5*, 37-57. doi: 10.1207/s15327647jcd0501_3

Namy, L. L. (2008). Recognition of iconicity doesn't come for free. *Developmental Science, 11*, 841-846. doi: 10.1111/j.1467-7687.2008.00732.x

Newtson, D. (1976). Foundations of attribution: The perception of ongoing behavior. In J. H. Harvey, W. J. Ickes & R. F. Kidd (Eds.), *New directions in attribution research* (pp. 223-248). Hillsdale, N. J.: Erlbaum.

Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychological Science, 25*, 903-910. doi: 10.1177/0956797613518351

Novack, M., & Goldin-Meadow, S. (2015). Learning from Gesture: How Our Hands Change Our

Minds. *Educational Psychology Review, 27,* 205-412. doi: 10.1007/s10648-015-9325-3

Novack, M. A., & Goldin-Meadow, S. (2016). Gesture as Representational Action: A paper about function. *Manuscript submitted for publication*

Novack, M. A., Goldin-Meadow, S., & Woodward, A, (2015). Learning from gesture: How early does it happen? *Cognition, 142,* 138-147. doi: 10.1016/j.cognition.2015.05.018

Novack, M. A., Wakefield, E. M., & Goldin-Meadow, S. (2016). What makes a movement a gesture? *Cognition*, *146*, 339-348. doi: 10.1016/j.cognition.2015.10.014

Novack, M.A., Wakefield, E. M., Congdon, E, C., Franconeri, S., & Goldin-Meadow, S. (2016) There is more to gesture than meets the eye: Visual attention to gesture's referents cannot account for its facilitative effects during math instruction. T*o appear in Proceedings of the 38th Annual Conference of the Cognitive Science Society*

Özçalışkan, Ş., & Goldin-Meadow, S. (2005). Do parents lead their children by the hand? *Journal of Child Language, 32.* 481-505. doi: 10.1017/S0305000905007002

Özçalışkan, Ş., & Goldin-Meadow, S. (2011). Is there an iconic gesture spurt at 26 months. *Integrating gestures: The interdisciplinary nature of gesture. Amsterdam, NL: John Benjamins*.

Ozyurek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of Cognitive Neuroscience, 19*, 605-616. doi:10.1162/jocn.2007.19.4.605

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

Ping R. M., & Goldin-Meadow, S. (2008). Hands in the air: using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology, 44*, 1277-1287. doi:10.1037/0012-1649.44.5.1277

Pulverman, R., Song, L., Hirsh-Pasek, K., Pruden, S. M. and Golinkoff, R. M. (2013) Preverbal Infants' Attention to Manner and Path: Foundations for Learning Relational Terms. *Child Development*, *84*, 241–252. doi: 10.1111/cdev.12030

Pulverman, R., Golinkoff, R. M., Hirsh-Pasek, K., & Buresh, J. S. (2008). Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition, 108,* 825–830.

Rotem-Kohavi, N., Hilderman, C. G., Liu, A., Makan, N., Wang, J. Z., & Virji-Babul, N. (2014). Network analysis of perception-action coupling in infants. *Front. Hum. Neurosci*, *8*. doi: 10.3389/fnhum.2014.00209

Roth, W., M. (2007). Making use of gestures, the leading edge in literacy development. In W. Saul (Ed.), *Communicating science: Examing the discourse*.

Roth, W., M., & Lawless, D. (2002). Science, culture, and the emergence of language. *Science Education*, *86,* 368-385. doi: 10.1002/sce.10008

Rowe, M.L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science, 323*, 951-953. doi: 10.1126/science.1167025

Schachner, A., & Carey, S. (2013). Reasoning about 'irrational' actions: When intentional movements cannot be explained, the movements themselves are seen as the goal. *Cognition, 129*, 309-327. doi:10.1016/j.cognition.2013.07.006

Searle, J., R. (1980). Minds, brains, and programs. *The Behavioral and Brain Sciences, 3*, 417-457. doi: 10.1017/S0140525X00005756

Sekine, K., Sowden, H., & Kita, S. (2015). The development of the ability to semantically integrate information in speech and iconic gesture in comprehension. *Cognitive Science*, *39*, 1855-1880. doi: 10.1111/cogs.12221

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science, 16*, 85-89. doi: 10.1111/j.0956-7976.2005.00786.x

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. In: Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008).

Stanfield, C., Williamson, R., & Özçalışkan, Ş. (2014). How early do children understand gesture-speech combinations with iconic gestures? *Journal of Child Language, 41*, 462-471. doi: 10.1017/S0305000913000019

Stieff, M. (2011). When is a molecule three dimensional? A task-specific role for imagistic reasoning in advanced chemistry. *Science Education, 95*, 310-336. doi: 10.1002/sce.20427

Thomas, L. E., & Lleras, A. (2009). Moving eyes and moving thought: on the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review, 14*, 663-668. doi: 10.3758/BF03196818

Tolar, T. D., Lederberg, A. R., Gokhale, S., & Tomasello, M. (2008). The development of the ability to recognize the meaning of iconic signs. *Journal of Deaf Studies and Deaf Education, 13*, 225-240. doi: 10.1093/deafed/enm045

Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes, 15*, 249-275. doi: 10.1080/01638539209544812

Trabasso, T., Stein, N. L., Rodkin, P. C., Munger, M. P., & Baughn, C. R. (1992). Knowledge of goals and plans in the on-line narration of events. *Cognitive Development, 7*, 133-170. doi: 10.1016/0885-2014(92)90009-G

Trofatter, C., Kontra, C., Beilock, S., Goldin-Meadow, S. (2015). Gesturing has a larger impact on problem-solving than action, even when action is accompanied by words. *Language, Cognition and Neuroscience*, 30, 251-260. doi:10.1080/23273798.2014.905692.

Valenzeno, L., Alibali, M. W., & Klatzky, R. L. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28*, 187-204. doi:10.1016/S0361-476X(02)00007-3

Vouloumanos, A., Onishi, K. H., & Pogue, A. (2012). Twelve-month-old infants recognize that speech can communicate unobservable intentions. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 12933–7. doi:10.1073/pnas.112105710

Wakefield, E. M., Congdon, E., Novack, M., Goldin-Meadow, S., & James, K. H. (2016). Learning math through your hands: The neural effects of gesture-based instruction in 8-year-old children. *Manuscript submitted for publication*

Wakefield, E. M., James, T. W., & James, K. H. (2013). Neural correlates of gesture processing across human development. *Cognitive neuropsychology*, *30*, 58-76. doi: 10.1080/02643294.2013.794777

Wakefield, E. M., & James, K. H. (2015). Effects of learning with gesture on children's understanding of a new language concept. *Developmental Psychology, 5*, 1105-1114. doi: 10.1037/a0039471

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*, 625–36. doi: 10.3758/BF03196322

Willems, R. M., Ozyurek, A., & Hagoort, P. (2007). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex, 17*, 2322-2333. doi:10.1093/cercor/bhl141

Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1-34. doi: 10.1016/S0010-0277(98)00058-4

Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors, *Infant Behavior and Development*, *22*, 145-160. doi: 10.1016/S0163-6383(99)00007-7

Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development*, *17*, 1061-1084. doi: 10.1016/S0885-2014(02)00074-6

Zacks, J., Tversky, B., & Iyer, G. (2001). Perceiving, remembering and communicating structure in events. *Journal of Experimental Psychology: General, 130*, 29-58. doi: 10.1037//0096-3445.130.1.29