# Using Gestures and Body Movements for Thinking and Learning

Jing Zhao

Submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy
under the Executive Committee of the Graduate School of
Arts and Sciences

COLUMBIA UNIVERSITY

2018

**Abstract**

Using gestures and body movements for thinking and learning

Jing Zhao

Gestures have been found to be helpful to people in many cognitive and daily activities, such as speaking, counting, learning, and problem solving. However, different gestures benefit people to different degrees, and people use gestures in different ways to assist thinking and learning.  From an embodied cognition perspective, gesture is seen as a simulated action. Therefore, to further understand the mechanisms of gesture's effects on thinking will directly help us harness embodied cognition theories to guide teaching and learning. In the literature, it is widely known that gesture not only reflects thinking, but also actively promotes thinking and learning. However, the mechanisms that account for gesture's effects on cognition remained obscure to us.

To better understand how different types of gestures benefit thinking and learning, Study 1 was conducted with 31 participants to investigate how teaching big (*n*=15) and small gestures (*n*=16) as a problem solving strategy influenced the actual gesture use and performance. The results suggested that the small gesture might possibly be a more effective gesture, because people who were taught small and used small gestures had the highest accuracy percentage on the primary task. However, using the small gesture did not significantly lower cognitive load compared to using the big gesture.

Based on these findings, Study 2 was conducted with 100 adults to further investigate how teaching different types of gestures influenced learners' gesture use, performance, learning, and cognitive load. In this study, the participants were randomly assigned to three groups, where they were taught to solve a molecular structure problem using small

(*n*=25), big (*n*=50), or no gestures (*n*=25). Then they were left in a quiet room to solve 15 molecule questions independently. Their answers and time spent on each question were recorded. A dual-task paradigm was used as an objective measure of cognitive load, and a NASA Questionnaire was used as a subjective measure of cognitive load. At the end, participants were asked to answer some transfer questions. Throughout the study, all participants' gestures and body movements were recorded by two cameras.

The findings from the two studies suggested that teaching different types of gestures had some influence on people's gesture use, performance, learning, and cognitive load. Specifically, small gestures taught as a problem-solving strategy were adopted more easily and more effectively used than big gestures and body movements. Questions that were answered through small gestures seemed to have a slightly higher accuracy percentage, but were not necessarily related to lowered cognitive load. The study also found that when people were taught gesture as a problem solving strategy and then asked to use it, they took some time at the very beginning to try and practice, and then gradually transitioned to using no gestures. In both studies, their thinking time, gesture time, gesturing density decreased gradually, without sacrificing accuracy. These findings contributed to both embodied cognition theories and gesture literature, and also shed light on instructional design in an educational setting.

## TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGEMENTS

First of all, I would like to thank my Ph.D advisor, Professor John Black, for your continuous guidance throughout the past five years. During all my years at Teachers College, you have been extremely open-minded and supportive to all of my ideas in research. You always encourage me to read, to try, to experiment, and learn on my own. I am very grateful for the freedom you gave me. With it, I explored various projects till I found what I am truly interested in. You are also a role model to me. You not only taught me how to do research, but also taught me to have a fearless and curious mind forever.  I will bring the spirit with me forever, wherever I go.

I would like to thank Professor Matthew Johnson, for teaching me in your courses: Multivariate Analysis and Statistical Inference, and advising me from 2014 on my Master of Science in Applied Statistics to my doctoral studies in Cognitive Science. Although these are the only two courses I took from you, your teaching made my learning experience such an enjoyable journey. I am also very grateful for your advice on my master's studies, without which I would not be able to finish it on time and love statistics to this degree. I appreciate your continuous invaluable insights and guidance.

I would like to thank Professor Catherine Chase, for allowing me to be a member of your research team in the Invention Tutor Project. It is the most exciting research project that I have worked on in the field of Artificial Intelligence. When I was in your team, you have been extremely patient and supportive to me. As a woman scientist, you are absolutely leading the way by example.

I would like to thank my dissertation committee members, Professor Steven Peverly and Professor Lisa Son. Your thoughtful comments and insightful critiques make this dissertation much better.

I would like to thank Professor Barbara Tversky. You are the first person I have ever known in person from Teachers College. In 2012, you led me into Teachers College, my dream school. I am very grateful that you gave me the opportunity to sit in your practicum, join the discussions, and learn from your wisdom. The years I spent in your practicum and your courses will remain to be one of the most cherished memories of graduate school.

I would like to thank Professor Frank Tang, from New York University, my advisor and my life mentor. You changed my life! Without you, I would not come to New York, go to graduate school, become a teacher, and start a research career. Especially, thank you so much for backing me up in all situations, when I changed jobs, applied for internships and graduate schools, switched careers, went to new cities, and everything else. You are like a father to me. Your warmth is contagious!

I would like to thank the Ben Wood Fellowship for the generous financial support. It has been a great honor for me.

I would like to thank my colleagues at the Cognition and Learning Lab and Institute for Learning Technologies, Michael Swart, Ben Friedman, Tiantian Jin, Ilya Lyashevesky, Alison Lee, Jenna Marks, Marianna Lamnina, Douglas Huang, Na Li, Mel Cesarano, Mellissa Bradley, Sor Kornkasem, Dr. Robin Walker, Jung-Hyun Ahn, Laura Malkiewich, and Woonhee Sung. The time we spent discussing research and exchange minds is one of the best part in graduate school.

I would like to thank my friends at Teachers College, Huacheng Li, Zhuangzhuang Han, Ahram Choi, Carol Lu, Seokmin Kang, Sat Virk, Yaoli Mao, Azi Jamilian, Julia Xing, Huei-yi Lai, Yang Liu, Danfei Hu, Chucheng Hu, Helena Connolly, and Shimin Kai, for the great time we spent together.

I would like to thank all mothers at Teachers College! Thank you for being so supportive and understanding. Also, thank you TC for the great lactation room. This is super helpful to me!

I would like to thank my colleagues Kennix Lee, Audrey Tse, Jenna Marks, Alison Lee, and Ilya Lyashevesky for helping me finalize the last version of this thesis.

I would like to thank my father Zhao, Quan Qiu and my mother Ai, Fang. Thank you very much for always challenging me from being good to the best, for the unconditional love, and for all the support you could provide along the way. For you, I am forever in debt. ☺

I would like to thank my father-in-law Zhong, Jiao, and mother-in-law, Li, Gui Lan. Thank you very much for explaining the definition of a loving family, for all of your emotional and financial support, and for taking care of my baby son so that I can work on my dissertation. I appreciate every meal you cooked, every gift from you, and everything you did.

I would like to thank my husband Zhong, Xiao, for being a supportive husband and a wonderful daddy, and for building this family together with me. Without you, I would not make the decision to apply for graduate school. Thank you for holding my hand tightly, especially during the final years of our Ph.D careers, when both of us were jobless and poor. Of course, we have worked together to conquer these difficulties.

I would like to thank my son, Aaron Junshen Zhong. I became a new mother in graduate school, and I am still adjusting to this new mommy life. After you were born, I learned that parenthood starts with a series of sleepless nights for several months. I had to run to the Mother's Room every three hours to pump breast milk to feed you, even when I was collecting data for this thesis. You made me understand what it means to be a parent, and your chunky cheeks have brought so many moments of laughter to everyone in this family. We know you will have a wonderful life ahead.

Finally, I would like to thank all who have been supportive to me. The world is so small. I am sure we will meet again and work together in the future!

*To my parents*

# Chapter I Introduction

## A. How gesture as a simulated action promotes thinking

Literature in embodied cognition has shown that human cognition is grounded and embodied, and action plays central roles in cognition and learning. Therefore, actions that are compatible with perception can promote thinking. At the same time, gesture, as a form of simulated action, can reflect and promote thinking and learning in a similar way. This line of theories suggests that embodiment, by enacting experience through our body, is supposed to promote thinking and learning. This point of view has been supported by a growing body of research demonstrating that incorporating bodily movements and gestures into instruction can promote thinking and learning (e.g., Alibali & Nathan, 2012; Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004; Goldin-Meadow, Cook, & Mitchell, 2009). Furthermore, studies with new technologies and user interfaces that recognize natural bodily movements (i.e., gestures, bodily movements) seemed to even suggest that the richer the perceptual experience is, the better the learning result will be. For example, Johnson-Glenberg, Birchfield, Tolentino, and Koziupa (2014) showed that allowing students to manipulate a virtual molecule using a trackable wand in an immersive and highly interactive platform consistently led to greater learning gains, compared to learning in a regular classroom setting. Johnson-Glenberg, Birchfield, and Usyal (2009) reported that a multimodal and immersive learning environment that incorporated multiple modalities (motion, sound, and graphics) could also lead to better learning and retention in geology as well. In a more recent paper, Johnson-Glenberg et al. (2014) proposed a taxonomy of embodiment, claiming that a higher level of embodiment is likely to bring a more efficacious learning result. Therefore, we could predict that a higher level of embodiment (e.g., a larger scale of body movements) will be associated with better learning than a lower level of embodiment (e.g., small-scale gestures).

Gesture can be seen as a kind of action, thus has the potential to promote learning. In the gesture literature, it has been widely agreed that gesture can promote thinking and learning in many cognitive activities. However, gesture's effect on cognitive load seemed to be a bit more controversial. One point of view is that while gesture serves to lighten working memory load like diagrams, this role is not the primary reason that gesture promotes learning. In a recent study by Jamalian, Giardino, and Tversky (2013b), they found participants' proportion of time spent on gesturing during studying materials did not increase, as memory load increased from light to heavy. Based on it, they conjured that gesture's role in offloading memory appeared to be less important than its other features like creating embodied representations. Another point of view even argued that gesture may impose cognitive load. In a study by Mol, Krahmer, Maes, and Swerts (2009), they proposed that many gestures produced by speakers were for communicative purposes, so speakers needed to put cognitive effort into producing gestures. They reported that in a communicative setting, when a narration task was not very demanding, a speaker produced a lot of gestures. However, when the narration task became more demanding and cost more load, the amount of gestures decreased. Beyond that, Warburton, Wilson, Lynch, and Cuykendall (2013)'s study also challenged this gesture-lightens-load explanation by proposing that the size of gesture influenced cognitive load. By instructing experienced dancers to learn and practice a dance routine either by reduced body movement ("marking") or by full-scale body movement (dancing "full out"), they found that dancers who practiced with the "marking" method performed much better than those who practiced in a "full out" way. Therefore, they hypothesized that large-scale body movement itself might impose its own cognitive difficulties and resulted in reducing the cognitive benefits of movement. If we think along this line of theories, we would predict that big-scale bodily movement is associated with less effective learning, compared to small-scale hand gestures.

## B. Overview of the dissertation

The dissertation is organized into four chapters. Chapter II is a literature review on grounded cognition theories, its application in the field of education, and how gesture as a simulated action promotes thinking and learning. It started with a review on the history of embodied cognition theories and actions' central role in human cognition. Then, I reviewed two theoretical frameworks that support the application of embodied cognition theories in classroom instruction. Finally, I dove deeper into the gesture literature and discussed how and when gesture could promote thinking and learning as a simulated action and a form of embodiment.

Chapter III presents Study 1. In Study 1, I investigated how teaching different types of gestures influenced people's spontaneous gesture, how people's use of gestures changed over time throughout the experiment, and how using big and small gestures influenced people's performance and cognitive load.

Chapter IV presents Study 2. Based on findings from Study 1, I investigated how teaching small, big, and no gestures influenced people own gesture use, how using small, big, mixed, and no gestures influenced people's performance and cognitive load, how people's gesturing behavior changed over time, and how they influenced people's performance in transfer questions.

In the final chapter, I included discussions on theoretical contributions, practical implications, and limitations of the studies. The dissertation concluded with possible directions for future research.

# Chapter II Literature Review

The literature review is organized into three sections. The first section reviews the background of the grounded cognition theories, including how grounded cognition is different from traditional views of cognition, the diverse foci of embodied cognition theory, and empirical evidence demonstrating how perception, action, and mental simulation work together in forming human cognition. The second section goes more specific to review literature on how the embodied cognition research and theory are applied in a specific field, education. It begins by reviewing two frameworks that support this application, and provides empirical evidence on how physical action and simulated action promote thinking and learning. The third section hones in on how gesture, which is a simulated action and a form of embodiment, connects with a wide range of cognition activities. It also discussed how gesture promotes thinking and learning.

## A. Perception, action, and embodied cognition

The grounded cognition theory proposes that human perception is grounded in action and perception. Within the grounded cognition theory, researchers have focused on different aspects of it. A growing body of studies has provided empirical evidence supporting this view.

### 1. Cognition is grounded and embodied.

Traditional views of cognition assumed that all knowledge is processed in our mind as arbitrary, abstract, and amodal symbols (Glenberg et al., 2004). However, this view was challenged in the past forty years. Increasingly more evidence has provided support to the perspective that cognition is grounded and embodied. They provided support to the view that human cognition is based on the interaction among bodily state, situated action, and the environment. Therefore, to develop a full understanding of something, we needed to construct a modal simulation of it in our mind (Black, Segal, Vitale, & Fadjo, 2012).

### *a) Bodily state, situated action, and the environment interact with a modal simulation to build human cognition*

Where does knowledge reside? Humanity's efforts to understand the world itself and its knowledge can be traced back to the time when humans took the first step in an attempt to understand learning and thinking. Till today, the question of how knowledge is acquired and stored in the brain remains to be one that is still open to debates and discoveries.

Contemporary research in psychology and human cognition has provided increasingly more support to the new perspective of grounded cognition, both theoretically and empirically.

In the past, traditional views of cognition assumed that all knowledge is represented as amodal symbols, which are processed in a semantic memory system in our brain. The propositional symbols serve as the basis for the spectrum of our cognitive processes from perception to thoughts.  Since 1980s, this point of view was challenged by grounded cognition researchers from various fields, including linguistics, computer science, philosophy, artificial intelligence, and cognitive science (Barsalou, 1999, 2008; Glenberg et al., 2004). They questioned the traditional view of cognition for lacking support in empirical evidence, and for its insufficient explanation on how cognition interfaces with perception and action (Barsalou, 2008). One example is that Searle (1980) proposed the Chinese Room Problem, and argued that simple processing of pure symbols would not lead to knowledge and learning.

Rather than explaining cognition in an amodal system, grounded cognition researchers proposed that modal representation is actually central to human cognition (e.g.Barsalou, 1999; Decety &

Grèzes, 2006; Goldman, 2006). They believe that when experience occurs (e.g., seeing a horse), our brain captures the information of the horse across modalities (e.g., how it looks, smells, and neighs, how it gallops, the environment it is in, and how we feel when we ride it, etc.,), integrates the information, and stores it as multimodal representations in memory. Later, when the information needs to be retrieved, a simulation of the horse will be activated in our brain, and we will have access of all the perception, action, and introspection information we gathered before. Grounded cognition researchers, in this way, united cognition with human perception, action, and introspection. Cognitive linguistics theories took this point even further and argued even the abstract concepts could be grounded metaphorically in our experience as well. They brought up extensive evidence across daily language and literature to show that people widely use concrete metaphors to talk about abstract ideas (e.g., *Love is a journey*; using *up* for good, positive experiences, and *down* for bad, negative ones) (Gibbs, 1994; Lakoff & Johnson, 1980, 1999; Turner, 1996).

In general, contrary from traditional views on cognition, proponents of grounded cognition hold that processing of pure symbols to build cognition is quite questionable. Bodily state, situated action, modal simulation, and even the environment interact to build human cognition.

### *b) Diverse foci in grounded cognition theories*

Although proponents of grounded cognition all reject to see human brain as a processor of abstract symbols, and believe that human's minds must be understood in the context of its relationship to a physical body and/or an interactive world, there remains a wide range of diversity in the term of *ground cognition*, especially on what aspects of it is the most central (Gibbs, 2006; Wilson, 2002). In a comprehensive review, Barsalou (2008) in particular, used the term of *grounded cognition* to emphasize that cognition actually could be grounded in multiple sources, including bodily state, situated action, social interaction, and the environment. At the same time, he also pointed out that the grounded cognition literature has different foci on each of these sources.

Situated action is considered one of these foci. As early as 1979, Gibson (1979a) explained from an ecological approach that visual perception is a situated activity. In this line, literature claimed that all actions are situated and cognition emerges from the interaction of brain, body, and environments (Chiel & Beer, 1997; Clark, 1997; Pfeifer & Scheier, 2001; Steels & Brooks, 1995; Thelen & Smith, 1996). Other researchers laid emphasis on social interaction theories. They proposed that when we try to perceive and understand other people's mind, we represent what's going on in others' mind by using a simulation in our own (Cattaneo & Rizzolatti, 2009; Goldman, 2006). Still many other aspects of the grounded cognition theory focused on the role of simulation particularly. They believe that when we interact with the world, a simulation is built upon our perceptual, motor, and introspection state. It is the presence of this simulation mechanism, as a means to represent information that allows us to develop an understanding of the world and draw inferences (Barsalou, 1999; Decety & Grèzes, 2006; Goldman, 2006; Kosslyn, 1980). Apparently, even more accounts of grounded cognition began to focus on the human body, as widespread evidence has shown mutual connections between bodily state and cognition (Barsalou, Niedenthal, Barbey, & Ruppert, 2003; Lakoff & Johnson, 1980; Smith, 2005).

In particular, Barsalou (2008) used the term *embodied cognition* to emphasize the dynamic interactions between human body and the physical world. Gibbs (2006) and Shapiro (2007) also used this term widely to emphasize the importance of the interaction between physical body and the world. Therefore, for the rest of the paper, I will use *embodied cognition* when discussing the role human body plays in cognition and learning.

### *c) Modal simulation in perception and action*

Embodied cognition has received wide empirical support, which suggests that mental simulations that are grounded in perception and action are essential in human cognition process. Below are some examples demonstrating the relation between action and a variety of perception models (i.e., visual, auditory, olfactory, gustatory and somatosensory). They have shown that action and perception are connected with modal simulations in our mind.

Among the various perception models, the visual perception model is the most widely studied one. It can be traced back to Kosslyn's mental imagery theories (Kosslyn, 1980, 1994). Embodied cognition researchers found that visual perception of an object can activate people's actions on it. When people see an object, a simulation of potential situated action on the objects will be activated so that people can be prepared to act on it. For instance, Tucker and Ellis (1998) reported a study showing that simply seeing a graspable object with a handle (e.g., a saucepan) can activate corresponding hand actions on it. Human are not sensitive to the objects, but could be sensitive even to the orientation of objects. Symes, Ellis, and Tucker (2007) reported that the angle of an object can facilitate spatially compatible responses. Moreover, our motor-based knowledge of how an object can be utilized also can come into play when we see things. Gerlach, Law, and Paulson (2002) reported that easily manipulated objects such as fruits, vegetables, and articles of clothing activated brain's motor areas more strongly than animals and nonmanipulable man-made objects. Interestingly, these simulations of objects can still occur even when the object itself is invisible to human, like when their names are shown as text (Tucker & Ellis, 2004; Zwaan, Van Der Stoep, Guadalupe, & Bouwmeester, 2012).

Human perception can influence action, but actions can also influence perception. In a series of studies, Proffitt (2006) reported that perceived steepness of a hill and perceived geographical distance could be influenced by actions and the state of the human body (e.g., wearing a heavy backpack or feeling tired). When people were tired or in a poor health condition, they felt the hills were steeper and the distance was longer. When people felt better, the hills became not that steep and the distance shorter. Franklin and Tversky (1990) also proposed that our perceived environment is shaped by the structure of our body. They found that when people were standing upright, locating objects along the head/feet axis was the easiest, followed by front/back and then by left/right. When people were in a reclining posture, the front/back was the easiest and were followed by head/feet and then left/right. This is because human body is symmetric along the left/right axis but perceived asymmetry with respect to the ground. Therefore, locating objects along the left/right axis is the most difficult when the environmental and bodily cues are relatively lacking. As Dewey (1896) said, "The motor response determines the stimulus, just as truly as sensory stimulus determines movement"(p.4).

The connection between perception and action can be found in human auditory system as well. Halpern, Zatorre, Bouffard, and Johnson (2004) reported that when making judgments about heard timbres, fMRI showed activation in people's primary and secondary auditory areas. Haueisen and Knösche (2001) found that listening to piano pieces can involuntarily trigger the respective finger movements of pianists. These movements, however, were not found for non-pianists who lacked the association between the auditory perception and the actions producing the music. In another study, Repp and Knoblich (2004) found that pianists, in order to recognize their own playing from auditory recordings, created a simulation of their own motor actions. This simulation helped them to match the anticipated and perceived action effects. In the domain of language, Pulvermüller et al. (2006) reported that when processing language, hearing syllables spoken not only activated the superior temporal cortex in the brain, but also the lips and tongue movements that produced them. Neurobiological models have provided additional support by claiming the connections between speech perception and its production mechanisms as well (Fry, 1966; Fuster, 2003; Pulvermüller, 1999).

The effects between perception and action are not limited to the visual and auditory domains. Research has found that perception of stimuli can activate many other brain areas too. For example, reading odor-related words (e.g., garlic, cinnamon, jasmine) can elicit activation in the primary olfactory cortex (González et al., 2006). Viewing pictures of appetizing foods can activate gustatory processing areas (Simmons, Martin, & Barsalou, 2005). Deficits in one's motor and somatosensory systems can make it difficult for people to judge weights when observing others lifting them (Bosbach, Cole, Prinz, & Knoblich, 2005).

The above empirical evidence all support the theory that property information is distributed in our brain's modality-specific areas. People process different information in different modals and represent the information by their perceptual simulations. Therefore, in verification tasks (e.g., Is *face* a property of *gorilla*?) when simulated perception information was accessed and activated, variables like property size were found to affect people's verification time and error (Solomon & Barsalou, 2004). Across modalities, when people switch simulation from one modality to another, a switch cost could incur (Marques, 2006; Pecher, Zeelenberg, & Barsalou, 2003, 2004; Spence, Nicholls, & Driver, 2001). For instance, Pecher et al. (2003) reported that when participants were asked to verify if an object (e.g., a lawn mower) had a certain property (e.g., loud), if the perceptual dimension switched between tasks, it took them more time to respond.

## 2. Action plays central roles in cognition and learning

Compared with the term *grounded cognition*, *embodied cognition* lays more emphasis on the dynamic interactions between the human body and the physical world. This view also implies that actions play an important role in human perception and cognition (Barsalou, 2008). The following reviews empirical evidence on how action influences cognition and learning.

### *a) Action and perception have mutual effect on each other*

Actions can have direct influence on how we perceive the world even since the time when we were infants. Campos, Bertenthal, and Kermoian (1992) reported that actions can change perceptions of infants who were younger than one-year old. In their study, they found that seven-month-old babies who already experienced crawling refused to cross over a visual cliff and showed signs of fear (i.e., accelerated heart rate) when placed at the edge of it. However, babies of the same age but having not been crawling yet did not show these signs of fear. They went further to make pre-crawling babies stand in a "walker" so that they can push with their feet and receive some self-generated motion. They found that these motor experiences allowed by the walker changed babies perception of the visual cliff, and made them fear. However, babies of the same age but having not received the experience from the walker still did not show fear at the visual cliff (Bertenthal, Campos, & Kermoian, 1994). This evidence showed that it is the developed motion, instead of age, that changed infants' perception of the world.

Action can also influence people's text comprehension. In an early study by Klatzky, Pellegrino, McCloskey, and Doherty (1989), priming a certain hand shape (e.g., pinch, poke, clench, and palm) was found to facilitate people's judgments about whether a phrase is sensible (e.g., "crumple a newspaper") or not (e.g., "crumple a window"). This showed that the influence of action can be found even for abstract and metaphorical concepts as well (Wilson & Gibbs, 2007).

### *b) The action-compatibility effect*

Studies in the context of embodied cognition not only reported the mutual effect between action and perception, but also a more robust congruent effect between the two, called action-compatibility effect (ACE). Many studies have shown that congruence between action and perception could make information processing easier, compared to incongruence. For example, Wexler, Kosslyn, and Berthoz (1998) reported that in mental rotation tasks, people were faster and more accurate when they concurrently perform a manual rotation in the same direction as the required mental rotation. Furthermore, a change in the speed of motor rotation can affect the speed of mental rotation correspondingly. Similarly, Wohlschläger and Wohlschläger (1998) reported that the spontaneous use of rotational hand movements resulted in shorter response times; and that compatible hand rotational directions facilitated mental rotation, whereas incompatible directions inhibited it.

This *action-compatibility effect* can be found in language comprehension as well. Glenberg and Kaschak (2002) asked subjects to judge whether sentences were sensible by making a response that required moving toward or away from their bodies. They reported that subjects found it easier to make a sensible judgment when the action implied in the sentence was in the same direction (e.g., to making an action moving away from the body when the sentence is "Close the drawer.") than in an opposite direction. They found this effect still presented in sentences describing the transfer of abstract entities as well, like "Liz told you the story." Wilson and Gibbs (2007) reported that both producing and imagining an appropriate body movement prior to reading a sentence containing compatible metaphorical phrases (e.g., "Grasp a concept") can facilitate comprehension of it. This effect can be found between one's action and emotive state too. Barsalou et al. (2003) reported that participants made faster responses when asked to indicate liking by pulling a lever towards their body, than those who were asked to indicate liking by pushing the lever away. Therefore, Solomon and Barsalou (2001) suggest that concepts are grounded in sensory-motor simulations.

The section above reviewed the literature in the perspective of grounded cognition, and discussed the connection between human perception, action, and cognition. In the next section, I will move on to talk about how embodied cognition theories are applied to the field of education to augment thinking and learning.

## B. Embodied Cognition as an Instructional Approach

How can we apply the embodied cognition theories into pedagogy? How embodied cognition can be applied to augment thinking and learning? What's gesture's role in classroom teaching and instructional design? When does gesturing promote learning and when does it not? As the theory of embodied cognition evolves, these and many other questions have begun to draw researchers' attention when considering their applications.

Educators and researchers have recognized the value of actions and movements in teaching and learning a long time ago. Dewey (1938) raises that each child is active, inquisitive, and wants to explore. Thus, instructors should integrate learning with experiences that are meaningful and useful to them and allow students to "learn by doing". Montessori (Montessori & Carter, 1936) also highlight that physical activity is an essential factor in children's intellectual growth, and state that "through movement we come in contact with external reality, and it is through these contacts that we eventually acquire even abstract ideas."

### 1. Theoretical framework that supports the application of embodied cognition theories into classroom instruction

Researchers of embodied cognition have proposed different theoretical frameworks explaining how embodied cognition perspectives can provide insights into cognition and learning. In this section, I will first discuss two of them in detail, because these two have been supported by a rich body of empirical evidence.

#### *a) The Indexical Hypothesis*

The first framework is the Indexical Hypothesis (IH) developed by Glenberg and Robertson (1999, 2000). The IH postulates that there are three steps when we convert text to meaning, which are *indexing, affordance*, and *mesh* (Glenberg, 2008, p. 45). When we try to understand a sentence such as "Art flicked the snake off the porch using the chair," we will go through the three steps one by one. First, the *indexing* process happens so that we can map the words (e.g., "art") and phrases (e.g., "the chair") to either objects in the environment or perceptual symbols that we are familiar with. Then, in our mind, *affordances* (see Gibson, 1979a) are derived from the indexed objects or perceptual symbols, so that we understand what can be done with the objects (e.g., the chair can be sit on, or be lifted as a weapon to flick away a snake), depending on our bodily state. For example, the affordance of a chair is different for an adult and a child. An adult can use the chair in a lot of different ways, such as sitting, standing on, or even lifting to use as a weapon. However, for a child, a chair may be used to sit on or to hide behind, but it would not be lifted up to use as a weapon. Third, grammatical knowledge will be used to *mesh* the affordances into a coherent set of actions, under the guidance of syntactic constructions. In the example, we will then understand that it is the subject, “chair”, that causes the motion of the object, “snake”, to go to a location, “off the porch”.

Glenberg and colleagues tested this Indexical Hypothesis in a series of studies. In one study, Glenberg et al. (2004) asked elementary school children to play with toy objects that were referred to in text (e.g., a barn, a tractor, and a horse, in a text about a farm) to simulate the actions described in the text. They found that children who manipulated text referents after reading a sentence performed better in memory tasks. In a following study, they asked the children to only imagine manipulating the toys. They found that the imagined manipulation strategy were as effective as physical manipulation, and it was even maintained when tested on new texts several days later. In another study, Glenberg, Brown, and Levin (2007) applied this strategy in a small group setting, where children took turns reading the sentence aloud and were then asked to manipulate the objects. They found physical manipulation had a positive effect on students’ reading performance when executed in small groups, and that watching others manipulate objects was as effective as manipulating on their own.  Similar findings were reported with Native American students who had learning difficulties as well (Marley, Levin, & Glenberg, 2007). This series of findings was further generalized to virtual images. Glenberg, Goldberg, and Zhu (2011) reported that manipulating images of text referents on a computer screen benefited children as much as manipulating real toys, when it was used as an instructional strategy to promote reading skills. This facilitative effect could be found even one week later.

### *b) The Imaginary Worlds*

Another framework is Black (2007)’s *Imaginary World* framework. Black’s *Imaginary World* was inspired by the concept of *Story World* (Black & Bower, 1980), which proposes that a *Story World* is a level of memory representation that the story text refers to. It includes the symbols, relationship, and the propositional content of the story (Carnap, 1956). Black (2007) stated that the same idea was echoed by a series of proposals, including the mental model proposals by Johnson-Laird (1983) and Gentner and Stevens (1983), and the situation model proposal by Van Dijk and Kintsch (1983). The *Story World* involves the visual and spatial imagery of the story, and allows people to imagine the different possibilities the story could play out. Black (2007) shifted the terminology *Story World* to *Imaginary World* to extend its application to content beyond stories. His *Imaginary World* explains a more general cognitive mechanism. He proposes that learners can construct a mental representation of learning materials, and this representation functions like a simulation that is essential for a full understanding.

A series of studies by Black and colleagues has demonstrated the application of the *Imaginary World* in education. Under the guidance of this framework, Hachey, Tsuei, and Black (2001) and Tsuei and Black (2004) reported teaching functional relations and formal system diagramming to middle school students through a Mars Colonies project. They found that although formal system diagraming had been proven to be too difficult for precollege students, their approach worked well for these middle school students. The students who were taught to think in *Imaginary World* learned to diagram dynamic earth science phenomena and were able to transfer these concepts and skills in later designs. Kuhn, Black, Keselman, and Kaplan (2000) reported another study in which scientific reasoning skills were taught to middle school students. In their *Imaginary World*, which was assisted by a computer simulation, students were able to conduct “thought experiments” by isolating variables one at a time. After several weeks, both their scientific inquiry skills and understanding of the content materials were improved. Another example is the Reflective Agent Learning Environment developed by Bai, Black, and Vitale (2006) . To play in this environment, students needed to teach a computer agent by building concept maps and system diagrams. Based on students' input, the agent then provided feedback on its thoughts on how the virtual world worked and how the virtual world itself was supposed to work. Students learned from teaching the agent, and reflected on the contrast between their *imaginary worlds* and the agent's worlds.

In the same line, Fadjo (2012) proposed the Instructional Embodiment Conceptual Framework. He suggested that under the instructional embodiment, there are two primary levels: physical embodiment and imagined embodiment. Under physical embodiment, there are four forms: direct, surrogate, augmented, and gestural embodiment; under the imaged embodiment, there are two forms: explicit and implicit embodiment. Students first construct an *Imaginary World* space through physical or imaged activities on the learning content, and then through the sixed proposed forms of embodiment, an embodied experience can be incorporated into a traditional, formal instructional environment. He also provided evidence showing that physical embodied and imaged embodied activities can help students construct *Imaginary Worlds* that can promote the development of computational thinking skills and abstract concepts.

Applications of the embodied cognition theories in instructional design were further discussed in Black et al. (2012), where the authors proposed that "there are three steps involved in a grounded cognition approach to learning something: (1) Have an embodied experience, (2) Learn to imagine that embodied experience, and (3) Imagine the experience when learning from symbolic materials." Black et al. (2012) also listed several examples using the embodied learning environment to promote learning. One example involves teaching elementary school students number sense and addition rules. It is hypothesized that addition is a discrete mental activity, and is thus congruent with discrete, pointing gestures, and number estimation is a continuous mental activity, and is thus congruent with continuous gestures. Therefore a gestural interface that elicits gestures congruent with these respective learning tasks will likely benefit learning. Studies with preschool children showed that a well-designed haptic activity on a gestural-controlled interface can elicit "effective" gestures to promote learning. The gestures were "effective" in that they were congruent with their respective learning task (i.e., discrete or continuous). Therefore, they facilitated better use of learning strategies and resulted in better performance. Another example involves teaching the conception of shapes. In this study, elementary school children played in a video game environment, where they had to construct polygons with different features to help the agent navigate an obstacle course. The study found that children who represented crucial properties that determined polygon class (e.g., congruency, parallelism, and right angle) with gestures performed significantly better than who did not. They were also better at overlooking irrelevant properties in favor of class-defining properties. This "hand metaphor" was considered to be the key to provide a spatial-grounding experience to promote learning. A third study showed that this embodied experience not only benefited the learning of concrete concept, but it can also promote the learning of abstract concept, such as in computational skills and mathematical thinking. Fadjo (2012) reported that students who experienced direct physical embodied activities and were instructed to explicitly practice imagining them in their mind showed the best learning results.

Black et al. (2012) pointed out that unlike other pedagogical frameworks where it is the instructor's responsibility to model and embody, instructional embodiment found ways to "engage the student in a sequence or system of movement, imagination, and exploration."

### 2. Physical and simulated actions influence learning

Since action plays a central role in human cognition, specific actions such as physical manipulation, imaged manipulation, and activities through haptic channel on physical and simulated objects, are supposed to influence thinking and learning as well. Researchers have attempted different ways to involve action and movements to promote teaching and learning.

Here are some examples on the application of embodied cognition theories to thinking and learning. Early embodied cognition researchers Lakoff and Núñez (2000) proposed that children's mathematical concepts should be developed through metaphors rooted in perception and action. Based on that, Siegler and Ramani (2008) reported that having children play board games with consecutively numbered, linearly arranged and equally sized squares can improve their knowledge of numerical magnitudes. Martin and Schwartz (2005) also showed that physically manipulating pie wedges and tiles can facilitate children's ability to develop a correct interpretation of fractions, and this physical experience yielded better performance in further transfer tasks. Bara, Gentaz, Colé, and Sprenger-Charolles (2004) demonstrated that incorporating a visuo-haptic and haptic exploration of letters (i.e., having children explore letters with their fingers and run their index finger along its outline in a fixed order corresponding to its writing) helped children develop phonemic awareness, knowledge of letters, letter-sound correspondence, and alphabetic principle usage. Glenberg et al. (2004) found that having children manipulate toy objects referred to in a text (e.g., a barn, a tractor, a horse, in a text about a farm) helped children map worlds and phrases into real world objects and experiences, and therefore facilitated derivation of meaning. This manipulation resulted in better memory and comprehension. The benefit of physical movement is not limited to learning about concrete materials, but they can benefit learning abstract content as well. Fadjo (2012) reported that students who acted out coding scripts with their bodies were able to write more lines of code and more complex scripts, and were better at implementing computational thinking skills, comparing to students taught with traditional methods.

Moreover, the benefit of action can still be seen even when the manipulation did not occur with real objects. It can be found with images on a computer screen or imagined images. Glenberg and colleagues completed a series of studies showing that compared with manipulation of real objects, both manipulation in imagination and manipulation on a computer screen yielded as effective results (Glenberg et al., 2011; Glenberg et al., 2004). Han and Black (2011) reported that usage of a haptic-augmented simulation about gear rotation can improve students' learning of physics. Their results indicated that when solving physics problems about how input force on one gear can cause output force in an adjacent gear, students who learned by playing with a joystick that enabled actual feeling of arm movements and were shown an in-time simulation on the screen demonstrating the force the gears received with gear rotation speed, learned much better than those who learned with equivalent but non-haptic simulations. Jang (2010) examined the facilitative effects of actions in a virtual-reality program that aimed to teach the complex internal anatomical structure. It is found that students who manipulated the images of anatomical structures in a 3-D Virtual Reality environment outperformed their peers who only viewed the manipulation.

### C. Gestures as a form of embodiment promotes learning

Gesture is movement in the air, but is also a concrete and important form of embodiment. This section reviewed evidence demonstrating that gesture, as a specific form of embodied simulation, promotes cognition and learning.

Many theories have proposed that perception elicits and primes actions (Gibson, 1979b; Sperry, 1952). Based on these theories and empirical evidence, Hostetter and Alibali (2008) posit that gesture can be seen as manifestation of the simulation resulted from human perception and action. They also propose a Gesture as Simulated Action (GSA) Framework, which asserts that "gestures emerge from the perceptual and motor simulations that underlie embodied language and mental imagery". They explain that in this framework, in line with embodied cognition theories, the interaction between perception and action is central to human cognition. As agreed by many researchers, the interaction is mutual: perception can determine potential action, just as action can determine what can be perceived. By simulating perception and action, language and mental imagery are processed. When the strength of activation of the simulation becomes sufficiently strong to spread to one's motor areas and surpass one's inhibition to express, gesture will emerge and be realized as an overt movement.

McNeill (1992) classifies gesture movements into four major categories: iconic, metaphoric, beat, and deictic gestures. *Iconic gestures* bear a close formal relationship to the semantic content of speech it accompanies. The gesture and its accompanying speech often refer to the same event and are partially overlapping. Usually looking at the speech or the iconic gestures alone would not reveal a complete picture of the speaker's memory and mental representation of the scene. An example of this type of gesture is: when describing people bending a tree back to the ground, a speaker may use a gesture that appears to be gripping something and pulling it backwards. *Metaphoric gestures*, similarly, are pictorial as well, but present an abstract idea instead of a concrete object or event. For example, a speaker may use a gesture to represent "hollow words" or "a deep book" as if a word is a container and a book has vertical dimension (Reddy, 1979). *Beat gestures* are hand movements that are like beating musical time. Regardless of the content it accompanies, beating gestures always stay in the same form (McNeill & Levy, 1982), and move along with the rhythmical pulsation of speech. *Cohesive gestures*, which can come in the form of iconic, metaphoric, or pointing gestures, are the ones serving to "tie together thematically related but temporally separated parts of the discourse." For example, a speaker may make a hand movement to inform the listener to go back to the main story line. Lastly, the *deictic gestures* are gestures in pointing form. Pointing gestures can aim at a physical place or an abstract concept. For example, a speaker may point at a space when asking "where did you come from before?" Each type of gesture has effects on cognition and learning in their own ways.

### 1. Using gesture in cognitive activities

As an important form of embodiment, gesture is ubiquitous in daily activities. People across all cultures (Feyereisen & De Lannoy, 1991) and ages (Iverson & Goldin-Meadow, 1998a) use gesture -- even those who have been blind from birth and have never seen other people gesturing use some gestures (Iverson & Goldin-Meadow, 1998b). Individuals gesture when they are by themselves, in the darkness, or talking on the telephone (Bavelas, Gerwing, Sutton, & Prevost, 2008). Of course, gesture is used by people in a lot of cognitive activities.

A very commonly seen gesture is the "co-speech gesture". An early work by Feyereisen and De Lannoy (1991) stated that "the use of (co-speech) gesture was thought to represent a former, 'natural' state of language." Kendon (1986) and McNeill (1985) proposed that gesture and speech both relate to mental representations in thinking. More recently, Hostetter and Alibali (2008) proposed a *Gesture as Simulated Action* Framework. They pointed out that "gestures are not simply an epi-phennomenon of active mental images", and that gesture actually facilitates speech. They mentioned that there are three theories explaining this facilitative effect. One theory is the *image maintenance theory* (De Ruiter, 1998; Wesp, Hesse, Keutmann, & Wheaton, 2001), stating that gesture activates visuo-spatial information in working memory to prevent them from decaying too quickly. Wesp et al. (2001) provided support to this view by showing that speakers gestured more when describing paintings from memory than from their physical presence. Another theory is the *lexical access hypothesis* (Krauss, Chen, & Chawla, 1996; Krauss, Chen, & Gotfexnum, 2000). This hypothesis posits that gesture facilitates the retrieval of lexical items for spatial and motor ideas. Morsella and Krauss (2004) provided support to this theory by showing that gesture can directly affect spatial memory and lexical retrieval. In their studies, they found people gestured more when describing visual objects from memory, and when the objects were difficult to remember and encode verbally. Notably, people also gestured when describing a visually accessible object, but restriction on gesture produced dysfluent speech even when spatial memory was untaxed. A third theory is the *information packaging hypothesis* proposed by Kita (2000). This hypothesis suggests that gesture is involved in the conceptual planning stage. It helps speakers organize ("package") spatial information into units that are suitable to verbalize. To support this hypothesis, Melinger and Kita (2007) showed that speakers produced more gestures at moments of relatively high conceptual load (e.g., when they were asked to describe a picture with higher complexity). Hostetter, Alibali, and Kita (2007) also showed that when people talked about visual patterns that were more difficult to conceptualize, they also gestured more.  Morrel-Samuels and Krauss (1992), similarly, provided evidence suggesting that when describing less familiar information, speakers may need to explore alternative ways to package information. Therefore, the onset of speech may be delayed and occur later than the onset of gesture. Still another theory is the *growth point theory* by McNeill (1992, 2005). Rather than focusing on the facilitative effect of gesture on speech, this theory proposed that gesture and speech sometimes work in collaboration to express and form ideas. Consistent with it, Iverson and Goldin-Meadow (1998c) agreed that gesture that occurs with speech can benefit both the speaker and the listener by facilitating speaking, communication, and comprehension.

Another cognitive activity gesture is also quite widely associated with is counting (Graham, 1999). It is commonly seen that children spontaneously use gesture while counting objects (Fuson, 1988; Gelman, 1980). When children counted, regardless of whether the counting gestures were done by themselves or by a puppet, children counted more accurately than when they were not gesturing (Alibali & DiRusso, 1999). Regarding the function of gesture itself, Gelman and Gallistel (1978) proposed that children gestured to keep track of what had been counted. Fuson (1988) argued that children gestured to help themselves coordinate numbers with corresponding objects. Alibali and DiRusso (1999) proposed that active gestures produced by children themselves helped them to both keep track of the counted items and coordinate speech and the items.

Gesture has been shown to play an important role in problem solving as well. Chu and Kita (2011) reported that when people tried to solve spatial visualization tasks (such as mental rotation tasks and paper folding tasks), they spontaneously produced gesture to help themselves think. The gesture produced actually helped enhance their performance. Alibali, Spencer, Knox, and Kita (2011) found out that gesture can also influence the strategy people chose for problem solving. In their study, when solving gear movement problems, participants who were allowed to gesture were more likely to use perceptual-motor strategies. Meanwhile, those who were not allowed to gesture were more likely to use the parity strategy instead.

Gesture also benefits reasoning. Schwartz and Black (1996) reported that adults spontaneously produced gestures when making inferences and constructing mental models to solve gear problems. Even children as young as five years old can be taught to gesture to reason. Ehrlich, Levine, and Goldin-Meadow (2006) provided evidence showing that gesturing helped children with their spatial reasoning, and even improved their mental rotation skills. Gesture can also benefit reasoning in a domain that is even not inherently spatial at all. For example, Beaudoin-Ryan and Goldin-Meadow (2014) found that in a moral reasoning lesson, fifth-grade children who were told to gesture produced significantly more responses involving multiple perspectives in speech than children who were told not to gesture. Gesturing successfully helped them to think “on one hand”, and also “on the other hand”.

## 2. Using gesture in learning

Gesture can be used in a lot of cognitive activities, but how is gesture connected with learning in particular? Goldin-Meadow (2010) proposed that gesture is connected with learning in two ways. One way is that gesture reflects an individual’s knowledge state. The other is that gesture alters people’s cognitive state and therefore promotes learning and understanding.

### *a) Gesture reflects the state of learning*

A series of studies provided evidence supporting that gesture serves as a window to a learner’s knowledge state. Church and Goldin-Meadow (1986) conducted a study in which five- to eight-year-old children were asked to explain their judgments about quantity invariance in a Piagetian conservation task. They reported that when explaining, some children’s gestures contained different information from their accompanying speech. For example, they said “The dish is wide”, while gesturing both the shortness and the wideness of the dish. The researchers found that children who produced explanations with mismatch in speech and gesture gained more improvements in the later training session than children whose speech and gesture matched. The authors argued that these mismatches between gesture and speech may reveal that these children were at the edge of making conceptual progress, therefore they were more receptive to training on this concept. This transitional stage can be observed in older children when they learned about mathematical equivalence concepts as well (Perry, Church, & Goldin-Meadow, 1988). In the same vein, researchers proposed that the mismatch between gesture and speech can be an index of transitional knowledge in other context (Perry et al., 1988), and that the mistach can offer insight into understanding the problem-solving process in adults and children alike (Garber & Goldin-Meadow, 2002).

The gesture-speech mismatch theory not only provided information to help researchers identify individuals who are at a transitional stage of learning and are ready to learn, it could also reflect people's mental representation during problem solving. Alibali, Bassok, Solomon, Syc, and Goldin-Meadow (1999) reported that when people were talking about how they solved a math word problem, the information conveyed in their gesture and speech did not always match. From information represented from people's verbal description and gesture, the researchers found that gesture could reinforce, be neutral, or conflict with speech. They even found that the strategies people used to solve problems varied systematically as a function of how those problems were represented in both speech and gesture. Therefore, they proposed that gesture and speech together could provide an index of people's mental representation of problems.

Gesture reflects a learner's knowledge state and even untrained adults could glean information of learners' knowledge state based on their speech and accompanied gestures (Goldin-Meadow, Wein, & Chang, 1992; Perry, Woolley, & Ifcher, 1995). All of these lay out the potential for educators and instructors to gather clues from a learner's gesture and use them to inform teaching and learning.

### *b) Gesture promotes learning*

Goldin-Meadow (2010) stated that gesture can also promote learning in (at least) two ways: (1) by influencing the communicative input for learners, and (2) by altering people's cognitive state directly.

Since gestures can reveal knowledge state and help instructors to identify who is ready for instruction, Goldin-Meadow and Singer (2003) conducted a study to investigate if adults can change their teaching according to learners' gesture. In the study, adults were asked to teach children who could not yet solve mathematical equivalence problems. They found that the adults offered more instruction strategies to children who produced mismatches than to children who produced no mismatches. Their findings supported the hypothesis that instructors could tailor their instructional input to help students learn.

Gesture can also alter people's cognition in a more direct way. Cook, Mitchell, and Goldin-Meadow (2008) found that requiring children to gesture a strategy to solve math problems prior to a math lesson and then asking them to produce the gestures themselves can actually help them learn and retain the knowledge. On the other hand, requiring children to speak out the strategy without gesturing did not make a difference. Their finding suggested that the act of gesturing itself was actively involved in the construction of new knowledge and led to learning. Similarly, Jamalian (2014) reported that preschoolers who were required to perform a grouping gesture when learning counting strategies significantly outperformed those in the control group. She suggested that gesture, which is actions by nature, could add one more layer of meaning by presenting information in two modalities (i.e., visual and motor).

## 3. The mechanism of gesture's effects on learning

Gesture plays an important role in cognition and learning, but what are the mechanisms that underlie the effect? A growing number of studies have provided evidence accounting for the effect from different angles. I summarized them into the following four main mechanisms: gestures can: (1) be used as a tool for representation, (2) add information to people's mental representation, (3) bring out implicit and new knowledge, and (4) offload working memory and lighten cognitive load.

### *a) Gesture can be used as a tool for representation*

Gesture, as action in space, can be used to model space. Emmorey, Tversky, and Taylor (2000) found that when describing scenes with landmarks and routes, English speakers and American Sign Language users adopted survey and route perspectives differently and used different gestures. In this case, gesture is used to represent space.

Gesture can represent one's motions on an object, and grounds mental simulations in actions. Beilock and Goldin-Meadow (2010) conducted a study in which participants were asked to explain how they solved the Tower of Hanoi problems. For some participants, the size and weight of the disks were switched (i.e., the smallest disk was the heaviest and could not be lifted with one hand). They found that the switch group's performance was hindered, because the switch made the perception and action incompatible. The authors proposed that their finding led support to the hypothesis that gesture can help ground people's mental representation in action via its representational properties.

Gesture can also represent perceptual-motor information. Alibali et al. (2011) asked participants to solve physics problems in which they had to predict the direction of gear movement. They found that participants who were prohibited from gesturing used an abstract strategy (the parity strategy) more often. Meanwhile, those who were allowed to gesture used perceptual-motor strategies more often, which were based on simulation of gear movements. They reasoned that the use of gesture can help highlight and structure the perceptual-motor information and thereby made such information more likely to be used.

### *b) Gesture adds information to people's mental representations*

Gesture adds perception information to people's mental representation. Beilock and Goldin-Meadow (2010) asked participants to solve a Tower-of-Hanoi task (TOH1), explain (with gesture) how they solve it, and then solve another Tower-of-Hanoi problem (TOH2). In TOH2, for some participants (in the experimental group), the disk weights were switched, with the smallest disk being the heaviest one that cannot be lifted with just one hand, and the largest disk being the lightest. They found that for the experimental group, the more gestures depicted moving the smallest disk, the worse people's performance was on TOH2. However, for the participants who worked with regular disks (in the control group), their performance in TOH2 was not negatively affected. Furthermore, if participants skipped the explanation step and did not gesture, their performance was not affected either. The authors reasoned that gesture not only grounded people's mental representation in action, but also added the information of the weight of the smallest disk to people's mental representation. Therefore, when the weight information contradicts with size, it interfered with thinking and performance.

Gesture highlights perceptual information for speakers. Alibali and Kita (2010) asked children to explain how they solved a Piagetian conservation task. They found that when children were prohibited from gesturing, they expressed more non-present information and less perceptually present information.  They proposed that gesture promoted thinking on perceptually present information, and therefore influence a speaker's decision on what to talk about.

### *c) Gesture brings out implicit knowledge and new knowledge*

Gesture can bring out implicit knowledge. A study by Broaders, Cook, Mitchell, and Goldin-Meadow (2007) showed that when students were asked to gesture, those who had been unable to solve the problems often added new and correct problem-solving strategies to their explanations. These added strategies were expressed only in gesture, but not in speech. Furthermore, in later instruction, children who were asked to gesture learned better than those who were asked not to.

Gesture can bring out the emergence of new knowledge. Boncoddo, Dixon, and Kelley (2010) examined gesture's representational function from a developmental perspective. They found that preschoolers, when trying to solve simple gear problems, initially used gesture to simulate the movement of gears. Based on their own gestures, most of them discovered the abstract rule (i.e., turning direction of gears alternates) on their own later. Their results provided support to the embodiment hypothesis that gesture, as an embodied action, can promote the development of new representation.

### *d) Gesture lightens cognitive load*

Another explanation of gesture's effects on learning is through the cognitive load theory.  The *cognitive load theory* is primarily based on our knowledge of human cognitive architecture, and the limited capacity and duration of the human working memory (Paas & Sweller, 2012). Its history traces back to Miller (1956)'s experiments on the limitation of the human working memory capacity, and Chase and Simon (1973)'s theory on how human chunks memory components into schema to organize information. In the late 1980s, the theory was outlined and further developed by Sweller (1988) out of a series of studies on problem solving. The cognitive load theory has had wide implications for instructional design and has provided guidelines to optimize learning conditions and instructional materials. This theory differentiates people's cognitive load into three types: *intrinsic cognitive load*, *extraneous cognitive load*, and *germane cognitive load*. The intrinsic cognitive load is the inherent level of difficulty associated with a specific instructional topic. The extraneous cognitive load is generated by the manner in which information is presented to learners and could be controlled by instructional designers (Chandler & Sweller, 1991). The germane cognitive load is the load devoted to the processing, construction, and automation of schemas (Sweller, Van Merrienboer, & Paas, 1998). According to cognitive load theory, an ideal instructional design will limit the extraneous cognitive load and promote the germane cognitive load.

There are a lot of different methods to measure cognitive load, directly or indirectly. Among all the objective methods to measure cognitive load, a widely-used one is the dual-task paradigm, which is a method to assess a learner's working memory load using a secondary task (Britton & Tesser, 1982; Kerr, 1973) in combination with a primary task. The secondary task usually requires learners to engage in an additional cognitive activity (e.g., holding irrelevant numbers or letters in mind) that is secondary to the primary task of learning. If the primary learning task imposes a heavy cognitive load, performance on the secondary task deteriorates. In contrast, a low cognitive load in the primary task will result in improved performance on the secondary task (Sweller, Ayres, & Kalyuga, 2011). In literature, the secondary task could be visual (Brünken, Steinbacher, Plass, & Leutner, 2002), auditory (Brünken, Plass, & Leutner, 2004), or verbal (Myerson, Hale, Rhee, & Jenkins, 1999).

Empirically, this dual-task paradigm has been widely used in studies to demonstrate gesture's cognitive benefits in offloading cognitive load. For example, in a study, Goldin-Meadow, Nusbaum, Kelly, and Wagner (2001) observed adults and children explaining their solution to a math problem (primary task) while trying to hold a list of letters or words in memory (secondary task). They expected that if gesture lightened cognitive load during speech, more mental and cognitive resources would be saved for the memory task to allow speakers to perform better. They found that both adults and children remembered more letters or words if they gestured during explanation compared to those who did not gesture, regardless of whether they were instructed not to gesture or spontaneously chose not to gesture. They argued that gesturing while talking helped lighten speakers' cognitive load so that they had more cognitive resources allocated to the memory task. Similar results were found in Wagner, Nusbaum, and Goldin-Meadow (2004)'s study as well. In the study, they asked adults to hold strings of letters or visual grid patterns in memory (secondary task) while explaining how they solved factoring problems (primary task). They found that participants remembered significantly more items when they gestured than when they did not gesture, regardless of whether the memory task was a verbal memory task or a visual memory task. Furthermore, Ping and Goldin-Meadow (2010) also asked children to hold two words in their memory (secondary task) while explaining answers to a Piagetian liquid quantity conservation task (primary task). Their study suggested that gesture's cognitive benefits could be found not only when speakers used gestures to refer to objects that were visible in the immediate environment—the benefits continued to be found even when speakers talked and gestured about objects that were not present and could not be directly indexed by gesture. These studies all showed that gesture lightened people's cognitive load.

## 4. Variation in gesture and individual difference

Variation in gesture is another important area in the gesture literature. Research has proposed some factors associated with gesture variation. Researchers have found that people's cognition, thinking, language and culture could all influence the usage of gesture. Different gestures benefit different cognitive activities.

### *a) Variations of gesture by people*

Many studies suggested that people's visual and verbal abilities are correlated with individual differences in gesture production. Bucci and Freedman (1978) found that individuals with high referential competence (a type of verbal ability) produced much more representational gestures than individuals with low referential competence. Vanetti and Allen (1988) also looked for differences among their participants divided into four even groups: high spatial high verbal, high spatial low verbal, low spatial high verbal, and low spatial low verbal. They found that participants with high spatial and low verbal ability produced the highest number of representation gestures, compared to participants from other groups, although the difference was not significant.

Cognitive abilities also correlate with how people use gesture to encode information. In a more recent study, Göksun, Goldin-Meadow, Newcombe, and Shipley (2013) found that high-spatial individuals used gesture in a different way than low-spatial individuals. They showed that high-spatial individuals were more likely to use gesture to encode the internal structure of target blocks, while low-spatial individuals tended to use gesture to convey the static state of those rotation blocks instead of the dynamic form. Furthermore, when the low-spatial participants used static gestures, these gestures were often iconic gestures that highlighted the entire structure of the blocks (e.g., a curved handshape gesture), whereas the high-spatial participants' gestures emphasized the internal structural relationship of the blocks (e.g., an L-shaped gesture).

Cultural background influences gesture variation in an even more complex way. Kita (2009) reviewed four different factors governing the variation: (1) conventions for form-meaning associations, (2) cognition, (3) language, and (4) pragmatics for communication. These four factors dictate issues such as how gestures are associated with meaning, how gestures are used representationally for motion, time, and space, how cross-linguistic differences causes gesture differences, and gestural pragmatics (e.g., gesture rate, gesture size, gesture space, and gestural politeness).

### *b) Variations of gesture in cognition and learning*

Gestures are different in terms of their rate, size, encoding, and could be used for different communicative purposes. Variations in gesture could have different effects on cognition and learning.

While some gestures help cognition and learning, some do not. Regarding speech-accompanying gestures, Cook, Yip, and Goldin-Meadow (2012) proposed that only gestures that were in coordination with the content of speech lightened working memory, but meaningless hand movements that were produced rhythmically with speech did not. Regarding gestures for thinking, Göksun et al. (2013)'s study suggested that gestures that highlighted the inner structure of objects but not other features promoted people's performance on a mental spatial transformation task.

Some gestures promote learning only when they are compatible with the learning content they represented. Segal (2011) proposed a concept of *Gestural Conceptual Mapping*, which hypothesized that only the gestures that were congruent with the learning concept would promote learning. She provided evidence showing that when preschool children learned about arithmetic (a discrete task), discrete but not continuous gestures supported learning the best; for estimation (a continuous task), continuous but not discrete gestures supported learning the best. She argued that the “right” gestures should be congruent with learning concepts, and compatible with the mental representation and operations needed to solve problems. These gestures would elicit the best performance. Another example is Kang (2012)’s study, which showed that different types of gestures primed different types of knowledge. In his study, he asked participants to learn from an instructional video about how an engine system worked. The instructor in the video taught the topic with either iconic gestures or action gestures. He found that iconic gestures that highlighted the structural knowledge of the system helped learning of the system’s structure, and action gestures that highlighted movements helped learning of the casual relationship between components within the system. He reasoned that the type of instructional gesture influenced a learner’s mental representation of the complex system.

Good gesture promotes learning, while bad gesture may impede it. Jamalian (2014) pointed out that in preschool children, although pointing gestures accompanying counting can be assumed as an accurate counting strategy, they may possibly impede children’s understanding of the cardinality concept. For example, when a child counts a set of apples with pointing gesture and recites “one, two, three, four, five,” he may fail to realize that “five” refers to the cardinality of the whole set, rather than just the fifth apple. Therefore, simple and repeated pointing gestures can fail to represent the concept of *set* as a collection. However, asking children to count and add a grouping gesture around the items to highlight the concept of *set* would promote their understanding of cardinality and improve their overall math competence.

Some gesture leads to cognitive changes, but not all. Goldin-Meadow (2010) stated that in many cases, people produced gestures that contained and reflected more information than its accompanying speech. These meaning-loaded gestures can translate differently for novices and experts. For experts, the added information may be only an adjustment to small variations in the discourse, so these gestures would not lead to learning. However, for novices of knowledge in the speech, this added information in gesture may mean a developmental difference. Their gestures could reflect speakers' experimentation with "new and not-yet-solidified ways of solving a task," therefore had the potential to lead to cognitive changes.

Since not all gestures promote learning equally and gestures are different in nature, I propose an effectiveness spectrum (see Figure 1). This spectrum demonstrates that we should not see gesture dichotomously as effective or ineffective when gesture is used for thinking and learning, but as a spectrum from very ineffective to very effective.


***Figure 1.*** Gesture Effectiveness as a spectrum.

### *c) Co-speech gesture and co-thought gesture*

In the gesture literature, *co-speech gestures* (communicative gestures that accompany concurrent speech) have already been very widely studied (e.g., De Ruiter, 1998; Kita, 2000; Krauss et al., 1996; Krauss et al., 2000; Wesp et al., 2001).

However, people also produce a lot of *co-thought gestures*, which are the gestures that are produced when people think in silence. As Chu and Kita (2011, 2015) pointed out, there have been very few studies that explicitly explore the role of co-thought gestures in spatial problem solving. The mechanisms underlying the production of co-thought gestures are also largely unknown.

From the very limited number of studies that specifically focus on gesture used in problem solving, we know that spontaneous gesture can reveal the strategies people chose to solve problems (Chu & Kita, 2008). For example, Schwartz and Black (1996) found that when solving gear movement problems, the participants transitioned from using a perceptual-motor strategy (e.g., depicting the movements of each gear) to using an abstract rule-based reasoning strategy (e.g., reasoning based on whether the number of gears was odd or even). This change was reflected in the decrease of gestural depictions of gear movements over the course of the experiment. Chu and Kita (2008) also reported that when solving mental rotation problems using their hands, people tended to use gestures depicting their manipulation on an object at the beginning, and then moved to gestures depicting object movements only. They also reported that people produced more manipulation gestures in earlier trials of the experiment than in later trials.

Co-thought gestures not only reflect the strategy people use when solving mental problems, they also influence people's performance in problem solving. In a study by Schwartz and Black (1999), participants had to imagine two glasses of different diameters filled to the same level with water. They were asked to judge which glass would start to spill first if tilted. The researchers found participants rarely answered the questions correctly when asked verbally. However, when the participants were asked to close their eyes, tilt the glasses with hands, and imagine the water level, they produced the correct answers more frequently. Consistent with this study, Chu and Kita (2011) also reported that when people had difficulty solving visual spatial problems, they spontaneously produced gestures to help themselves, and gestures can indeed improve their performance. Moreover, people who were encouraged to gesture also performed better than people who were prohibited from gesturing and people who were allowed but not encouraged to gesture.

Although a considerable amount of research has consistently shown that co-thought gesture can be used to benefit thinking and learning as effectively as the co-speech gesture, the mechanisms of co-thought gestures' benefits on thinking and learning are still relatively understudied. Especially, I think it will be beneficial to further explicitly explore how to involve co-thought gestures in the field of education as an instructional approach.

## 5. Using gestures and body movements for chemistry learning

The literature on embodied cognition, gesture and body movements, and the application of embodied cognition as an instructional approach has shown the potential of using gestures and body movements in teaching spatial thinking across science, technology, engineer, and mathematics (STEM) disciplines. In this dissertation study, I chose chirality in chemistry as the learning content based on the following considerations.

First of all, in the past decades, more and more researchers have considered the role of body and movements in visualizing and meaning making in STEM teaching and learning. A growing body of research has demonstrated the positive effects of gestures and body movements in various fields of STEM education. For example, Goldin-Meadow et al. (2009) found that during a math lesson, children who were required to produce correct gestures learned more than children who were required to produce partially correct gestures or no gestures at all. They suggested that when learning something new, body movements were not only involved in processing ideas, they also helped in extracting implicit meaning and creating new ones. Adults can benefit from gestures and body movements as well. By analyzing the gestures produced in an undergraduate physics class, Scherr (2008) found that gesturing helped students articulate emerging ideas, organize information, and facilitate construction of new ideas. Similarly, Singer, Radinsky, and Goldman (2008) also reported that when learning geoscience using a data visualization tool, gesturing made it possible for learners to concretize phenomena that were otherwise not directly observable in space and time and were difficult to be captured by speech alone, therefore promoted their learning of abstract concepts. In general, by functioning as a simulation of the physical world and facilitating the linking between sensorimotor experiences and mental representations, gestures help learners better comprehend abstract and complex ideas (Wilson, 2002).

Second, the discipline of chemistry itself involves extensive study of dynamic spatial relationships in entities at a molecular level, therefore it provides a lot of good opportunities to study how gestures and body movements support thinking and learning in STEM education (Stieff, Lira, & Scopelitis, 2016). Flood et al. (2014) pointed out that many chemistry phenomena are frequently submicroscopic (thus inaccessible to our senses), require visual-spatial thinking in a three-dimensional space, and involve dynamic motion or change. Therefore, gestures of hands and body parts can become a very powerful medium to simulate and enact different vibrational and rotational motions at a molecular level. In this line of thinking, researchers and educators have demonstrated the benefits of explicit training on the use of gesture, as a pathway to support spatial thinking and improve student success. Their research has helped gain insight into instructional techniques and the mechanism of learning. For example, Chu and Kita (2011) showed that students who were encouraged to gesture performed better on mental rotation tests than those who were instructed not to. Stull, Barrett, and Hegarty (2013) demonstrated that manually manipulating virtual molecular models on a haptic device yielded comparable improvements in spatial problem solving than learning with high-fidelity concrete models. Stieff et al. (2016) found that by physically simulating spatial transformations with gestures, students learned equally well as those who learned from concrete models. Their learning was long lasting even when the models were taken away. These studies all supported the important role of gesture in promoting spatial thinking and learning in the field of chemistry.

Third, as an important, foundational, yet complicated concept, *chirality* is a very challenging topic for students in introductory level chemistry courses in college. The concept of chirality refers to a geometric property of some molecules and ions. By definition, if a molecule or ion is non-superimposable on its mirror image, it is chiral; otherwise, it is achiral. In an introductory chemistry course, the concept of chirality is often introduced using a pair of hands and other chiral and achiral objects. However, to learn the concept, students need to represent three-dimensional molecular structures with multiple two-dimensional diagrams of the molecule, and then translate between these diagrams to make analyses. Translating among these diagrams and depicting spatial-relational information to visualize molecular structure can be exceptionally challenging for novices who primarily rely on imagistic strategies or even no strategies at all (Stieff, 2011; Stieff, Dixon, Ryu, Kumi, & Hegarty, 2014; Stieff et al., 2016). Therefore, using chirality as the learning content in this study will have immediate implications for teaching and learning in the classroom.

# Chapter III Study 1

## A. Research questions for this study

A review of literature reveals three reasons for Study 1:

(1) There are variations in gesture and gesture use. Not every kind of gesture promotes thinking and learning.

(2) The mechanism of how gesture promotes thinking and learning is understudied, especially regarding gesture's effects on cognitive load.

(3) A very limited number of studies have explicitly investigated the co-thought gestures.

With this background, I decided to further study the mechanism of gesture's effects on thinking. In this study, I investigated whether and how different types of gestures (including big gestures and small gestures) influenced people's thinking and learning, when people were engaged in problem solving activities in a non-communicative environment.

There are four guiding research questions for this study:

1. How would teaching different types of gestures influence learners' own gesture uses?
2. How would co-thought gestures change over time when people use them to solve problems?
3. How would different types of gestures influence learners' performances?
4. How would different types of gestures influence learners' cognitive load?

## B. Method

### 1. Participants

Thirty-one graduate students from Teachers College, Columbia University participated in this study. They were all at novice level in the learning content. All of them participated this study for course credit.

Participants were randomly divided into two groups. One group ($n$ =15) was taught to use big gestures as a strategy to solve a molecule configuration problem. The other group ($n$ =16) was taught to use small gestures to solve it.

In data analysis, four participants were excluded from the sample. From the small gesture group, one participant was excluded because she completed each of the questions partially, and another participant was excluded because his body blocked the camera, obscuring some of his gestures and making it impossible for us to code his gesture use. From the big gesture group, two participants were excluded because one was playing with the system (keying in the same answer for each question), and the other one spent too little time on each question (the experimenter believed it would be impossible to complete the questions in such a short time). Therefore, the final sample consisted of 27 individuals, with 13 in the small gesture group and 14 in the big gesture group.

### 2. Procedure

Participants first signed a consent form and a form of participants' rights. After they read and signed the documents and assented to participating in the study and being video recorded, they were asked whether they would like to give permission for us to show their videos in presentations of the research.

The main procedures of the experiment included:

1. Participants completed a paper version of the Mental Rotation Task (MRT) (Vandenberg & Kuse, 1978).
2. Participants were divided randomly into two groups. One group was taught how to use small gestures to solve a molecule configuration problem. The other group was taught how to use big gestures to solve it.
3. Participants solved 15 sets of problems. Each set contained a primary task (one molecule problem) and a secondary task (one block tapping problem).
4. Participants completed a NASA Task Load Index (Hart & Staveland, 1988) and an exit survey.

## C. Results and discussion

### 1. Teaching small and big gestures as a problem solving strategy

Out of the 27 participants included in the final data analysis, 13 of them were taught to solve the molecule problems using small gestures. The rest of them were taught big gestures. After teaching, participants were asked to solve problems on their own. They were encouraged to use the taught gestures but were informed that using the taught gestures was not mandatory. Thus, participants voluntarily chose to use small gestures, big gestures, no gestures, and even mixed gestures (a combination of small and big gestures) on each trial.

**Small gestures were more easily accepted after they were taught.** As shown in Figure 2, for the people who were taught big gestures, the number of participants who accepted them was very low. Out of 14, 9 persons (64.29%) used the taught gestures for only a few trials (less than 5 trials). Two (14.29%) of them used the taught gestures moderately (between 5 to 9 trials). Only 3

(21.43%) accepted the instruction and used the taught gestures frequently (in more than 10 trials out of the 15 trials). However, for the people who were taught small gestures, the pattern was reversed. Out of 13, 9 persons (69.23%) applied the taught gestures frequently (in more than 10 trials out of the 15 trials). Two (15.38%) of them used the taught gestures moderately (between 5 to 9 trials). Two (15.38%) of them used the taught gestures for only a few trials (less than 5 trials). Because of the small sample size, the difference between groups was not significant. However, the pattern was very obvious.

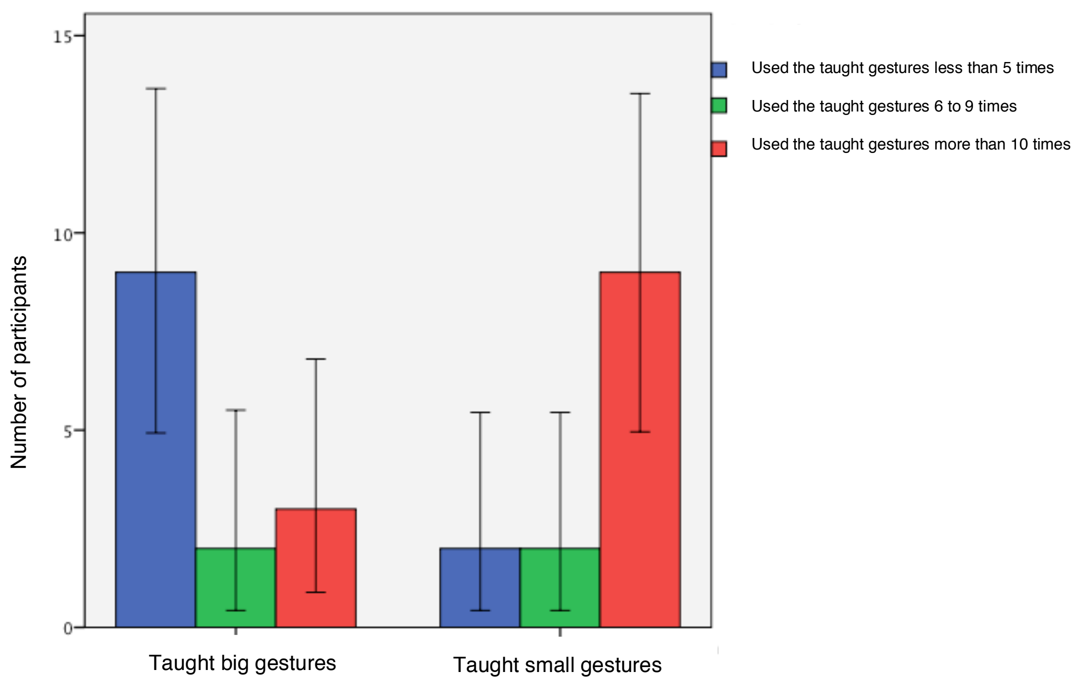



***Figure 2***. Gesture used by participants who were taught big and small gestures. Error bars represent 95% confidence interval.

Similarly, the type of small gesture was the most popular choice by participants from both groups, regardless of what type of gesture was taught. Figure 3 shows a breakdown of all 405 trials by the gesture type and group.

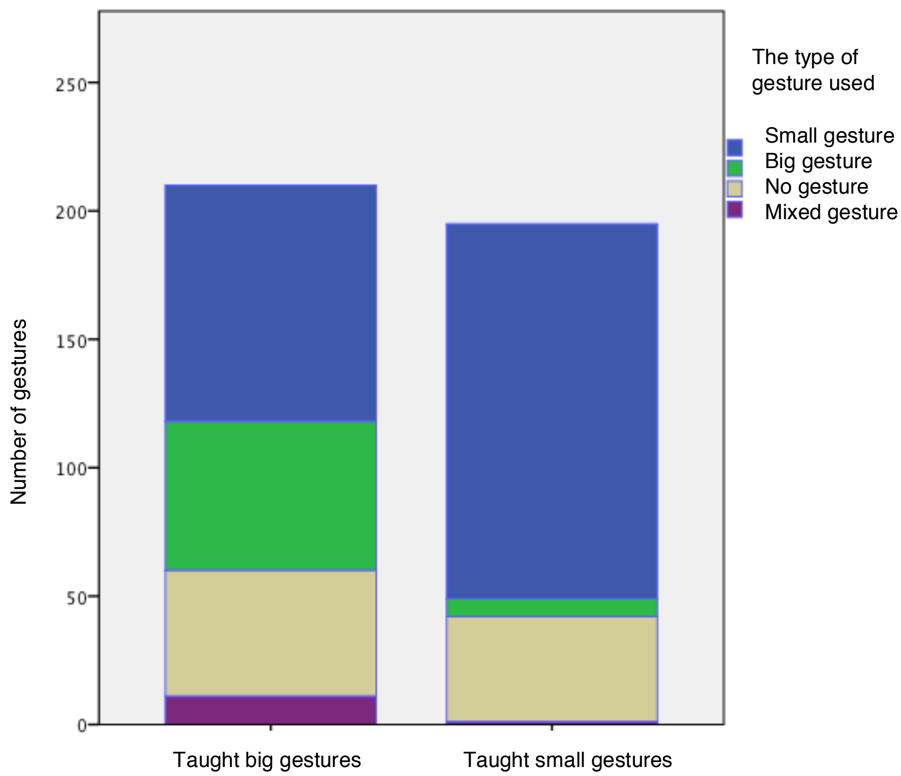



***Figure 3.*** Gesture taught and used by all participants.

Altogether, the type of small gesture was used in 238 trials (58.8%) by participants. From the bar on the right, we can see that those who were taught to use small gestures used them the most frequently in 146 (36.05%) trials. Notably, people who were taught to use big gestures used small gestures the most frequently as well in 92 (22.72%) trials, although they had not been taught this type of gesture at all.

For those who were taught small gestures, the type of *no gesture* was the next most frequent, accounting for 41 (10.12%) trials.

For those who were taught big gestures, the second and third most frequently used gesture types were the *big gesture* and *no gesture*. They were used for similar number of trials, in 58 (14.32%) and 49 (12.10%) trials respectively.

The least frequent gesture type was *mixed gesture*, which was used in 12 trials (3%). Notably, 11 mixed gestures were used by people who were taught big gestures. Only one was used by the group who was taught small gestures.

These results indicate that most people who were taught small gestures seemed to be very satisfied with using this strategy. The type of small gesture was their primary choice. However, for people who were taught big gestures, although a good proportion of them continued to use them when asked to solve problems on their own, an even larger proportion of them switched to small gestures or even no gestures. Interestingly, those who where taught to use big gestures tried mixed gestures on more trials at the beginning than those who were taught small gestures did. I think using mixed gestures could possibly reflect some confusion from the participants when they tried to look for a suitable strategy for themselves.

### 2. Gesture type and performance in the primary task

Figure 4 shows participants' performances on the primary task (the molecule problems) when using small, big, no gestures, and mixed gestures. We can see that when the type of small gesture was used, it resulted in much more correct answers than incorrect answers. A similar pattern can be seen for questions answered by no gestures. However, for questions answered by big gestures and mixed gestures, there did not appear to be a huge difference between the number of correct and incorrect answers.

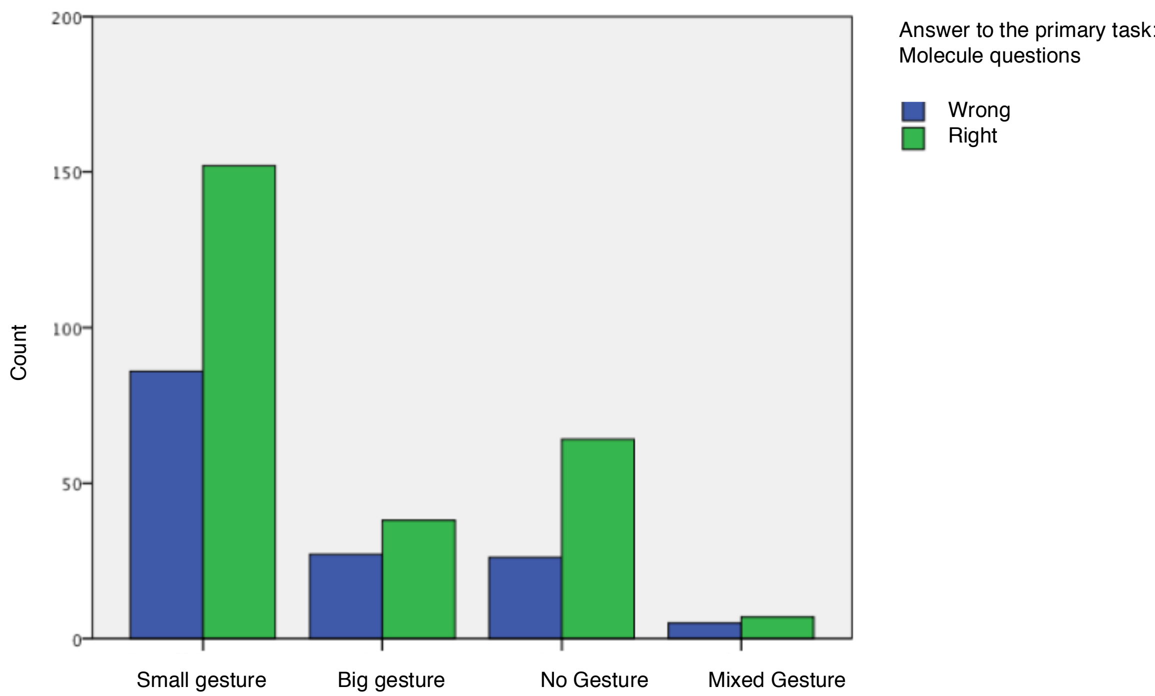



***Figure 4.*** Performance in the primary task by gesture type.

To better understand how gesture type affected the performance score on the primary task, I constructed a binary logistic regression model, using the actual performance on the primary task (*right* or *wrong*) as the dependent variable, the type of gesture used (*small gesture, big gesture, mixed gesture,* or *no gesture*) as a fixed effect and subject ID as a random effect. The model showed that the effect of gesture type on performance was not significant (Table 1)

Table 1

*Effects of Gesture Used on the Primary Task*

| Source | *F* | *df1* | *df2* | *p-value* |
|---|---|---|---|---|
| Corrected Model | 0.507 | 3 | 401 | .677 |
| Gesture used | 0.507 | 3 | 401 | .677 |

### 3. Performance in the secondary task: an objective measure of cognitive load

Figure 5 shows how participants' performance on the secondary task (the block tapping problem) was influenced by using small, big, no gestures, and mixed gestures.

The range of the scores for the secondary task was 0 to 5. I found that when small gestures were used to solve problems in the primary task, the mean score of cognitive load ($M$ = 3.56, $SD$=1.18) associated with the questions correctly answered was higher than the mean score of cognitive load associated with the questions incorrectly answered ($M$=3.10, $SD$=1.19). Similarly, for the no gesture group, the mean score of cognitive load ($M$= 3.51, $SD$=1.23) associated with the questions correctly answered was also higher than the mean score of cognitive load associated with the questions incorrectly answered ($M$=2.77, $SD$=1.25). However, this pattern was not shown in the questions answered by big gestures and mixed gestures.

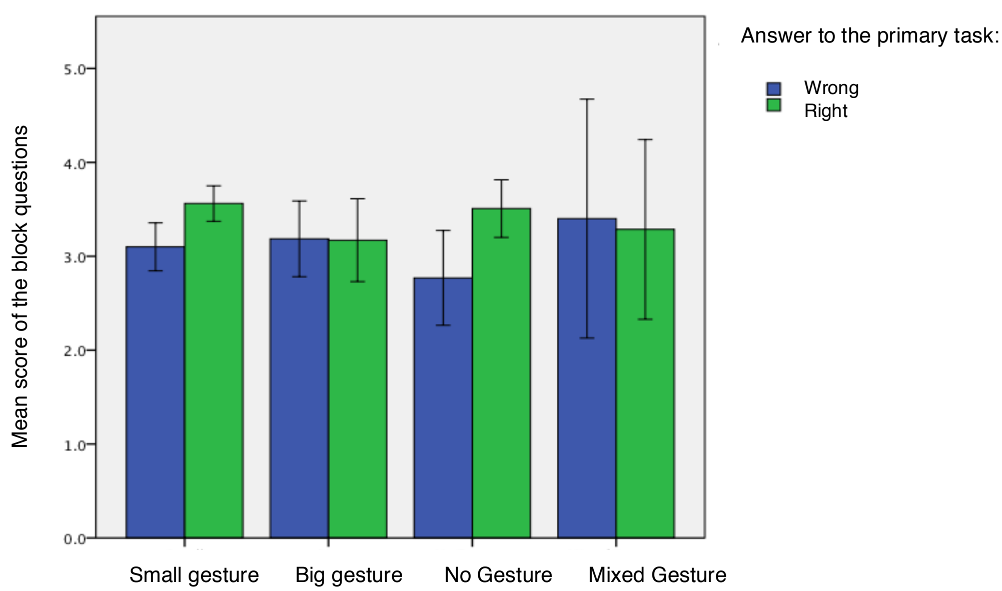



***Figure 5***. Performance in the secondary task by gesture type.

To understand how the type of gesture used affected the performance score on the secondary task, I constructed a binary logistic regression model (a generalized linear mixed-effect model), using the actual performance on the secondary task as the dependent variable, the type of gesture used (*small gesture, big gesture, mixed gesture,* or *no gesture*), answer to molecule questions (*right* or *wrong*), and the interaction between these two variables as fixed effects, and subject ID as a random effect. The model showed that the effects were significant (Table 2).

Table 2

*Effects of Gesture Used on the Secondary Task*

| Source | *F* | *df1* | *df2* | *p-value* |
|---|---|---|---|---|
| Corrected Model | 3.064 | 7 | 120 | .005 |
| Gesture Used | 2.315 | 3 | 120 | .079 |
| Answer to MQ | 0.014 | 1 | 120 | .907 |
| Answer to MQ * Gesture | 1.596 | 3 | 120 | .194 |

The above results indicate that the type of gesture used might possibly have some effects on performance on the primary task, but also on the cognitive load as shown in the secondary task.

## 4. Gesture duration time, thinking time, and gesture density

Figure 6 shows that the total time spent on thinking on each primary task decreased significantly and steadily as people worked from question 1 to question 15. Regardless of the type of gesture taught, the gesture time decreased in a very similar pattern.

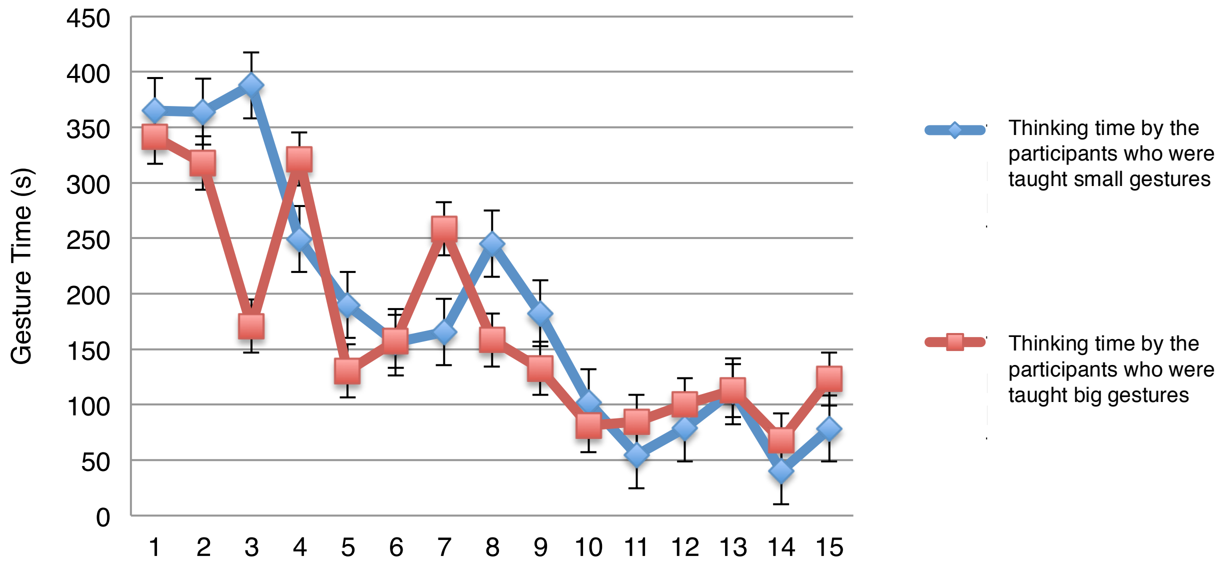



***Figure 6.*** Total time spent on thinking on primary task by participants who were taught small and big gestures. Error bars represent standard error.

Figure 7 shows that average gesture time and average thinking time both decreased from question 1 to question 15.

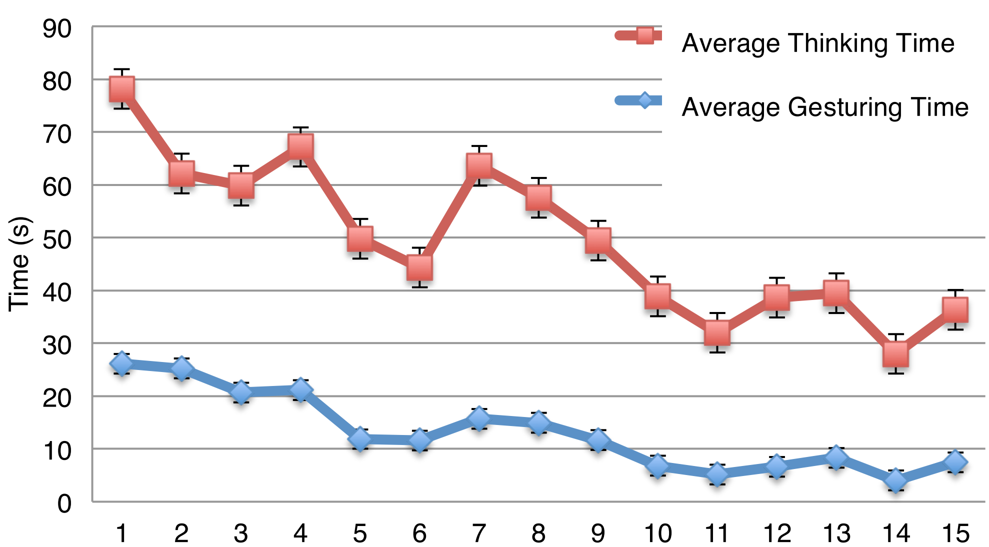



***Figure 7.*** Average gesture time and average thinking time for primary task. Error bars represent standard error.

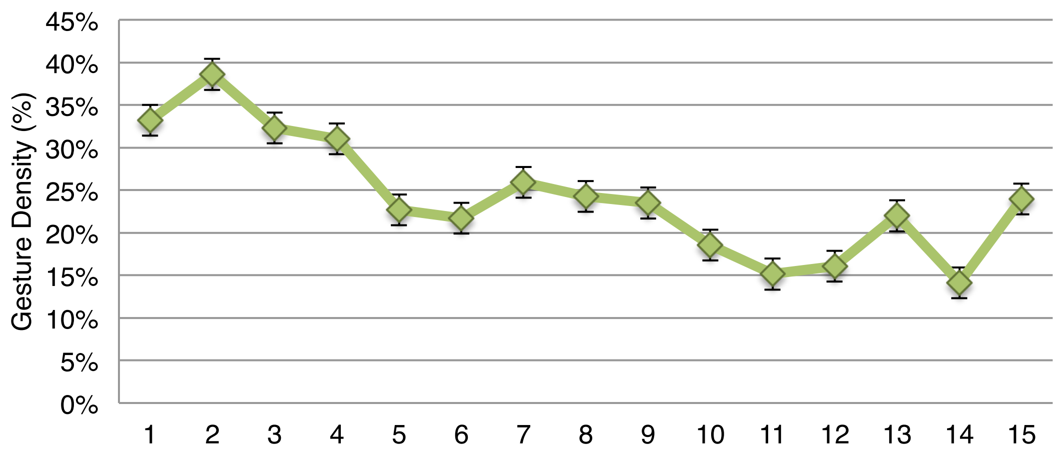



***Figure 8.*** Average gesture density for primary task. Error bars represent standard error.

Figure 8 shows that gesture density (gesture density = gesture time/thinking time) also decreased from question 1 to question 15.

Together, the above results show that from question 1 to question 15, people spent less time thinking and less time gesturing. Their gesture density decreased in the same trend as well.

In terms of the type of gesture, Figure 9 shows that the number of participants using small, big, and mixed gestures for each primary task decreased slowly from the beginning to the end.

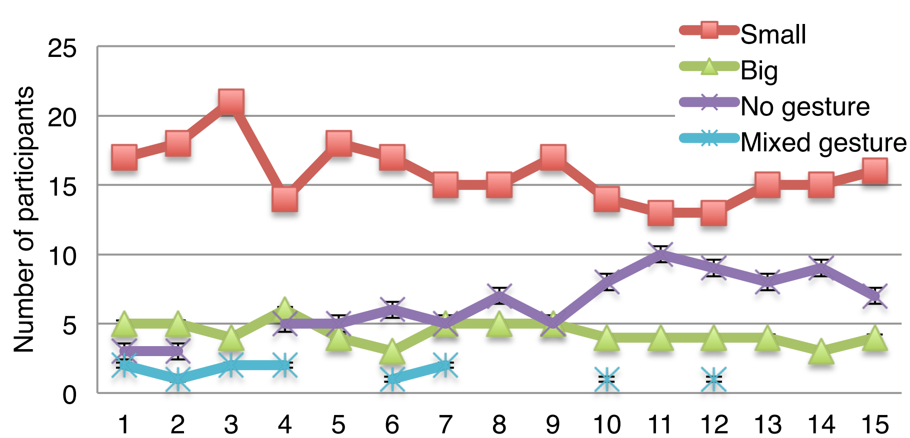



***Figure 9.*** The type of gesture used for the primary task. Error bars represent standard error.

Interestingly, from question 1 to question 15, the number of participants who used no gestures increased steadily and significantly. It seemed that into the second half of the time (questions 8 to 15), a small number of people gave up small and big gestures and switched to no gestures. In addition, a few people tried the mixed gestures in the first half of the trials, but few used them during the second half. It seemed that people tried a mixture of small and big gestures for a few trials at the beginning, but dropped them later.

To understand the effect of the type of gesture taught on the type of gesture used, I constructed a multinomial logistic regression model using the type of gesture used (*small gestures, big gestures, mixed gestures*, and *no gestures*) as the dependent variable and the type of gesture taught (*small gestures, big gestures,* and *no gestures*) as the fixed effect. The model showed that the type of gesture taught significantly influenced the type of gesture used (Table 3).

Table 3

*Effects of Gesture Taught on Gesture Use*

| Source | *F* | *df1* | *df2* | *p value* |
|---|---|---|---|---|
| Corrected Model | 15.142 | 3 | 399 | p <. 001 |
| Gesture Taught | 15.142 | 3 | 399 | . p <. 001 |

### 5. Accuracy percentage

Figure 10 shows that the accuracy percentage for the primary task remained steady and even increased slightly towards the end, in spite of a few dips (in No.4, No. 8, and No. 13). This trend, together with the decrease of time spent on each trial, suggested that people became more and more skillful at solving the problems, because they managed to keep a steady accuracy percentage while spending less and less time on each trial.

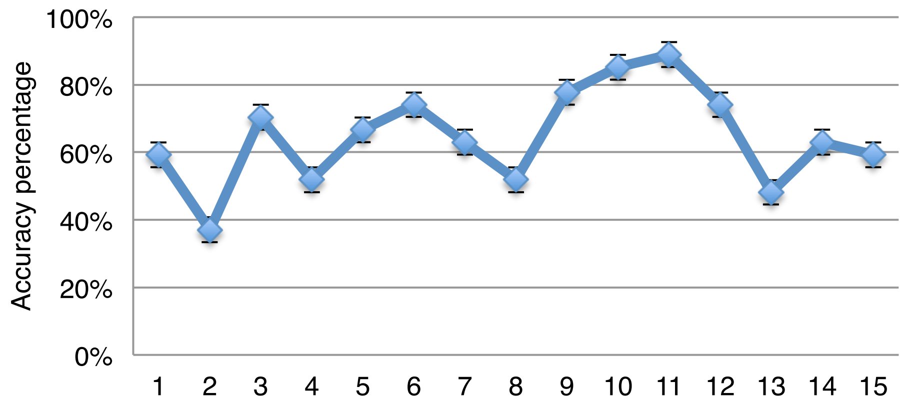



***Figure 10***. Accuracy percentage for the primary task. Error bars represent standard error.

Participants were broken up into the four sub-groups TBUS (taught big used small), TBUB (taught big used big), TSUS (taught small used small), and TSUB (taught small used big). As shown in Figure 11, we can see that participants who were taught small gestures and used small gestures performed the best. In the TSUB group, we do not have a calculation for accuracy percentage because there were too few people (*n* = 4).

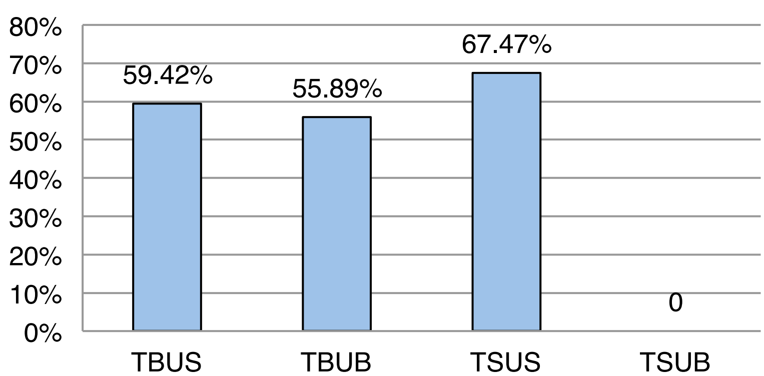



***Figure 11.*** Average accuracy percentage for different gesture types.

### 6. Subjective Measure of cognitive load

For each participant, when they solved the molecule questions, the type of gesture (*small, big,* and *no gesture*) were recorded and coded.

The result showed that our participants rarely consistently used one type of gesture throughout the 15 questions. Many of them started with one certain type of gesture, switched to a different type of gesture for some questions now and then, and gradually dropped all gestures as the time moved on. Some switched between big and small gestures on one question during thinking. Others did not even use gesture at all. In my coding, if a participant used one type of gesture to answer more than two thirds of the questions, this individual was coded as the gesture user of this type. For example, throughout the 15 questions, if a participant used small gestures on more than two thirds of the questions, he was coded as a *small gesture user*. If a participant did not use any particular type of gesture for more than two thirds of the questions, he was coded as a *mixed gesture user*. So all participants were coded into the following four types of gesture users*: small gesture user, big gesture user, no gesture user,* and *mixed gesture user* (Figure 12).

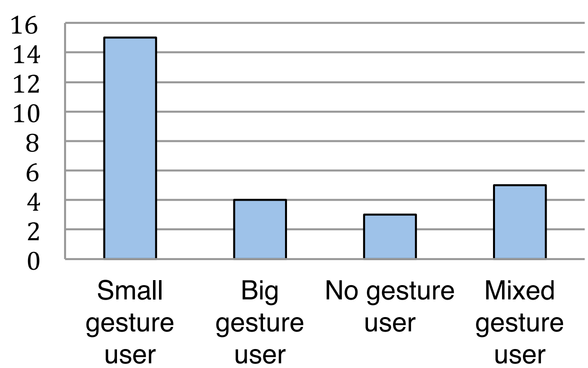



***Figure 12.*** The four types of gesture users.

Each individual completed a NASA Task Load Index, which was used as a subjective measure of cognitive load. The subjective measure of cognitive load revealed several interesting findings from the four types of gesture user. Figure 13 shows the results.

When asked “How physically demanding was the task?,” not surprisingly, people who were *big gesture users* and *mixed gesture users* reported the highest level of physical demandingness followed by the *small gesture users*. *No gesture users* reported the lowest level of physical demandingness.

When asked “How mentally demanding was the task?,” *no gesture users* reported the highest level of mental demandingness, with the other users not very different from each other.

When asked “How successful were you in accomplishing what you were asked to do? (performance),” it seemed that *mixed gesture users* were more confident in their own performance, with other groups thinking alike. Maybe it was the switch between gestures that made them feel they were trying different strategies and led to this higher level of perceived accomplishment.

When asked “How hard did you have to work to accomplish your level of performance? (effort),” the *small gesture users* and *big gesture users* were not vert different from each other. However, *the mixed gesture users* thought they had expended the most effort to accomplish their level of performance. *No gesture users* thought they had expended the least effort.

When asked about frustration level, “How insecure, discouraged, irritated, stressed, and annoyed were you?,” *big gesture users* reported the lowest frustration level. It is consistent with our intuition that when movements become bigger or when a large portion of the body is engaged, people tend to feel that their actions on learning are playful, game-like, or more fun, thus the frustration level was lowered. The *small gesture users'* levels of frustration were higher than those of *the big gesture users.*

When asked "How hurried or rushed was the pace of the task?," *big gesture users* thought they were much more hurried throughout the task than the other groups. However, as shown in Figure 14, they actually took neither more thinking time nor more gesture time than *small* and *mixed gesture users*. Interestingly, these *no gesture users* thought they were the least rushed, but they took the longest thinking time and shortest gesturing time out of all groups.

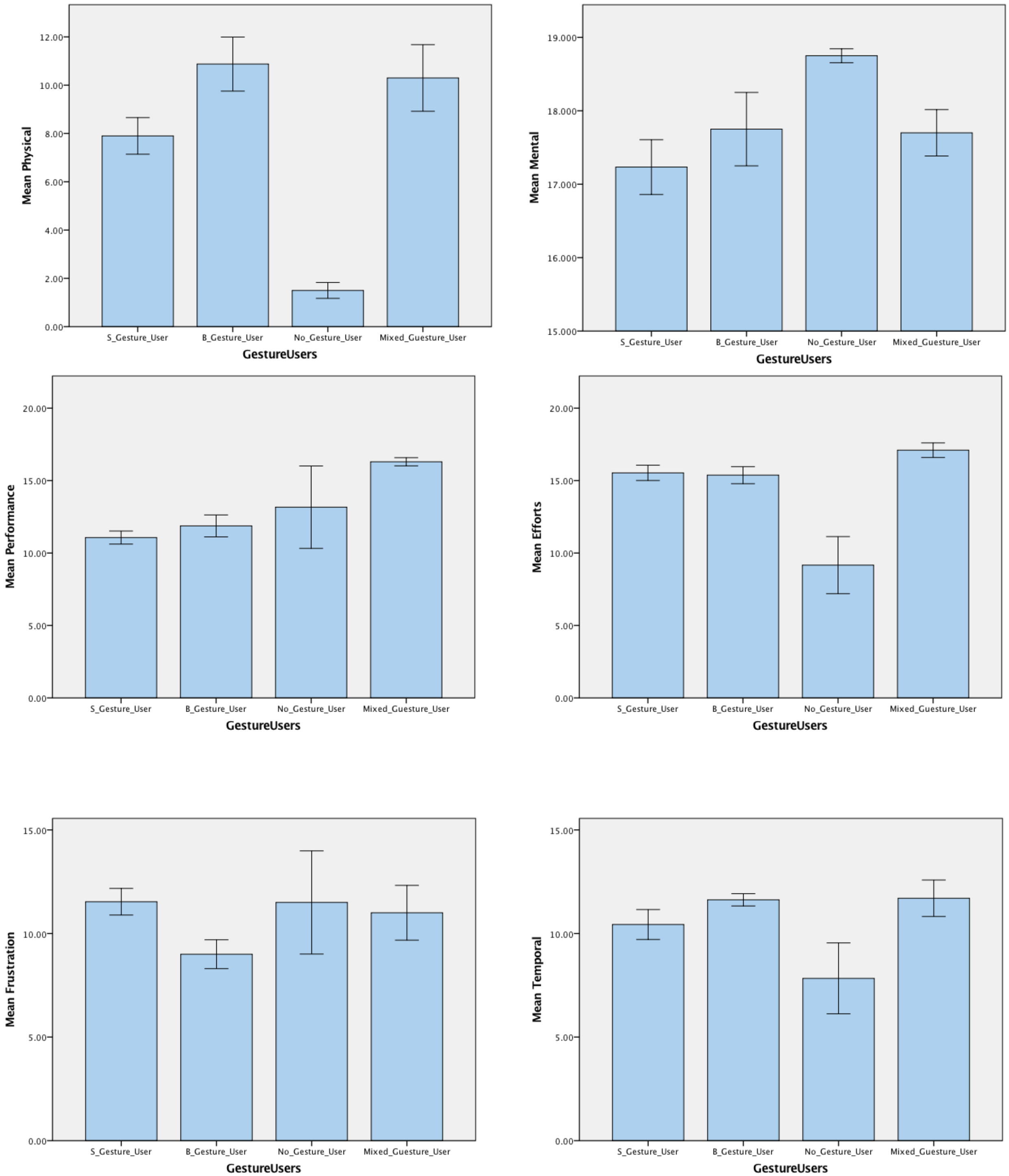



***Figure 13.*** Subjective measure of cognitive load by gesture user type. Error bars represent 95% confidence interval

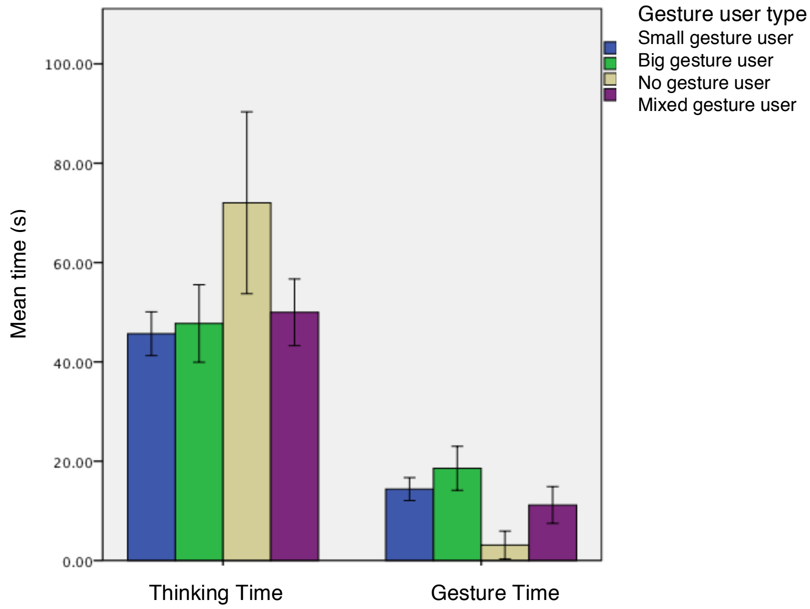



***Figure 14***. Thinking time and gesture time by gesture user type. Error bars represent 95% confidence internal.

## 7. Individual difference

The Mental Rotation Test (MRT) has been used in a lot of studies to assess participants' visual spatial ability. Some researchers used MRT to identify individuals as high ability if they scored > 50 %, and low ability if they scored $\leq$ 50 % (Geiser, Lehmann, & Eid, 2006). Others used the median split. Here, I used the median split at score = 9.00 to divide participants into low- and high-scoring group because of the distribution of scores in this sample (*Range* = 0 -18, *median* = 9). See Figure 15.

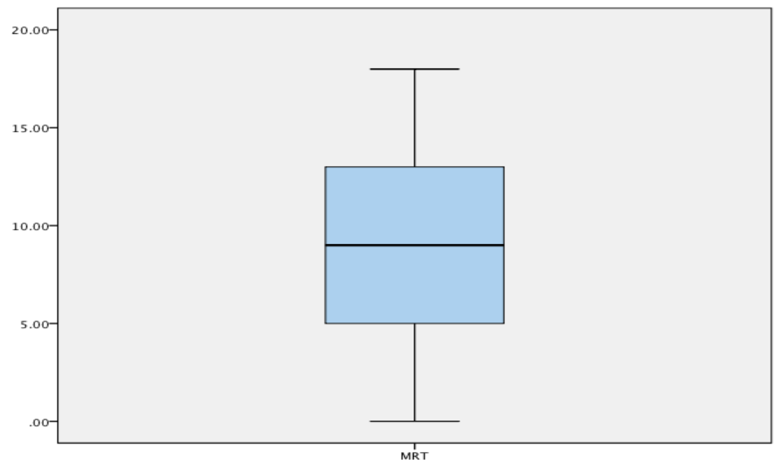



***Figure 15.*** Participants' Mental Rotation Task score.

Figure 16 shows that there was an interaction between the type of gesture user and the level of visual spatial ability. In high-spatial ability participants, *small gesture users* and *no gesture users* performed significantly better than *the big gesture users and the mixed gesture users.* However, in low-spatial ability participants, *the big gesture users* performed significantly better than *small gesture users* and *mixed gesture users*.

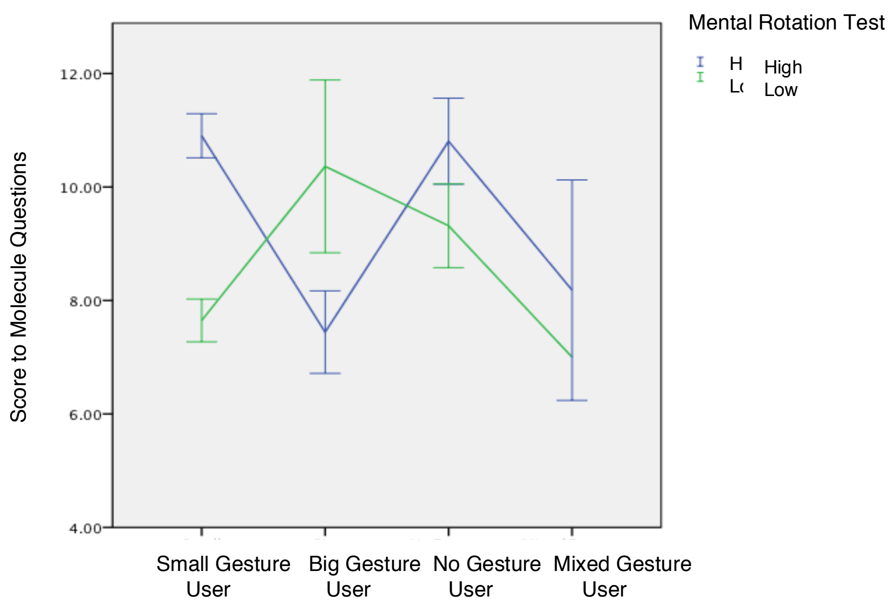



***Figure 16.*** Gesture user type and their mean scores of the primary task. Error bars represent 95% confidence interval.

### D. Discussion

In this study, participants were taught two types of gestures (*big* or *small*) and were asked to use them to solve molecule configuration problems. They were left alone in a quiet room to solve the problems on their own in a non-communicative setting. Their performance on the primary task (the molecule question) and on the secondary task (the block question) were analyzed. Based on the results, I have the following conclusions.

1. **Small gestures were better accepted than big gestures**. In our study, small gestures and big gestures were taught as two types of problem strategies to two groups of people. When people were encouraged to use the taught gestures to assist their problem solving, those who were taught small gestures were much more likely to keep using the taught gestures on most of the questions than those who were taught big gestures. People who were taught the big gestures spontaneously chose to use a lot of small gestures, which were never taught to them; some even did not try the big gestures at all. In addition, in the first half of the 15 questions, quite a number of people who were taught big gestures tried a mixture of gestures on several trials, but they eventually gave up the mixed gestures and switched to other types of gestures. Comparatively, only one person who were taught small gestures tried mixed gestures once in the first half. One possible explanation for this observation is that maybe those who were taught big gestures were not very satisfied with the taught strategy. They attempted a few other gestures at the beginning but gradually moved to gestures they felt confortable with.

2. **Small gestures could be more effective gestures than big gestures.** From the accuracy percentage (Figure 11), we can see that people who were taught small and used small gestures had the highest accuracy percentage of all groups. Small gestures also seemed to be related to a better performance in the primary task and a lighter cognitive load. However, the difference was

not statistically significant. The type of small gesture could possibly be a more effective gesture type than the type of big gesture.

**3. Using gestures influenced people's subjective judgment of cognitive load.** The results of the study showed that people's subjective judgment of cognitive load might not match their actual performance. For *small gesture users*, although small gestures seemed to have some positive effects on performance and cognitive load, they reported a much higher level of frustration than *big gesture users.* Moreover, when asked how successful they were in accomplishing what they were asked to do, *small gesture users'* rating of the performance was not significantly higher than that of *big gesture users'* (Figure 13).

*Big gesture users'* subjective judgments of cognitive load were also different from their actual performance in a number of ways. For example, although *big gesture users* reported a much lower level of frustration, their performance was not significantly better than other groups. They reported a significantly higher level of temporal demandingness, but they did not spend a longer time on gesturing or thinking than *small gestures users* did. In addition, *big gesture users'* rating of the difficulty to accomplish their performance level was also not significantly different from *small gesture users*. Taken together, these results imply that the type of big gesture seems to be related with a lower level of frustration, but its positive influence was not extended to other aspects. In other words, although using big gestures might look fun and make people feel less frustrated, it did not have significantly positive effects on the actual performance or the time spent on thinking.

**4. People became more skillful at gesturing.** The results also suggested that people spent less time thinking and less time gesturing as they moved forward in the tasks (Figure 6 and 7). Their gesture density decreased as well. However, their accuracy did not drop as the thinking time, gesturing time, and gesture density decreased (Figure 10). Moreover, the number of people who used no gestures increased significantly in the second half of the task (Figure 9). Considering the following three findings: (1) the type of *no gesture* increased during the second half, (2) people's gesture time, thinking time, and gesture density all decreased, and (3) the decrease was not at the cost of accuracy percentage, one possible interpretation is that people became more and more skillful at solving the problems.

## E. Limitations of Study 1

There are several limitations to this study. First, it would be better if a control group were added, in which no gestures were taught. It would help us to have a better idea regarding how the teaching of small and big gestures benefited learning, and if the benefits should be attributed to gesture type or the gesture teaching per se. Second, the 15 questions in the primary task were presented to participants in a fixed order. However, the accuracy percentage for the trials 2, 8, and 13 were much lower than other trials for no obvious reason. It would be better to present the trials in a random order in the future. Third, all of the instruction sessions on using small or big gestures were led by the same researcher. This might introduce some confounding factors as the research became more and more skillful at teaching gestures as the data collection process went on. It would be better if the instruction were presented in a video. Fourth, in this study, after a primary task question was presented, on some trials participants spontaneously used some time to rehearse and practice the secondary task so that they could better hold what they saw in memory. This period of time was counted into the thinking time in the analysis, because the cognitive boundary between stopping rehearsing patterns and starting thinking on the primary task was so unclear, such that a trained coder could not parse them apart. In the future, it would be good to add another fixation cross to indicate the beginning of the primary task, and instruct participants to move on when seeing the cross every time.

# Chapter IV Study 2

## A. Research questions

A review of literature and findings from Study 1 inspired me to investigate further into the co-thought gesture; how it can be taught, how it can be learned, whether teaching gesture will improve people's performance in solving spatial problems, and what would be the mechanism of gestures' effects on learning.

The leading research questions for this study include:

1. How will teaching big gestures, small gestures, and no gestures influence people's own gesture uses in a non-communicative setting?
2. How will using different types of gestures influence performances?
3. How will the use of gestures change over time when people solve problems?
4. How will using different types of gestures influence cognitive loads?

## B. Hypotheses

Based on previous research and findings of Study 1, I have the following hypotheses:

**Hypothesis 1:** Small gestures are better accepted than big gestures when taught as a problem solving strategy.

**Hypothesis 2:** Among the following four groups: (1) taught small gestures and use small gestures (TSUS), (2) taught small gestures and use big gestures (TSUB), (3) taught big gestures and use big gestures (TBUB), and (4) taught big gestures and use small gestures (TBUS), the first group will outperform all the other groups.

**Hypothesis 3:** When solving spatial problems, people's gesture time, thinking time, and gesture density will decrease; accuracy remains stable.

**Hypothesis 4:** When used as a strategy to solve spatial problems, small gestures lighten cognitive load more than big gestures.

## C. Method

### 1. Participants

One hundred graduate students from Teachers College of Columbia University participated in this study. They received either course credit or ten dollars for one hour of participation. Four cases were excluded from the sample, because of technical issues during the session. The final number of participants included in the analysis was 96, including 75 females and 21 males. There were 24 in the small gesture group, 50 in the big gesture group, and 22 in the no gesture group.

### 2. Design

The study employed a 1 × 3 factorial design. There were three treatment groups. Participants were randomly assigned to the three groups. They were taught small gestures, big gestures, and no gestures respectively. See Table 4.

Table 4

*Experiment Design and the Conditions*

| | **Gesture Type taught to participants** | | |
|---|---|---|---|
| | Small gesture | Big gesture | No gesture |
| **Groups** | Group 1<br>*n* =25 | Group 2<br>*n* =50 | Group 3<br>*n* =25 |

From what we observed from Study 1, we found that all participants rarely kept using the same gesture from the beginning to the end. Most of them used more than one type of gesture throughout the experiment. Moreover, from Study 1 we found that many people who were taught big gestures ended up not using them very frequently. Therefore, if we had assigned the same 25 participants in the big gesture group, we would expect that the number of people who actually used big gestures might be too low for us to make meaningful statistical inferences. In the hope of increasing the number of participants who ended up using big gestures during the independent problem solving session, we decided to double the sample size for the big gesture group from 25 to 50.

### 3. Procedure

Participants first signed a consent form and a form of participants' rights. After they read and signed the consent form, assented to participating in the study and being video recorded, they were asked whether they would like to give permission for us to show their videos in presentations of the research. Then they were asked to complete a paper version of the Mental Rotation Task (MRT) (Vandenberg & Kuse, 1978).

After that, participants were randomly divided into three groups: the big gesture group, the small gesture group, and the no gesture group. Each group of them was given ten minutes to watch an instructional video. The video content explained how to judge whether a chiral molecule is R- or S- configuration with big, small, or no gestures. Participants were allowed to stop the video at any time during the learning portion, and review the content as they felt necessary. Participants who were assigned into the small gesture group and the big gesture group were also required to move their hands or body parts when the video asked them to do so. Then participants were told: (1) They will stay in the room by himself or herself to solve fifteen problem sets on a laptop. (2) They will be encouraged to use their hands and any movements learned from the video because it will assist their problem solving. However, they are not required to. (3) Accuracy for all questions will be more important than the quickness of completion; the accuracy for molecule questions (rather than the block questions) is their first priority. (4) They can get started after the experimenter leaves the room, and take as much time as they want. (5) A fixation cross on the screen indicates they should move on to the next step. After every fixation cross appears for one second, the screen will advance automatically.

Participants were told that accuracy is the priority, so that they could take as much as time as needed for any problem. The quickness of responses is de-emphasized because we do not want participants' co-thought gestures to be suppressed due to time pressure. Participants were also told that the molecule questions were worth more points than the block questions, so that they would pay more attention to the molecule questions without being explicitly told that they were the primary task.

After ensuring participants understood the instructions, the experimenter left the room to allow the participants to work on the 15 trials on their own. Each trial included one primary task and one secondary task. Each primary task and secondary task began with a white fixation cross in the center of the screen. The fixation cross lasted for one second, and then the trial began automatically.

The procedure in Study 2 was the same as Study 1. Like Study 1, there were no practice trials before the experimental trials began. No feedback on the primary or secondary task was given during the time of experiment.

After completing the fifteen trials, participants completed a NASA Task Load Index (Hart & Staveland, 1988). After that, participants completed a near-transfer task. Finally, they took an exit survey.

Figure 17 shows the main procedure of the experiment.

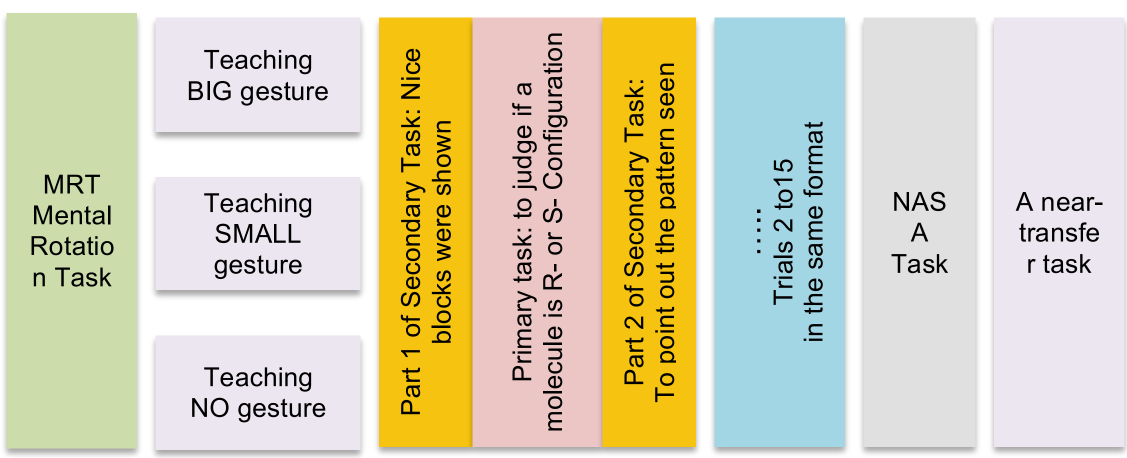



***Figure 17.*** Procedure of Study 2.

### 4. Content and Materials

In this study, the learning objective for all participants was to figure out whether a chiral molecule is a R-or S- configuration. In the video, they were taught to using small, big, or no gestures as a problem solving strategy.

In the transfer task, participants were asked to learn a related topic (the concept of *chirality*) from an illustration and two examples on their own. Then, they were shown six molecules diagrammed in dash-and-wedge notation, and were asked to judge whether each of them was chiral (non-superimposable with their mirror image) or achiral (superimposable with their mirror image) one by one. In the transfer task, participants were not taught any gestures, nor received any explicit instructions regarding whether they should gesture or not.

### 5. Measures

#### *a) An objective measure of cognitive load: Dual- Task Paradigm*

In this study, we used the dual-task paradigm to take the objective measure of cognitive load.

The primary task was a visual spatial problem. In the primary task, a molecule in a two-dimensional, wedge-and-dash notation was presented to participants on a laptop screen. Participants were asked to judge whether it was of R- or S-configuration. Strategy to solve the molecule problem using big, small, or no gestures had been already taught during the instruction session at the very beginning. Figure 18 shows a sample of the molecules in S-configuration.

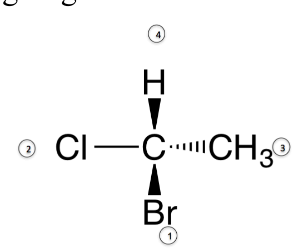



***Figure 18.*** A molecule presented to participants in dash and wedge notation.

Participant had to respond to each problem by hitting the R key for R-configuration and S key for S-configuration.

The secondary task was a block tapping task (see Figure 19). It is a visual spatial problem as well.

In the experiment, the dual-task was presented in two parts. Part 1 of the secondary task was presented before the primary task started. In this part, the nine blue static blocks were shown on a black background on the laptop screen, and then five of them turned yellow one after another. The duration of each yellow block shown was for one second.  Figure 19 shows a screening shot when one block was turning yellow.

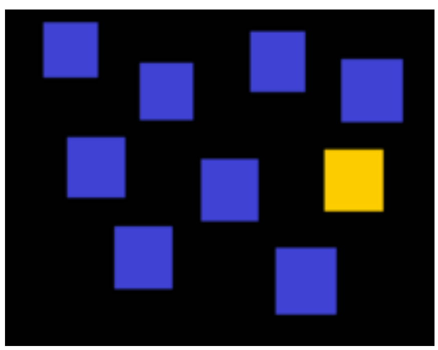


***Figure 19.*** Nine static blue boxes shown on a laptop screen. In the experiment, five of the blocks will turn yellow one after one, each for one second.

As required, participants had to hold what they saw in memory while working on the primary task. After the participants made a response to the primary task (typing "R" or "S"), part 2 of the secondary task began automatically. At this point, the original nine blue blocks were shown again, asking participants to recall what they had remembered and point the pattern out using a finger. After participants finished pointing the pattern out and hit the space key, the second trial of the dual task began.

This specific secondary task was adapted from the Corsi Block Tapping Task (Kessels, Van Zandvoort, Postma, Kappelle, & De Haan, 2000). It was designed as such out of the following considerations: (1) The number of blocks turning yellow was set to five. The reason is *the Corsi Span* (the average capacity of visual-spatial short-term working memory for normal human subject) is known to be five (Kessels et al., 2000), and we want this secondary task to be difficult enough to have participants' memories taxed. (2) According to the *Principle of Specificity*, this secondary task had to be visuo-spatial so that it can interfere participants' performance on the primary task, which is visuo-spatial as well.

According to the Dual-Task Paradigm, if different types of gestures induced different amounts of cognitive load in the primary task, we would expect that participants' performance in the secondary task will be influenced accordingly. To be specific, if a certain type of gesture lightens cognitive load when a participant is working on the molecule problem, the performance on its secondary task (the block question) will be better. On the contrary, if this type of gesture does not lighten cognitive load that much, the performance on the block question will be worse.

#### *b) A subjective measure of cognitive load: the NASA Task Load Index*

The Hart and Staveland (1988)'s NASA Task Load Index was used to assess participants' cognitive load as a subjective measure. Participants filled out this measure after they completed all tasks. They rated their cognitive loads from six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration.

### *c) Mental Rotation Task*

A paper version of the Mental Rotation Task (MRT) by Vandenburg & Kuse (1978) was administered to measure subjects' visual-spatial ability. In the test, each item consisted of one criterion figure, two alternatives, and two incorrect ones. The two correct alternatives were always identical to the criterion figure in structure but were shown in rotated position. The incorrect two were distractors. Subjects were asked to choose the correct alternatives. A sample item is shown in Figure 20.

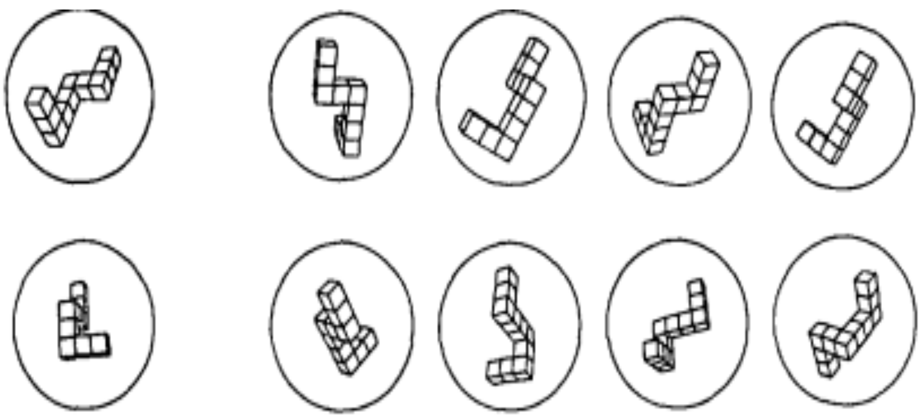


***Figure 20.*** A sample of the Mentation Rotation Task. The 1st and 4th alternative in the first item are correct; the distractors are mirror images of the criterion figure. In the second item, the 2nd and 3rd alternatives are correct; the distractors are rotated images of other criterion figures.

## 6. Apparatus

The primary task and secondary task were programmed in E-prime, and were presented to participants on a 15-inch Dell Laptop. Participants' gestures were captured by two cameras that were visible to them throughout the entire session. One camera was set at the left side or right side of the participant on a tripod. The second camera was set in front of the participant on a tripod too.

## 7. Gesture coding

To check inter rater reliability, two trained coders coded 25% of the videos on gesture type, *Kappa* = .89, $p < .001$, and for length of gesturing time, $r = .964$, $p < .01$. Disagreements were resolved by discussion.

All finger, hand, arm, and body movements by participants over the duration of the primary task trials, if not appearing to be for other purposes, such as scratching the head or pulling the hair, were coded as gestures. Beating gestures were not coded as gestures. *Gesture Time* was coded as from the beginning when a participant started making a movement to the moment when the movement was completed.

Participants' *gesture type* (*small gesture, big gesture, mixed gesture,* and *no gesture*) was also coded. *Small gesture* was defined as to use finger, hand, wrist, and part of the forearm as taught to represent molecule structure. When using small gestures, participants always used their palm to represent the central carbon in the target molecule. *Big gesture* was defined as to use arm, leg, and/or body movements to represent molecule structure. When using big gestures, participants use their torso to represent the center of the molecules and limbs to represent the molecule structure. The following were not counted as gestures nor into gesture time: (1) holding a gesture or holding body parts in a certain position for longer than 4 seconds, (2) the movement of dropping hands or arms after holding a gesture, and (3) moving hands and/or arms towards the key board to make a response to the primary or secondary task.

All gestures were coded from recorded videos in the ELAN software (European Distributed Corpora Project [EUDICO] Linguistic Annotator), developed by the Max Planck Institute for Psycholinguistics. In cases of disagreement, coders discussed to resolve the differences.

## D. Results

### 1. Teaching different types of gestures as a problem solving strategy

Each of 96 participants answered 15 molecule questions, which resulted in 1440 trials in the primary task. Out of the 1440 trials, the type of small gesture was the most frequently used, followed by no gesture, big gesture, and mixed gesture. Table 5 shows the gesture type, frequency, and their percentage.

Table 5

*Gesture Behavior for the Primary Task*

| Gesture Type | Frequency | Percentage of Total |
|---|---|---|
| Used Small Gesture | 615 | 42.7% |
| Used Big Gesture | 165 | 11.5% |
| Used Mixed Gesture | 93 | 6.5% |
| Used No Gesture | 567 | 39.4% |
| Total | 1440 | |

As expected, participants spontaneously chose to use gestures of preference on their own, although they were taught small, big, or no gestures respectively. As shown in Table 5, regardless of the type of gesture in instruction, small gesture was the most frequently used gesture type.

From Table 6, by percentage, we can see that for those who were taught small gestures, 70.56% chose to use small gestures, 7.09% big gestures, 1.94% mixed gestures, and 26.11% no gestures; for those who were taught big gestures, 43.20% chose to use small gestures, 20.13% big gestures, 11.07% mixed gestures, and 25.60% no gestures; for those who were taught no gestures, 11.21% chose to use small gestures, 2.73% big gestures, 0.91% mixed gestures, and 85.15% no gestures.

Mixed gesture was the least popular across all the three conditions. Figure 21 presents a bar graph of its percentage.

Table 6

*Gesture Behavior After Instruction*

| **Gesture Taught** | **Frequency and Percentage of Subtotal** |
|---|---|
| Taught Small | |
| Used Small Gesture | 254 (70.56%) |
| Used Big Gesture | 5 (7.09%) |
| Used Mixed Gesture | 7 (1.94%) |
| Used No Gesture | 94 (26.11%) |
| Taught Big | |
| Used Small Gesture | 324 (43.2%) |
| Used Big Gesture | 151 (20.13%) |
| Used Mixed Gesture | 83 (11.07%) |
| Used No Gesture | 192 (25.60%) |
| Taught No | |
| Used Small Gesture | 37 (11.21%) |
| Used Big Gesture | 9 (2.73%) |
| Used Mixed Gesture | 3 (0.91%) |
| Used No Gesture | 281 (85.15%) |
| Total | 1440 |

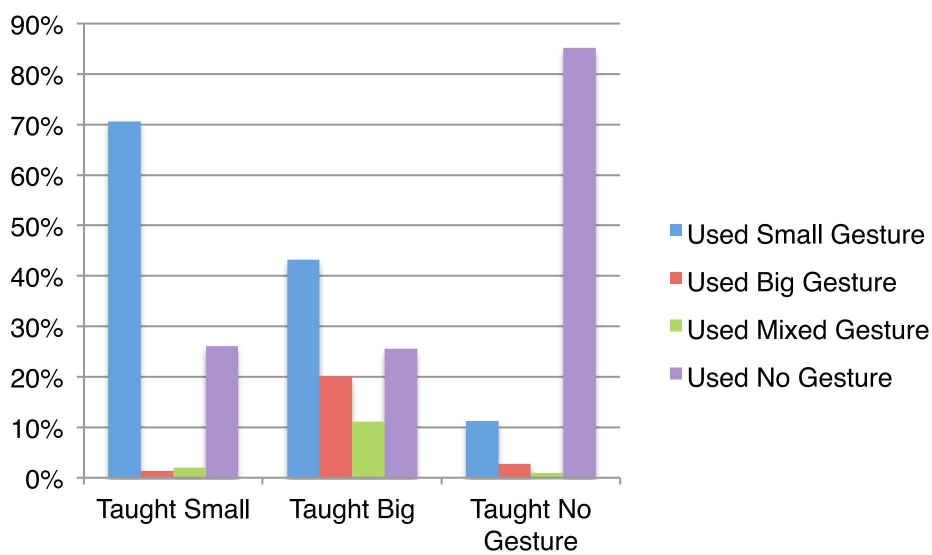



***Figure 21.*** The gesture type taught and the gesture type used by the three groups.

Based on these findings, I constructed a multinomial logistic regression model, using the type of gesture used (*small, big, mixed,* and *no gesture)* as the dependent variable, and the type of gesture taught (*small, big,* and *no gesture)* as the fixed effect. The model was statistically significant. It showed that the type of gesture taught significantly influenced the type of gesture used, $p < .001$. Table 7 shows a summary of the model.

Table 7

*Effects of Gesture Taught on Gesture Used*

| Source | *F* | *df1* | *df2* | *p-value* |
|---|---|---|---|---|
| Corrected | 61.416 | 6 | 1,431 | $p < .001$ |
| Gesture Taught | 61.416 | 6 | 1,431 | $p < .001$ |

If we see it from the perspective of time, from Figure 22, we can see that through question 1 to question 15, small gesture steadily remained to be the most frequently used gesture. The number of participants who used no gestures increased steadily and gradually. The number of participants who used big gestures remained around 10, except for a slight increase on the last few trials. The number of mixed gestures kept decreasing.

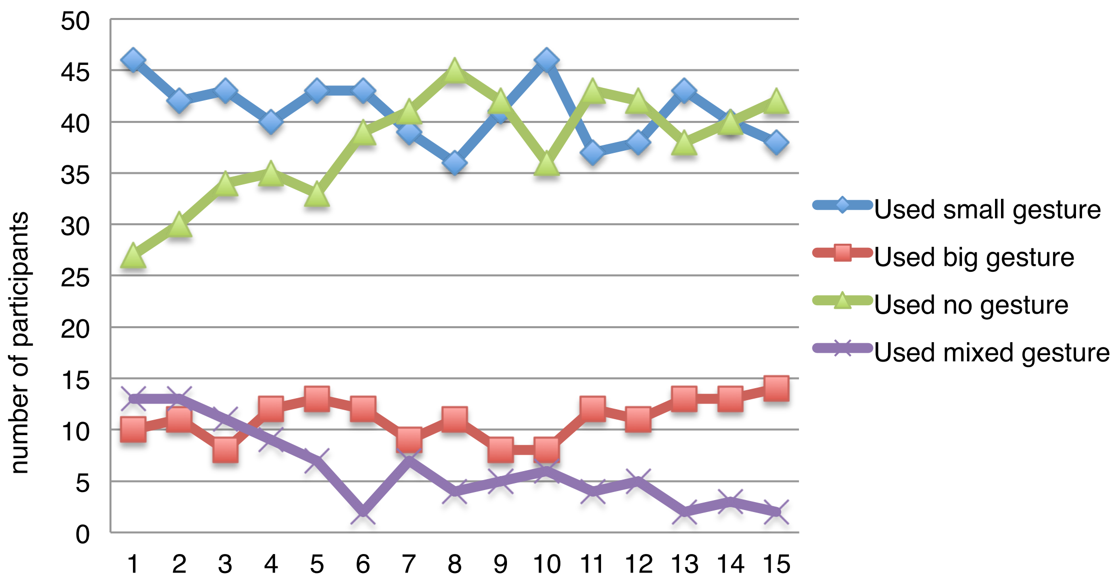



***Figure 22.*** The gesture type used on each trial.

### 2. Gesture and accuracy

Table 8 presents a summary of the gestures used, along with their frequency and accuracy percentage. Out of the 254 trials answered by small gestures, people got 371 right and 244 wrong, with a percentage of accuracy 60.33%. Out of the 165 trials answered by big gestures, people got 95 right and 70 wrong, with a percentage of accuracy 57.58%. Out of the 93 trials answered by mixed gestures, people got 48 right and 45 wrong, with a percentage of accuracy 51.56%. Out of the 567 trials answered by no gestures, people got 325 right and 242 wrong, with a percentage of accuracy 57.32%.

Table 8

*Gesture Behavior and Accuracy for the Primary Task*

| Gesture Type | Frequency | Percentage of Accuracy of Subtotal |
|---|---|---|
| Used Small Gesture | | |
| Right | 371 | 60.33% |
| Wrong | 244 | |
| Used Big Gesture | | |
| Right | 95 | 57.58% |
| Wrong | 70 | |
| Used Mixed Gesture | | |
| Right | 48 | 51.56% |
| Wrong | 45 | |
| Used No Gesture | | |
| Right | 325 | 57.32% |
| Wrong | 242 | |
| Total | 1440 | |

Based on Table 8, I continued my analysis from two perspectives. One is based on the type of gesture used by participants, and the other is based on the type of gesture we taught to them during instruction. The consideration is: in real life, instructors might be more interested in the gestures they chose to teach to students and would likely encourage students to use the taught gestures. On the other hand, instructors can not force students to use a certain type of gesture strategy, thus eventually the actual gestures students ended up using would affect their learning and performance as well.

From the perspective of gesture use, for each of the 15 trials, I calculated an accuracy percentage when the questions were answered by small, big, mixed, and no gestures respectively. Figure 23 shows that when the questions were answered by small gestures, the accuracy percentage is the highest out of all groups. This result is consistent with what we found from Study 1.

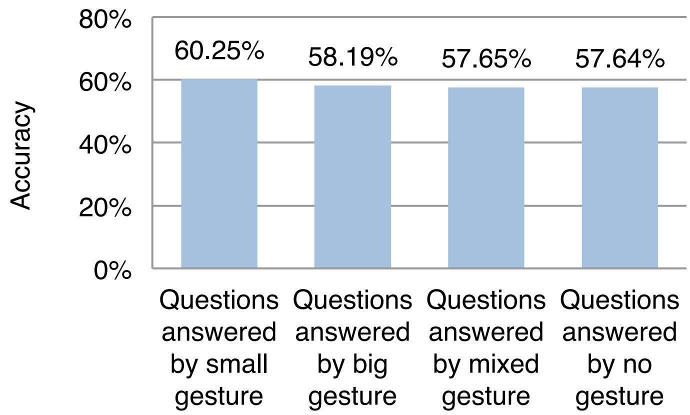



***Figure 23.*** Accuracy based on the type of gesture used.

To better understand how gesture use affected performances on the primary task, I constructed a binary logistic regression model using performance on the primary task (*right* or *wrong*) as the dependent variable, the level of visual-spatial ability (*high* or *low*), the type of gesture used (*small, big, mixed,* or *no gesture*) and the trial number as fixed effects, and subject ID as a random effect. The model showed that although the accuracy percentage of questions answered by small gesture was the highest among all groups (as shown in Figure 23), this difference did not reach the significance level. Table 9 presents a summary of the model.

Table 9

*Effects of Visual-Spatial Ability, Gesture Used, and Trial on Performance*

| Source | *F* | *df1* | *df2* | *p-value* |
|---|---|---|---|---|
| Corrected Model | 1.074 | 35 | 1,404 | .354 |
| Visual-Spatial Ability | 1.144 | 18 | 1,404 | .302 |
| Gesture used | 0.885 | 3 | 1,404 | .448 |
| # of the trial | 1.077 | 14 | 1,404 | .374 |

From the perspective of teaching, for each of the 15 trials, I calculated an accuracy percentage on the questions answered by subjects who were taught small, big, or no gestures. Figure 24 shows that the participants who were taught small gestures had the highest accuracy percentage, which was only slightly higher than other groups.

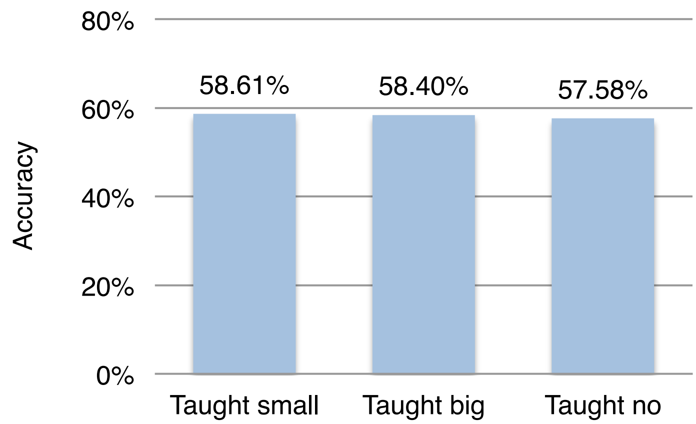



***Figure 24.*** Accuracy based on the type of gesture taught.

Again, I fitted a binary logistic regression model, using the actual performance on the primary task (*right* or *wrong*) as the dependent variable, the level of visual-spatial ability (*high* or *low*), the type of gesture taught (*small, big, mixed,* and *no gesture*), and the trial number as fixed effects, and subject ID as a random effect.

The analysis showed that the type of gesture taught did not have a significant effect on performance. Table 10 presents a summary of the model.

Table 10

*Effects of Visual-Spatial Ability, Gesture Taught, and Trial on Performance*

| Source | *F* | *df1* | *df2* | *p-value* |
|---|---|---|---|---|
| Corrected Model | 1.031 | 34 | 1,405 | .419 |
| Visual Spatial Ability | 1.182 | 18 | 1,405 | .267 |
| Gesture Taught | 0.017 | 2 | 1,405 | .983 |
| # of the trial | 1.035 | 14 | 1,405 | .415 |

Comparing the group who were taught and not taught gestures from the perspective of time, from Figure 25, we can see that in the first half of the trials, the accuracy percentage of the two groups were not very different from each other. In the second half onwards (trial 8 to trial 15), the gesture-taught group outperformed the gesture-not-taught group six out of eight times. However, this effect seemed to be subtle.

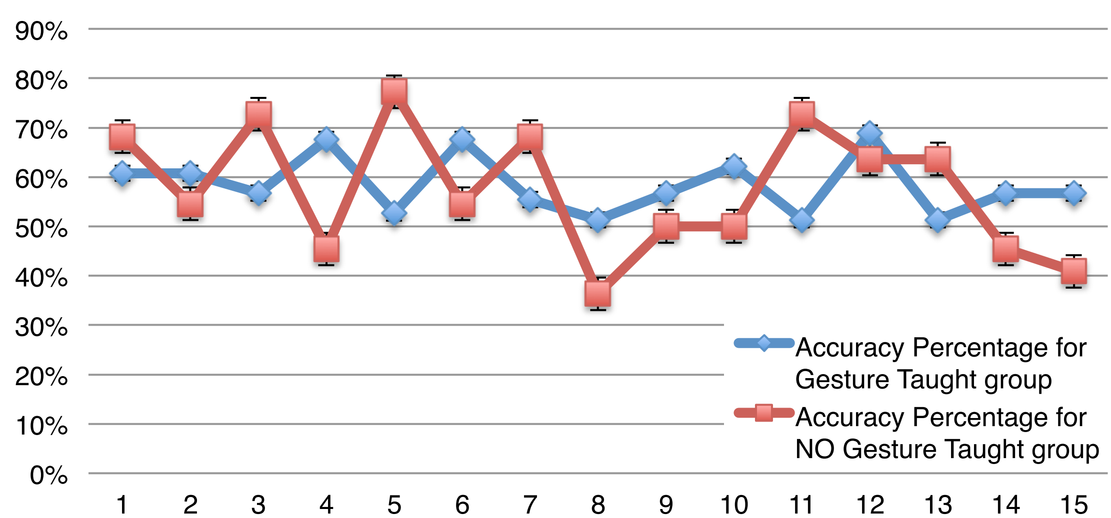



***Figure 25.*** Accuracy percentage for people who were taught and not taught gestures. Error bars represent standard error.

Notably, in the design of this study, the 15 primary questions presented should be of the same difficulty level, and they were presented to participants at a random order. In order to check whether the trial had an impact on the performance, I fitted two binary logistic regression models to test the effects of *trial number* and *question item number* on the actual performance respectively. The model showed that neither the trial number nor the question item had significant effect on performance, except that the $12^{th}$ trial, and the item 3, 11, 13 seemed to be easier than other questions at the $p$ <.05 level.

## 3. Gesture time, thinking time, and gesture density

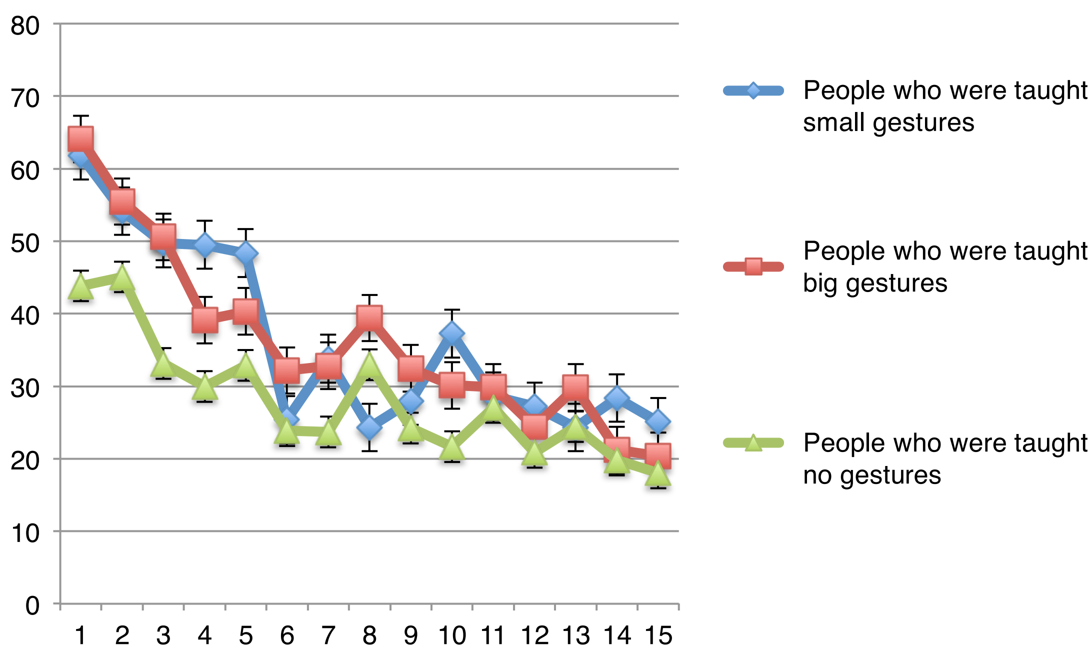



***Figure 26.*** Thinking time for each trial of the primary task by people who were taught small, big, and no gestures. Error bars represent standard error.

From Figure 26, we can see that the average time people spent on thinking decreased from question 1 to question 15, regardless whether they were taught small, big, or no gestures. In addition, the gesture-taught groups spent more time on thinking in the first half, but the time they spent became closer and closer to the no gesture group towards the end.

There were four sub-groups within the group who were taught small gestures: taught small used small gestures, taught small used big gestures, taught small used mixed gestures, and taught small used no gestures. Figure 27 shows that their thinking time decreased from the beginning to the end.

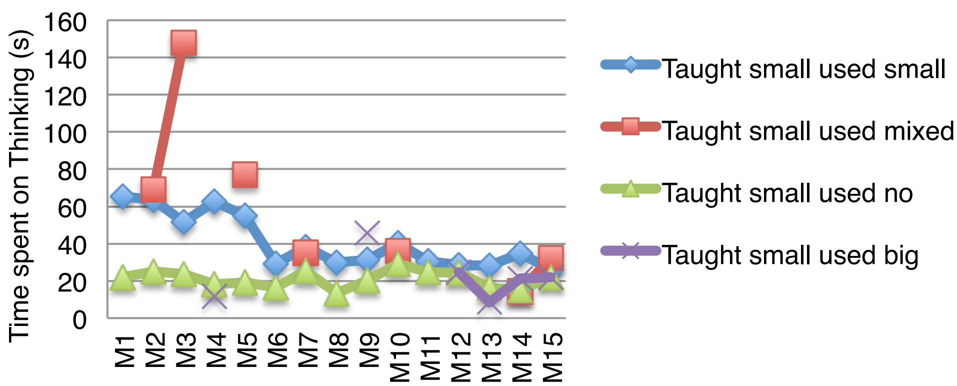



***Figure 27.*** Time spent on thinking by people who were taught small gestures.

There were also four sub-groups within the group who were taught big gestures: taught big used small gestures, taught big used big gestures, taught big used mixed gestures, and taught big used no gestures. Figure 28 shows that their thinking time also decreased from the beginning to the end.

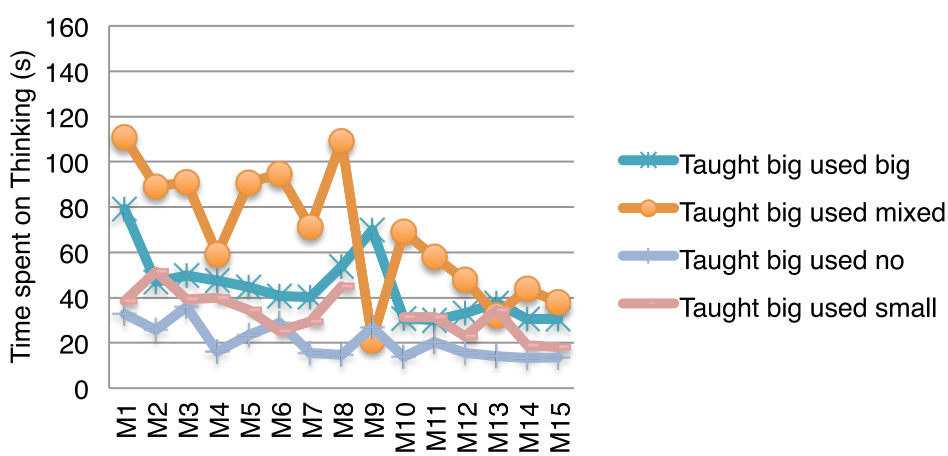



***Figure 28.*** Time spent on thinking by people who were taught big gestures.

People in the three groups not only spent less time thinking, they also spent less and less time gesturing (Figure 29). For the gesture-taught group, people who were taught small and big gestures showed very similar patterns in their gesturing time. The average time they spent on gesturing decreased gradually from the beginning to the end.

Average gesturing time for people who were taught small gestures decreased from 26.60 seconds to 8.71 seconds. Average gesturing time for people who were taught big gestures decreased from 24.84 seconds to 5.93 seconds. The decrease was especially more rapid in the first half than in the second half for both groups. For the gesture-not-taught group, participants did not move their hands and/or body parts for the majority of the session time. This group's average gesturing time remained below five seconds from the beginning to the end.

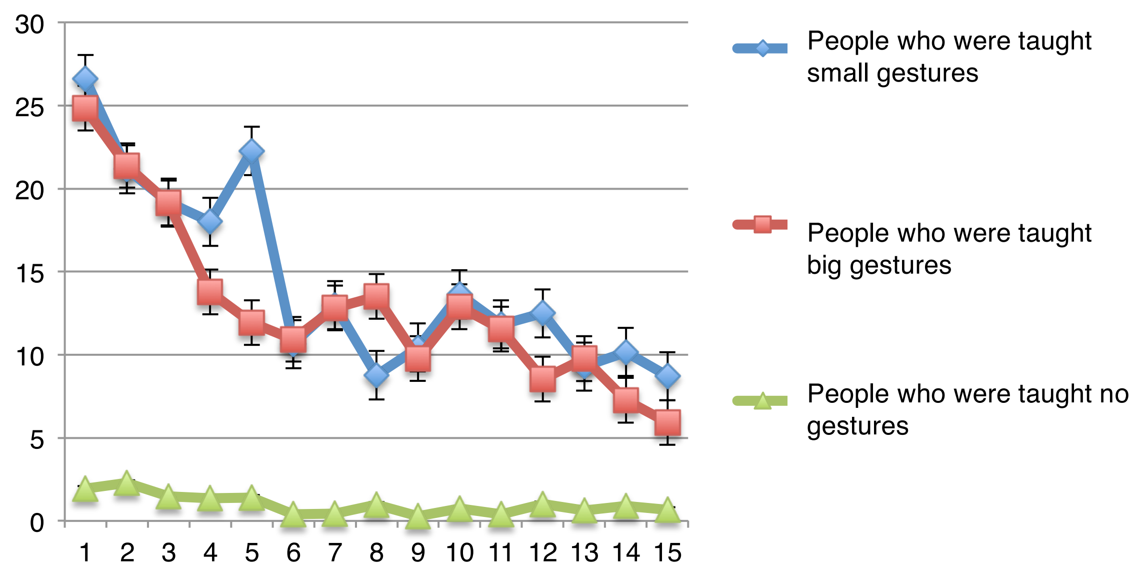



***Figure 29.*** Gesture time for each trial of the primary task by people who were taught small, big, and no gestures.

There were four sub-groups within the group who were taught small gestures: taught small used small gestures, taught small used big gestures, taught small used mixed gestures, and taught small used no gestures. Figure 30 shows that the gesture time decreased from the beginning to the end.

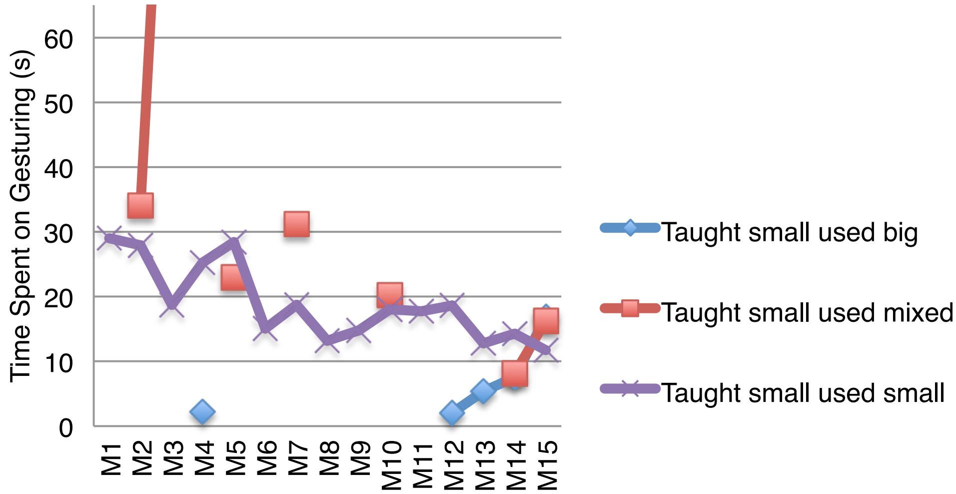



***Figure 30.*** Time spent on gesturing by people who were taught small gestures.

There were also four sub-groups within the group who were taught big gestures: taught big used small gestures, taught big used big gestures, taught big used mixed gestures, and taught big used no gestures. Figure 31 shows that the gesture time also decreased from the beginning to the end in a similar pattern.

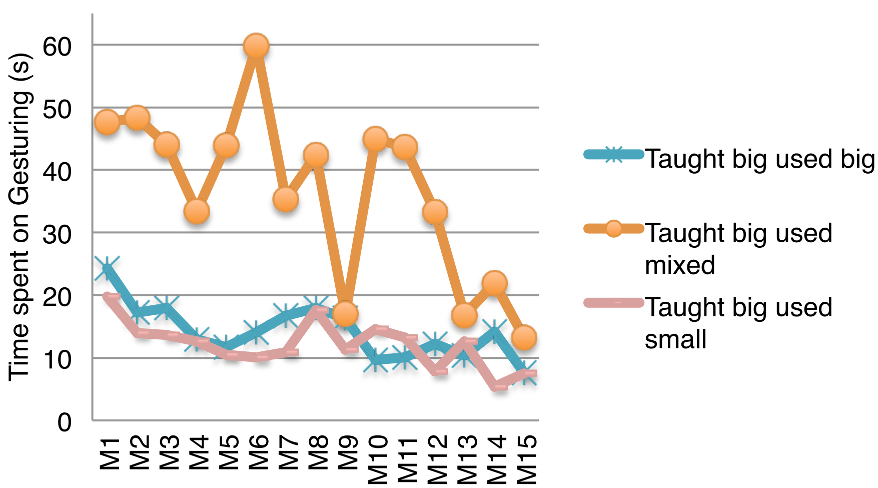



***Figure 31.*** Time spent on gesturing by people who were taught big gestures.

Although both thinking time and gesturing time decreased from question 1 to question 5, participants' gesturing density remained steady from the beginning to the end, between 31.23% and 40.46% (with a very slight decrease towards the end), as shown in Figure 32.

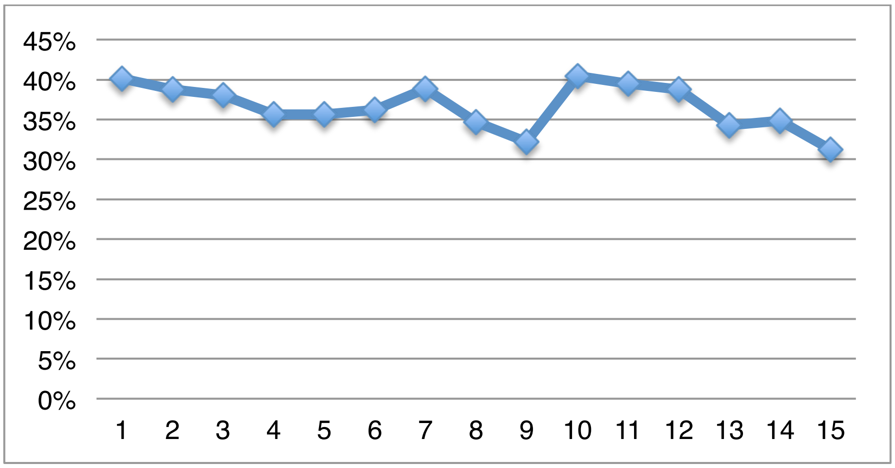



***Figure 32.*** Gesture density for each trial.

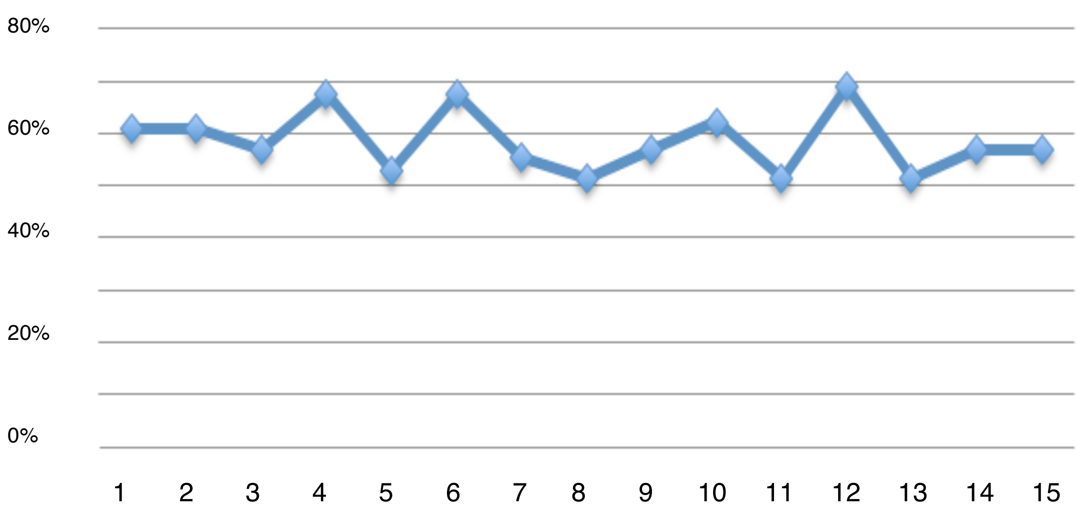



***Figure 33.*** Accuracy percentage for each primary task.

In addition, although participants' average thinking time, gesturing time, gesture density kept decreasing, people's accuracy percentage remained pretty steady (Figure 33).

### 4. Objective Measure of cognitive load

Participants' cognitive load was measured by their performance in the secondary task, which were the block tapping questions. The possible scores ranged from zero to five.

To understand how the type of gesture used affected performance scores on the secondary task, I constructed a binary logistic regression model, using performance on the secondary task (score range = 0 to 5) as the dependent variable, the type of gesture used (*small, big, mixed,* and *no gesture*), answer to molecule questions (*right* or *wrong*), and the interaction between these two variables as fixed effects, and subject ID as a random effect. The model showed that the type of gesture used did not have significant effect on the performance on the secondary questions (Table 11).

Table 11

*Effects of Gesture Used, Answer to Molecule Question, The Interaction, and the Actual Molecule Question ID on the Secondary Task*

| Source | *F* | *df1* | *df2* | *p-value* |
|---|---|---|---|---|
| Corrected Model | 0.743 | 21 | 1,414 | .790 |
| Gesture Used | 0.756 | 3 | 1,414 | .519 |
| Answer to MQ | 1.750 | 1 | 1,414 | .186 |
| Answer to MQ* Gesture Used | 1.259 | 3 | 1,414 | .287 |
| MoleculeQuestionID | 0.523 | 14 | 1,414 | .921 |

## 5. Subjective Measure of cognitive load

Our participants rarely consistently used one type of gesture throughout the 15 trials. As found from Study 1, they showed different gesture use patterns by switching between different types of gestures throughout the 15 trials. The same as Study 1, for a particular individual, if he chose to use a certain type of gesture as a problem solving strategy for more than two thirds of all the primary task questions, we coded this individual as the gesture user of this type. For example, for the 15 molecule questions, if a participant used small gestures on more than 10 questions, he was coded as *a small gesture user*. If a participant did not have a dominating gesture type, or used mixed gestures on one trial for two thirds of the trials, he was coded *a mixed gesture user*. Accordingly, all participants were divided into the following four types of gesture user: *the small gesture user, the big gesture user, the mixed gesture user,* and *the no gesture user.* Table 12 presents a summary of the four gesture user types.

Table 12

*Four Types of Gesture Users*

| Gesture User Type | Count | Percentage of the Total |
|---|---|---|
| Small Gesture User | 32 | 33.33% |
| Big Gesture User | 6 | 6.25% |
| Mixed Gesture User | 26 | 27.08% |
| No Gesture User | 32 | 33.33% |
| Total | 96 | |

Each individual completed a NASA Task Load Index, as a subjective measure of their cognitive load.

When asked "How mentally demanding was the task?," the *small, big, mixed,* and *no gesture users* reported no significant difference at the $p < .05$ level, $F(3,92) = 1.01$, $p = .394$ (Table 13).

Table 13

*Means and Standard Deviations of Gesture Users on Mental Demandingness*

| Gesture Type | *Mean (SD)* | *F* | *p-value* |
|---|---|---|---|
| Used Small Gesture | 16.91 (3.24) | 1.01 | .39 |
| Used Big Gesture | 19.17(1.33) | | |
| Used Mixed Gesture | 16.67(2.76) | | |
| Used No Gesture | 17.31(3.93) | | |

When asked "How physically demanding was the task?," *the small, big, mixed,* and *no gesture users* reported no significant difference at the $p$ <.05 level, $F$ (3,92) = .85, $p$ = .472 (Table 14).

Table 14

*Means and Standard Deviations of Gesture Users on Physical Demandingness*

| Gesture Type | *Mean (SD)* | *F* | *p-value* |
|---|---|---|---|
| Used Small Gesture | 7.95 (5.38) | .85 | .47 |
| Used Big Gesture | 8.50 (5.92) | | |
| Used Mixed Gesture | 7.58 (5.27) | | |
| Used No Gesture | 6.03 (5.59) | | |

When asked "How successful were you in accomplishing what you were asked to do? (performance)?, " *the small, big, mixed,* and *no gesture users* reported no significant difference at the $p$ <.05 level, $F$(3,92) = .36, $p$ = .785 (Table 15).

Table 15

*Means and Standard Deviations of Gesture Users on Performance*

| Gesture Type | *Mean (SD)* | *F* | *p-value* |
|---|---|---|---|
| Used Small Gesture | 13.39 (4.44) | .36 | .79 |
| Used Big Gesture | 15.00 (5.04) | | |
| Used Mixed Gesture | 13.04 (4.57) | | |
| Used No Gesture | 13.19 (5.59) | | |

When asked "How hard did you have to work to accomplish your level of performance? (effort)?," *the small, big, mixed,* and *no gesture users* reported no significant difference at the $p$ <.05 level, $F(3,92) = .83$, $p = .483$ (Table 16).

Table 16

*Means and Standard Deviations of Gesture Users on Effort*

| Gesture Type | *Mean (SD)* | *F* | *p-value* |
|---|---|---|---|
| Used Small Gesture | 15.50 (3.60) | .83 | .48 |
| Used Big Gesture | 17.00 (3.22) | | |
| Used Mixed Gesture | 14.39 (4.08) | | |
| Used No Gesture | 14.77 (4.89) | | |

When asked about the frustration level, "How insecure, discouraged, irritated, stressed, and annoyed were you?," *the small, big, mixed,* and *no gesture users* reported no significant difference at the $p$ <.05 level, $F(3,92) = 1.19$, $p = .319$ (Table 17).

Table 17

*Means and Standard Deviations of Gesture Users on Frustration*

| Gesture Type | *Mean (SD)* | *F* | *p-value* |
|---|---|---|---|
| Used Small Gesture | 11.83 (6.47) | 1.19 | .32 |
| Used Big Gesture | 15.50 (4.09) | | |
| Used Mixed Gesture | 13.96 (3.33) | | |
| Used No Gesture | 12.73 (5.99) | | |

When asked "How hurried or rushed was the pace of the task?," *the small, big, mixed,* and *no gesture users* reported no significant difference at the $p$ <.05 level, $F(3,92) = .48$, $p = .720$ (Table 18).

Table 18

*Means and Standard Deviations of Gesture Users on Temporal Demandingness*

| Gesture Type | *Mean (SD)* | *F* | *p-value* |
|---|---|---|---|
| Used Small Gesture | 11.73 (5.28) | .48 | .72 |
| Used Big Gesture | 11.83 (5.72) | | |
| Used Mixed Gesture | 10.13 (5.12) | | |
| Used No Gesture | 11.02 (5.94) | | |

## 6. Individual difference

As in Study 1, I used the median split at score = 9.00 to divide participants into low- and high-scoring group because of the distribution of scores in this sample (*Range* = 0-18, *median* = 9.00). See Figure 34.

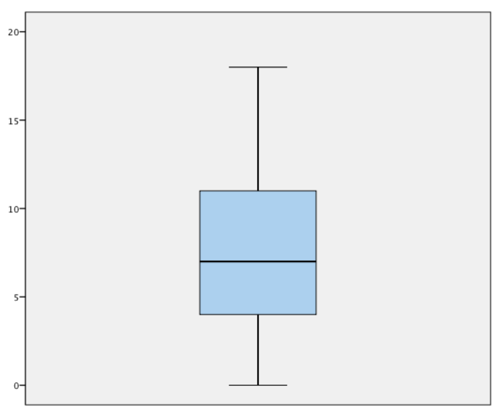



***Figure 34***. *High and low split of the participants.*

Table 19 shows a summary of the performance scores of *the small gesture users, the big gesture users, the no gesture users*, and *the mixed gesture users.*

Table 19

*Four Types of Gesture Users and Their Means Scores for the Primary Task*

| | Small Gesture | | Big Gesture | | No Gesture | | Mixed Gesture | |
|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | M | SD |
| Low Visual Spatial Ability | 7.32 | 2.14 | 8.33 | 3.06 | 8.28 | 2.20 | 9.00 | 2.00 |
| High Visual Spatial Ability | 10.77 | 3.06 | 9.67 | 3.06 | 8.93 | 2.08 | 9.85 | 3.08 |

Figure 35 shows that there was an interaction between gesture user type and the level of visual spatial ability on the primary task performance, $F$ (3,88)= 2.90, $p$ = .039. Specifically, high- and low- spatial ability participants performed similarly when they chose to use big, mixed, and no gestures as their primary strategy. However, low- spatial ability participants performed much worse than high-spatial participants, when using small gestures as the primary strategy.

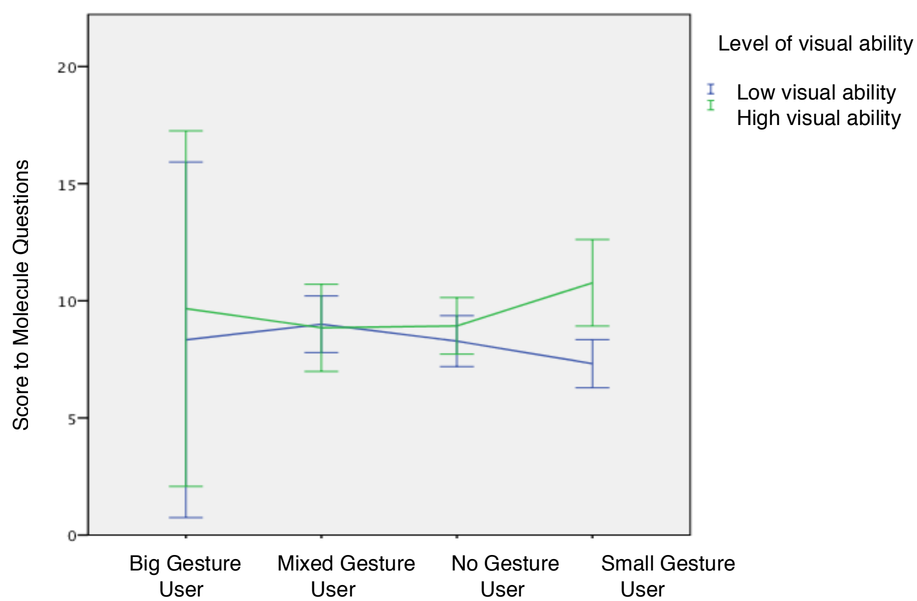



***Figure 35.*** Mean score of the primary task answered by high- and low- spatial ability participants. Error bars represent 95% confidence interval.

The result of the analysis of variance (*ANOVA*) is presented in Table 20.

Table 20

*Gesture User Type and Mean Scores for Primary Task -- ANOVA*

| | *df* | *MS* | *F* | *p* | *η2* |
|---|---|---|---|---|---|
| Gesture User Type | 3 | 1.109 | .183 | .907 | .006 |
| Visual Spatial Ability | 1 | 25.921 | 4.288 | **.041*** | .046 |
| Gesture User Type × Visual Spatial Ability | 3 | 17.555 | 2.904 | **.039*** | .090 |
| Error | 88 | 6.045 | | | |

* $p$ <.05
*R Squared* = .159 (*Adjusted R Squared* = .092)

## 7. Transfer tasks

After completing the primary and secondary task, participants were given a new topic to study, and then answered six transfer questions based on the topic. The result is shown in Table 21.

Eight participants (33.33%) who were taught small gestures for the primary task chose to gesture when learning a new topic on their own, and sixteen (66.67%) chose not. Nineteen participants (38%) who were taught big gestures for the primary task chose to gesture when learning a new topic on their own, and thirty-one (62%) chose not. Three participants (13.63%) who were taught no gestures for the primary task chose to gesture when learning a new topic on their own, and nineteen (86.36%) chose not. Notably, the type of gesture taught (*small, big,* or *no gesture*) did not significantly influence whether people gestured when learning a new topic, $\chi^2(2, N = 96) = 4.286$, $p = .117$. Overall, our participants tended to not gesture when asked to learn a new topic.

Sixty-six out of the 96 (68.75%) participants chose not to gesture when learning a new topic on their own.

Table 21

*Gesture Behavior When Learning a New Topic*

| Gesture Taught | Total | Used Gesture or not when learning a new topic (Transfer Task): Yes (% of subtotal) | No (% of subtotal) |
|---|---|---|---|
| Taught Small Gestures | 24 | 8 (33.33%) | 16 (66.67%) |
| Taught Big Gestures | 50 | 19 (38%) | 31 (62%) |
| Taught No Gestures | 22 | 3 (13.63%) | 19 (86.36%) |

When answering the transfer questions, participants tended to not gesture either. The number of participants who did not gesture slightly increased towards the end from 75 to 86. Those who did gesture slightly decreased from 21 to 10 (Figure 36).

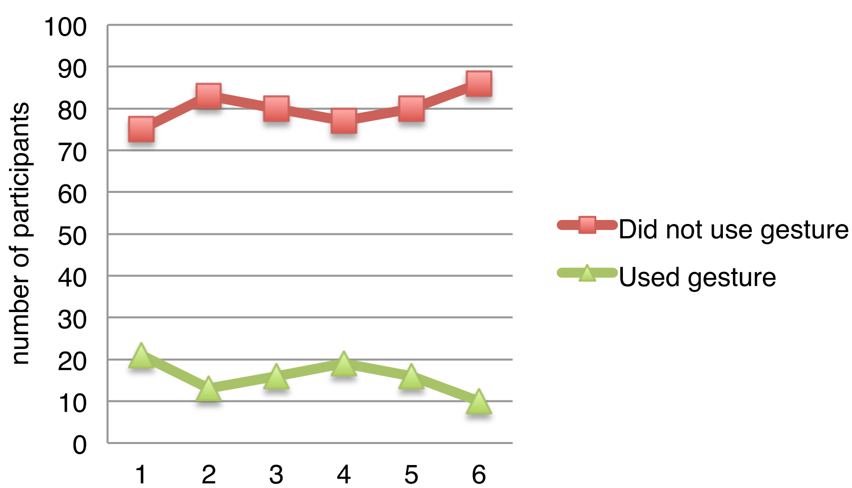



***Figure 36.*** The number of participants who gestured and did not gesture in each transfer question.

When answering the six transfer questions, the overall performances by people who were taught small, big, and no gestures were very similar (Figure 37). A binary logistic regression model using performance on the transfer question (*right* or *wrong*) as the dependent variable and the type of gesture taught (*small, big,* or *no gesture*) as the fixed effect showed that the type of gesture taught in the training beforehand did not have a significant influence on performance on the transfer task, $p$ =. 417.

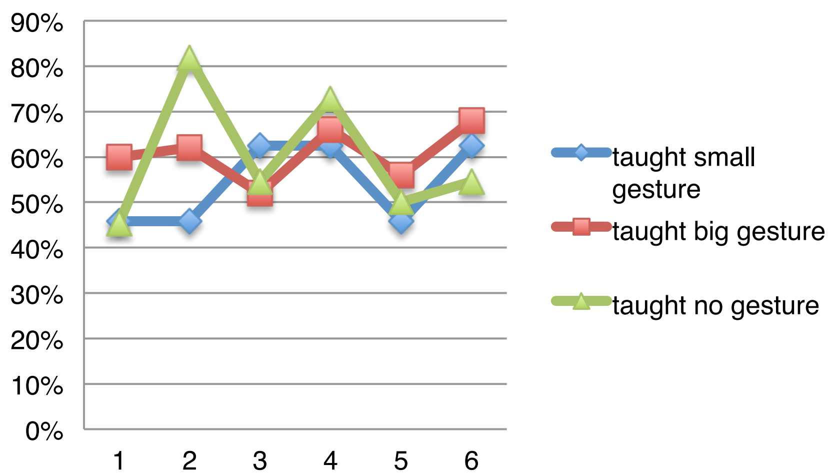



***Figure 37.*** Accuracy percentage in transfer questions by the gesture type taught before.

When answering the transfer questions, people who gestured and not gestured did not show a significant difference (Figure 38). A binary logistic regression model using performance on the transfer question as the dependent variable, and gesture used as a fixed effect, showed that gesture used or not on the transfer questions did not have a significant influence on transfer task performance, $p$ =. 836.

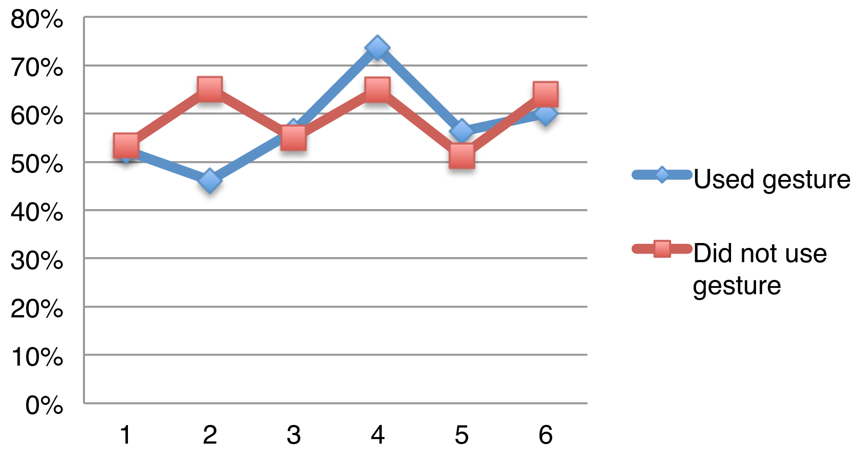



***Figure 38.*** Accuracy percentage of transfer questions by people who used and did not use gesture.

# Chapter V Discussion

## A. General discussion

In this dissertation, I investigated how teaching different types of gestures and body movements influenced people's thinking, learning, and problem solving. In Study 1, the instruction was conducted in person. In Study 2, the instruction was conducted through a recorded video. In both studies, participants completed 15 sets of questions after the instruction, and then rated their cognitive load. In Study 2, participants completed an additional near-transfer task that consisted of six questions.  Consolidating findings from Study 1 and Study 2, I came to the following conclusions.

The findings of the two studies are consistent with earlier studies on the teaching of gestures. First, gesture was teachable. In our study, we purposely taught people different types of gestures and body movements as problem solving strategies. Every participant chose to use the taught gestures to a lesser or greater extent. None of them tried to solve the problem without gesturing at all. Second, I found that when people were taught how to use gesture to solve problems, and were encouraged to gesture, they were receptive and subsequently used the gestures and body movements. This is consistent with previous findings that modeling gesture to people made them gesture (Cook et al., 2008), and encouraging or forcing them to gesture would bring out effective gesture as well (Broaders et al., 2007). Third, although our instructions were very brief, after our instruction, participants were found to use the taught gesture extensively. This is also consistent with previous findings that even brief sessions of training in gesture can yield beneficial results (Ehrlich et al., 2006).

When it comes to the specific gesture types, the conclusions are as follows.

1. **Small gestures were more readily accepted than big gestures**. In both studies, I found that small gestures were always better accepted than big gestures, no matter whether they were taught in person or through a video. My results also showed that even when people were taught big gestures as a problem-solving strategy and were encouraged to use them, the most preferred gesture type was still the small gesture. The type of gesture taught had a significant effect on the actual gesture use.

2. **Small gestures might have a more positive effect on thinking and performance than big gestures.** In both studies, questions answered by the small gestures had a slightly higher accuracy percentage. However, the effect on accuracy and on the cognitive load did not reach statistical significance. Further studies would be needed to further investigate this effect to have a better understanding of it.

3. **Using different types of gestures may or may not influence people's subjective judgment of their cognitive load differently.** Study 1 showed that using different types of gestures affected users' subjective judgment of their cognitive load differently in multiple respects. However, in Study 2, none of the subjective cognitive load judgments showed any significant difference. A possible explanation might be related to the mode of instruction, because the impact of a real person seems to be bigger than that of a recorded video. Specifically, the switch from the in-person instruction to the digital presentation might have made gesture's effects much less impactful, so that a lot of the differences reported in Study 1 were no longer significant in Study 2. In addition, it's possible that a lot of the in-person "fun" with whole body movement was dampened with the switch to digital instruction, so the frustration level of the big gesture users was also not significantly lowered in Study 2.

4. **People became more skillful at gesturing.** The results of both studies suggested that people in general spent less time thinking and less time gesturing as they moved forward with the trials. Their gesture density decreased slightly as well. However, their accuracy percentage did not drop as the thinking time, gesturing time, and gesture density decreased. Moreover, the number of people who started to use no gestures increased significantly in the second half of the trials.

5. **People tried mixed gestures at the beginning and dropped them gradually.** In both studies, we observed that a number of participants tried to use a mixture of small gestures and big gestures at the beginning, but the number dropped slowly when people moved on towards the end. Mixed gesture was not only the least favored gesture type across all conditions, and its accuracy was also the lowest. Therefore, trying a mixture of the small and big gestures did not seem to be a strategy that boosted people's problem-solving performance. On the contrary, it seemed that across conditions, people attempted this type of mixture at the very beginning when they were a little confused and were exploring different strategies, but gradually moved on to a gesture type they felt comfortable with or even transitioned to using no gestures as they became more competent at problem-solving.

6. **People who were taught to use gestures did not differ significantly from those who were not taught to use gestures at the beginning, but the gesture group seemed to perform a little better later.** In Study 2, we added a no gesture group. In this group, gesture was neither taught nor encouraged. We found that the overall performance of this group was not significantly different from the other two gesture groups. This finding seemed different from the literature in which gesturing groups usually outperformed non-gesturing groups (Chu & Kita, 2011; Goldin-Meadow et al., 2001; Jamalian, Giardino, & Tversky, 2013a). In my studies, I also found that

although there were no overall significant differences between the gesture groups and the no gesture group, the gesture groups began to catch up and outperformed the gesture-not-taught group in their accuracy during the second half of the trials. A possible explanation for my findings might be that participants in the gesture groups responded to the gesture instruction differently. For some people, it took some time and effort to learn to use the taught gesture skillfully and internalize them as their own strategy to increase their performance. Others might have learned the strategy much more quickly. They began by using the taught gesture as the main strategy, but as they became more and more skillful at using it, they gradually dropped the gestures while maintaining or even increasing their performance. The effect of these two groups might cancel each other as time moved forward. In the literature, when comparing the performance of a gesture group and a no-gesture group, the gestures people utilized were mostly spontaneous gestures that were produced by people naturally (Chu & Kita, 2011; Goldin-Meadow et al., 2001; Jamalian et al., 2013a).When the gestures were taught as a strategy and were required to be used (Cook et al., 2008; Jamalian, 2014), the literature rarely talked about the complexity of using the taught gesture as a strategy, or whether participants needed some time and effort to practice the taught gestures so that they could used them as a strategy effectively and skillfully. In our study, some participants confirmed to experimenters that they would prefer more time or practice for the taught gestures, so that they could use them on more questions comfortably, or use them more effectively.

### B. Theoretical contributions

First, this study contributes to embodied cognition theories. The findings from the two studies encourage us to interpret the embodied cognition theories from a new perspective. According to embodied cognition theory, action and gesture (as simulated action) can promote thinking by engaging the human body. However, not many studies in the field have provided direct evidence regarding whether the degree of embodiment is as important as well. The results of this dissertation suggested that it might make sense to look at the concept of embodiment as a spectrum, from a low degree of embodiment to a high degree of embodiment, rather than as a dichotomous methodology (e.g., this experience is embodied or not). My findings directly suggest that big gestures and body movements, although they could be seen as a higher degree of body engagement, might not necessarily lead to better performance compared to small gestures. Both teaching and use of the big gestures and body movements seemed to be less effective as the small gesture. It may further imply that teaching the right type of embodiment may be more important than eliciting a higher level or degree of an embodied experience.

This study also contributes to the gesture literature. In the literature, although it is widely agreed that gesture plays a role in promoting thinking and learning, the mechanisms that underlie this process are not yet fully understood (Goldin-Meadow, 2010). My studies shed light on these mechanisms by showing that different types of gestures might possibly involve different amounts of cognitive load, and that using different types of gestures (e.g., big or small) might benefit learners of different cognitive abilities (high and low visual spatial abilities) to different degrees. Future studies should further investigate this mechanism and others to better understand gesture's cognitive benefits in teaching and learning. Finally, regarding the manipulation of gesture, my studies also provide evidence that small gestures might be more easily accepted than big gestures, so it would be beneficial to study how people gradually pick up and drop gestures along the course of learning.

## C. Practical implications

Based on my findings, I have identified the following practical implications for instruction and instructional design. First, although in general gestures and body movements were found to promote thinking and learning, the literature has already shown that not every kind of gesture benefits learning. In line with this, my studies suggest that different types of gestures and body movements may promote thinking and learning to different degrees. This means that when teaching gesture as a problem solving strategy, instructors should explore, compare, and even experiment to find the best suitable gesture type(s) based on the learning materials and learners' ability, in order to get the greatest benefits from the gestures.

Second, instructors should be cautious about the cost of cognitive load that comes with gesturing, although the cognitive benefits of  gesture are relatively more well-known. In our study, I found that small gestures are not only more easily accepted, they might also possibly be related to higher accuracy and lower cognitive load in some cases. Therefore, it would be beneficial to try different gesture types when designing instruction. Ideally, a gesture type that both promotes learning and lightens cognitive load would be selected. Instructors should take these two factors into consideration when designing instruction.

Third, when instructors teach gesture to students as a learning strategy or problem solving strategy, if the learning content is complicated (e.g., the learning content in my studies is about the concept of chirality and molecular structure), or if the gesture strategy itself is not very straight forward, time and practice for the learner should be taken into consideration. Participants may benefit from being given experience practicing the taught gestures before they could use them comfortably and skillfully.

Fourth, for learning designers and education software designers, an instructional experience designed to engage full body movement (e.g., an educational game on Xbox) might seem more attractive and motivational than one on a small-scale interface that only engages bodily movement partially. However, the results of my studies suggest that larger-scale movement might not necessarily bring better learning results, although a subjective measure indicated that people may possibly feel less frustrated with such gestures. In other words, whole body movement in an educational setting could be perceived as cool, playful, fun, or less frustrating, but it may not be directly related to increased performance. On the contrary, for some learning content, small gestures might possibly elicit frustration (although we do not know if it is only temporarily, or whether it would be mitigated if any feedback or practice time were given), but they may possibly benefit thinking and performance more. An instructor or a designer would need to take all of these factors into consideration comprehensively, when making decisions on how to design the optimal instructional gesture. An ideal instructional gesture should bring out a positive learning experience and learning results.

## D. Limitations

There are four limitations to the present studies. First, the delivery methods of the instruction for Study 1 and Study 2 were different. In Study 1, the instruction was conducted in person, but in Study 2 the instruction was conducted in a pre-recorded video. Although in the second study, the video instruction made the instruction consistent for all participants and conditions, this difference made the results of the two studies less comparable.

Second, for both studies, participants' learning task performance was solely based on two degrees of measurement, namely, right or wrong (R- or S-configuration for the primary task, and chiral or achiral for the transfer task). While the analysis already showed some differences across groups, future studies should consider using quantitative measures to gain more information.

Third, participants' attitude towards gesture and their tendency to use gesture during learning and problem solving were not measured or controlled prior to the study session. Therefore, there is no way to conclude whether attitude or tendency to gesture would cause any systematic differences.

Fourth, the studies were conducted with a particular population (college students) on a specific domain (college level stereochemistry). Based on these two limitations, the findings cannot be generalized to other populations or other domains. Further research might consider teaching different types of gestures to learners of other ages and in other domains.

## E. Directions for future research

There are several research directions that could extend the theoretical and practical implications of the grounded embodiment theories, specifically in the applications of gesture. Although embodied cognition theories suggest that embodiment promotes learning, I believe cognition might not be simply seen as "embodied" or "not embodied". In my studies, I found that the big gestures and body movements did not benefit thinking and learning as much as the small gestures did. Further research could provide more insights on whether and how different degrees of embodiment would influence thinking and learning.

This study also inspired us to better harness the cognitive benefits of gesture in instruction, and put them to better use. First, in my studies, following instruction, gesture was only encouraged but not required. To draw more powerful conclusions, it would be meaningful for future studies to further investigate how teaching gestures would benefit people, if gesturing were explicitly required (rather than just encouraged). Second, in our study, it was found that people's gesture time, thinking time, and gesture density all decreased as they moved forward with trials without sacrificing accuracy. It would be interesting to further study the reason for this decrease. It would provide a more complete understanding on whether the decrease was due to people becoming so skillful that they did not need to gesture any more, or they just started to internalize the acquired gestures and replaced the full-scale gestures with "imagined gesture" or "simulated gesture" that were invisible to us. Third, in both studies during the second half of the trials, some people started to use "minimal gestures". These gestures looked like very minimal movements of fingers or hands, but seemed to be different from random hand movements. In my data analysis, we categorized this type of gesture into small gesture, but its gesture space was so much smaller that the gesture mechanism was suspected to be different from what we referred to as small gesture. It would be interesting for future studies to investigate the nature of these "mini gestures", to examine whether they were just a transitional stage to no gesture (which reflected that people were internalizing the problem solving strategy) or a totally different type of gesture, which was partially a physical gesture and partially an imagined gesture. Fourth, our study showed that for some but not all participants, time and practice may be needed to master the taught or self-generated gesture. Considering that literature on gesture manipulation has been limited, a worthwhile future research direction could be to look into how to better teach gesture as a problem solving strategy, especially when the learning content is complicated. Fifth, in my study, I only investigated the benefits of gestures in the learning of stereochemistry. Future research should consider testing gesture's effects across other content fields to gain a more comprehensive understanding. Finally, considering individual difference, people of high and low visual ability might benefit from different types of gestures. Future studies could consider exploring possible ways to better help learners of lower visual spatial ability.

# Reference

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science, 10*, 327-333.

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

Alibali, M. W., & Kita, S. (2010). Gesture highlights perceptually present information for speakers. *Gesture, 10*, 3-28.

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.

Alibali, M. W., Spencer, R. C., Knox, L., & Kita, S. (2011). Spontaneous gestures influence strategy choices in problem solving. *Psychological Science, 22*, 1138-1144.

Bai, X., Black, J., & Vitale, J. (2006). *REAL: an agent-based learning environment.* Paper presented at the Agent-Based Systems for Human Learning Conference, Hakodate, Japan.

Bara, F., Gentaz, E., Colé, P., & Sprenger-Charolles, L. (2004). The visuo-haptic and haptic exploration of letters increases the kindergarten-children's understanding of the alphabetic principle. *Cognitive development, 19*, 433-449.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and brain sciences, 22*, 637-660.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol., 59*, 617-645.

Barsalou, L. W., Niedenthal, P. M., Barbey, A. K., & Ruppert, J. A. (2003). Social embodiment. *Psychology of learning and motivation, 43*, 43-92.

Bavelas, J., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language, 58*, 495-520.

Beaudoin-Ryan, L., & Goldin-Meadow, S. (2014). Teaching moral reasoning through gesture. *Developmental science, 17*, 984-990.

Beilock, S. L., & Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *Psychological Science, 21*, 1605-1610.

Bertenthal, B. I., Campos, J. J., & Kermoian, R. (1994). An epigenetic perspective on the development of self-produced locomotion and its consequences. *Current Directions in Psychological Science, 3*, 140-145.

Black, J. B. (2007). Imaginary worlds. *Memory and mind: A festschrift for Gordon H. Bower*, 195-208.

Black, J. B., & Bower, G. H. (1980). Story understanding as problem-solving. *Poetics, 9*, 223-250.

Black, J. B., Segal, A., Vitale, J., & Fadjo, C. (Eds.). (2012). *Embodied cognition and learning environment design*. New York: Routledge.

Boncoddo, R., Dixon, J. A., & Kelley, E. (2010). The emergence of a novel representation from action: evidence from preschoolers. *Developmental science, 13*, 370-377.

Bosbach, S., Cole, J., Prinz, W., & Knoblich, G. (2005). Inferring another's expectation from action: the role of peripheral sensation. *Nature neuroscience, 8*, 1295-1297.

Britton, B. K., & Tesser, A. (1982). Effects of prior knowledge on use of cognitive capacity in three complex cognitive tasks. *Journal of Verbal Learning and Verbal Behavior, 21*, 421-436.

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: General, 136*, 539.

Brünken, R., Plass, J. L., & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. *Instructional Science, 32*, 115-132.

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental psychology, 49*, 109.

Bucci, W., & Freedman, N. (1978). Language and hand: The dimension of referential competence1. *Journal of Personality, 46*, 594-622.

Campos, J. J., Bertenthal, B. I., & Kermoian, R. (1992). Early experience and emotional development: The emergence of wariness of heights. *Psychological Science, 3*, 61-64.

Carnap, R. (1956). *Meaning and necessity* (2nd ed.). Chicago: University of Chicago Press.

Cattaneo, L., & Rizzolatti, G. (2009). The mirror neuron system. *Archives of neurology, 66*, 557-560.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive psychology, 4*, 55-81.

Chiel, H. J., & Beer, R. D. (1997). The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in neurosciences, 20*, 553-557.

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General, 137*, 706.

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

Chu, M., & Kita, S. (2015). Co-thought and co-speech gestures are generated by the same action generation process.

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

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT press.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition, 106*, 1047-1058.

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and cognitive processes, 27*, 594-610.

De Ruiter, J. P. (1998). *Gesture and speech production.* Radboud University Nijmegen Nijmegen.

Decety, J., & Grèzes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain research, 1079*, 4-14.

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological review, 3*, 357.

Dewey, J. (1938). *Experience and education*. New York: Touchstone.

Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology, 42*, 1259.

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

Fadjo, C. L. (2012). *Developing Computational Thinking through Grounded Embodied Cognition.* (Ph. D.), Columbia University, New York, NY. Retrieved from http://hdl.handle.net/10022/AC:P:13168 Available from Columbia University Academic Commons

Feyereisen, P., & De Lannoy, J.-D. (1991). *Gestures and speech: Psychological investigations*: Cambridge University Press.

Flood, V. J., Amar, F. o. G., Nemirovsky, R., Harrer, B. W., Bruce, M. R., & Wittmann, M. C. (2014). Paying attention to gesture when students talk chemistry: Interactional resources for responsive teaching. *Journal of Chemical Education, 92*, 11-22.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119*, 63.

Fry, D. B. (1966). *The Genesis of Language: A Psycholinguistic Approach* (F. Smith & G. A. Miller Eds. f ed.).

Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer-Verlag.

Fuster, J. M. (2003). *Cortex and mind: Unifying cognition*: Oxford university press.

Garber, P., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive science, 26*, 817-831.

Geiser, C., Lehmann, W., & Eid, M. (2006). Separating" rotators" from" nonrotators" in the mental rotations test: A multigroup latent class analysis. *Multivariate Behavioral Research, 41*, 261-293.

Gelman, R. (1980). What young children know about numbers. *Educational Psychologist, 15*, 54-68.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gentner, D., & Stevens, A. L. (Eds.). (1983). *Mental models*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gerlach, C., Law, I., & Paulson, O. B. (2002). When action turns into words. Activation of motor-based knowledge during categorization of manipulable objects. *Journal of cognitive neuroscience, 14*, 1230-1239.

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*: Cambridge University Press.

Gibbs, R. W. (2006). *Embodiment and cognitive science*. New York: Cambridge University Press.

Gibson, J. J. (1979a). *The ecological approach to visual perception: classic edition*. New York: Houghton Mifflin.

Gibson, J. J. (1979b). *The ecological approach to visual perception.* Hillsdale, NJ: Erlbaum.

Glenberg, A. M. (2008). Toward the integration of bodily states, language, and action. In G. R. Semin & E. R. Smith (Eds.), *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches.* (pp. 45). New York, NY: Cambridge University Press.

Glenberg, A. M., Brown, M., & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology, 32*, 389-399.

Glenberg, A. M., Goldberg, A. B., & Zhu, X. (2011). Improving early reading comprehension using embodied CAI. *Instructional Science, 39*, 27-39.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and Imagined Activity Can Enhance Young Children's Reading Comprehension. *Journal of Educational Psychology, 96*, 424.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic bulletin & review, 9*, 558-565.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse processes, 28*, 1-26.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of memory and language, 43*, 379-401.

Göksun, T., Goldin-Meadow, S., Newcombe, N., & Shipley, T. (2013). Individual differences in mental rotation: what does gesture tell us? *Cognitive processing, 14*, 153-162.

Goldin-Meadow, S. (2010). When gesture does and does not promote learning. *Language and cognition, 2*, 1-19.

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

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12*, 516-522.

Goldin-Meadow, S., & Singer, M. A. (2003). From children's hands to adults' ears: gesture's role in the learning process. *Developmental Psychology, 39*, 509.

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction, 9*, 201-219.

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford: Oxford University Press.

González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., & Ávila, C. (2006). Reading cinnamon activates olfactory brain regions. *Neuroimage, 32*, 906-912.

Graham, T. A. (1999). The role of gesture in children's learning to count. *Journal of Experimental Child Psychology, 74*, 333-355.

Hachey, A., Tsuei, L., & Black, J. (2001). *Fostering mental-model thinking during design.* Paper presented at the Proceedings of EdMedia, Norfolk, VA.

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia, 42*, 1281-1292.

Han, I., & Black, J. B. (2011). Incorporating haptic feedback in simulation for learning physics. *Computers & Education, 57*, 2281-2290.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology, 52*, 139-183.

Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of cognitive neuroscience, 13*, 786-792.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic bulletin & review, 15*, 495-514.

Hostetter, A. B., Alibali, M. W., & Kita, S. (2007). I see it in my hands' eye: Representational gestures reflect conceptual demands. *Language and Cognitive Processes, 22*, 313-336.

Iverson, J. M., & Goldin-Meadow, S. (1998a). *The nature and functions of gesture in children's communication*: Jossey-Bass Publishers.

Iverson, J. M., & Goldin-Meadow, S. (1998b). Why people gesture when they speak. *Nature, 396*, 228-228.

Iverson, J. M., & Goldin-Meadow, S. (Eds.). (1998c). *The Nature and Functions of Gesture in Children's Communication: New Directions for Child and Adolescent Development, No. 79* (Vol. 37). San Francisco: Jossey-Bass.

Jamalian, A. (2014). Grouping Gestures Promote Children's Effective Counting Strategies by Adding a Layer of Meaning Through Action.

Jamalian, A., Giardino, V., & Tversky, B. (2013a). *Gestures for Thinking.* Paper presented at the CogSci.

Jamalian, A., Giardino, V., & Tversky, B. (2013b). *Gestures for Thinking.* Paper presented at the Proceedings of the 35th Annual Conference of the Cognitive Science Society.

Jang, S. (2010). *From Vesalius to Virtual Reality: How Embodied Cognition Facilitates the Visualization of Anatomy*: ERIC.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology, 106*, 86.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Johnson-Glenberg, M. C., Birchfield, D., & Usyal, S. (2009). SMALLab: Virtual geology studies using embodied learning with motion, sound, and graphics. *Educational Media International, 46*, 267-280.

Kang, S. (2012). As I sow, so shall you reap: The different roles of different gestures in knowledge construction.

Kendon, A. (1986). Some reasons for studying gesture. *Semiotica, 62*, 3-28.

Kerr, B. (1973). Processing demands during mental operations. *Memory & Cognition, 1*, 401-412.

Kessels, R. P., Van Zandvoort, M. J., Postma, A., Kappelle, L. J., & De Haan, E. H. (2000). The Corsi block-tapping task: standardization and normative data. *Applied neuropsychology, 7*, 252-258.

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and cognitive processes, 24*, 145-167.

Kita, S. (Ed.) (2000). *How representational gestures help speaking*. Cambridge: Cambridge Unuversity Press.

Klatzky, R. L., Pellegrino, J. W., McCloskey, B. P., & Doherty, S. (1989). Can you squeeze a tomato? The role of motor representations in semantic sensibility judgments. *Journal of memory and language, 28*, 56-77.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.

Krauss, R. M., Chen, Y., & Chawla, P. (Eds.). (1996). *Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us?* (Vol. 28). San Diego: Academic Press.

Krauss, R. M., Chen, Y., & Gotfexnum, R. F. (Eds.). (2000). *Lexical gestures and lexical access: a process model* (Vol. 2). Cambridge: Cambridge University Press.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction, 18*, 495-523.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. *Chicago: University of Chicago Press*.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic books.

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic books.

Marley, S. C., Levin, J. R., & Glenberg, A. M. (2007). Improving Native American children's listening comprehension through concrete representations. *Contemporary Educational Psychology, 32*, 537-550.

Marques, J. F. (2006). Specialization and semantic organization: Evidence for multiple semantics linked to sensory modalities. *Memory & Cognition, 34*, 60-67.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive science, 29*, 587-625.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological review, 92*, 350.

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

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

McNeill, D., & Levy, E. (Eds.). (1982). *Conceptual representations in language activity and gesture*. Chichester, England: Wiley & Sons.

Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and cognitive processes, 22*, 473-500.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review, 63*, 81.

Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2009). *Communicative Gestures and Memory Load.* Paper presented at the Proceedings of the 31st Annual Conference of the Cognitive Science Society.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 615.

Morsella, E., & Krauss, R. M. (2004). The role of gestures in spatial working memory and speech. *The American journal of psychology*, 411-424.

Myerson, J., Hale, S., Rhee, S. H., & Jenkins, L. (1999). Selective interference with verbal and spatial working memory in young and older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 54*, P161-P164.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review, 24*, 27-45.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying Properties from Different Modalities for Concepts Produces Switching Costs. *Psychological Science, 14*, 119-214.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic bulletin & review, 11*, 164-167.

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

Perry, M., Woolley, J., & Ifcher, J. (1995). Adults' abilities to detect children's readiness to learn. *International Journal of Behavioral Development, 18*, 365-381.

Pfeifer, R., & Scheier, C. (2001). *Understanding intelligence*. Cambridge, MA: MIT press.

Ping, R., & Goldin-Meadow, S. (2010). Gesturing saves cognitive resources when talking about nonpresent objects. *Cognitive science, 34*, 602-619.

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on psychological science, 1*, 110-122.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and brain sciences, 22*, 253-279.

Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences, 103*, 7865-7870.

Reddy, M. (Ed.) (1979). *The conduit metaphor* (Vol. 2). Cambridge: Cambridge University Press.

Repp, B. H., & Knoblich, G. (2004). Perceiving action identity how pianists recognize their own performances. *Psychological Science, 15*, 604-609.

Scherr, R. E. (2008). Gesture analysis for physics education researchers. *Physical Review Special Topics-Physics Education Research, 4*, 010101.

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

Schwartz, D. L., & Black, T. (1999). Inferences through imagined actions: Knowing by simulated doing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 116.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences, 3*, 417-424.

Segal, A. (2011). *Do Gestural Interfaces Promote Thinking? Embodied Interaction: Congruent Gestures and Direct Touch Promote Performance in Math*: ERIC.

Shapiro, L. (2007). The embodied cognition research programme. *Philosophy compass, 2*, 338-346.

Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental science, 11*, 655-661.

Simmons, W. K., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex, 15*, 1602-1608.

Singer, M., Radinsky, J., & Goldman, S. R. (2008). The role of gesture in meaning construction. *Discourse Processes, 45*, 365-386.

Smith, L. B. (2005). Cognition as a dynamic system: Principles from embodiment. *Developmental Review, 25*, 278-298.

Solomon, K. O., & Barsalou, L. W. (2001). Representing properties locally. *Cognitive psychology, 43*, 129-169.

Solomon, K. O., & Barsalou, L. W. (2004). Perceptual simulation in property verification. *Memory & Cognition, 32*, 244-259.

Spence, C., Nicholls, M. E., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics, 63*, 330-336.

Sperry, R. W. (1952). Neurology and the mind-brain problem. *American scientist, 40*, 291-312.

Steels, L., & Brooks, R. (1995). *The" artificial life" route to" artificial intelligence": Building situated embodied agents*. Hillsdale, NJ: Erlbaum.

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

Stieff, M., Dixon, B. L., Ryu, M., Kumi, B. C., & Hegarty, M. (2014). Strategy training eliminates sex differences in spatial problem solving in a stem domain. *Journal of Educational Psychology, 106*, 390.

Stieff, M., Lira, M. E., & Scopelitis, S. A. (2016). Gesture supports spatial thinking in STEM. *Cognition and Instruction, 34*, 80-99.

Stull, A. T., Barrett, T., & Hegarty, M. (2013). Usability of concrete and virtual models in chemistry instruction. *Computers in Human Behavior, 29*, 2546-2556.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science, 12*, 257-285.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory* (Vol. 1): Springer.

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Symes, E., Ellis, R., & Tucker, M. (2007). Visual object affordances: Object orientation. *Acta psychologica, 124*, 238-255.

Thelen, E., & Smith, L. B. (1996). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT press.

Tsuei, L., & Black, J. (2004). Using simulation diagrams to facilitate reasoning about mechanisms and systems. *Retrieved from Dissertations and Theses database.(AAT 3129039)*.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human perception and performance, 24*, 830.

Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta psychologica, 116*, 185-203.

Turner, M. (1996). *The literary mind: The origins of thought and language*. New York: Oxford University Press.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills*.

Vanetti, E. J., & Allen, G. L. (1988). Communicating environmental knowledge the impact of verbal and spatial abilities on the production and comprehension of route directions. *Environment and Behavior, 20*, 667-682.

Wagner, S. M., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of memory and language, 50*, 395-407.

Warburton, E. C., Wilson, M., Lynch, M., & Cuykendall, S. (2013). The Cognitive Benefits of Movement Reduction Evidence From Dance Marking. *Psychological Science, 24*, 1732-1739.

Wesp, R., Hesse, J., Keutmann, D., & Wheaton, K. (2001). Gestures maintain spatial imagery. *The American journal of psychology, 114*, 591.

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition, 68*, 77-94.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review, 9*, 625-636.

Wilson, N. L., & Gibbs, R. W. (2007). Real and imagined body movement primes metaphor comprehension. *Cognitive science, 31*, 721-731.

Wohlschläger, A., & Wohlschläger, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human perception and performance, 24*, 397.

Zwaan, R. A., Van Der Stoep, N., Guadalupe, T., & Bouwmeester, S. (2012). Language comprehension in the balance: the robustness of the action-compatibility effect (ACE). *PLoS ONE, 7*, e31204.

# Appendix

## Appendix A: Mental Rotation Task

Name: ______________________________

Student no: __________________________

Date: _______________________________

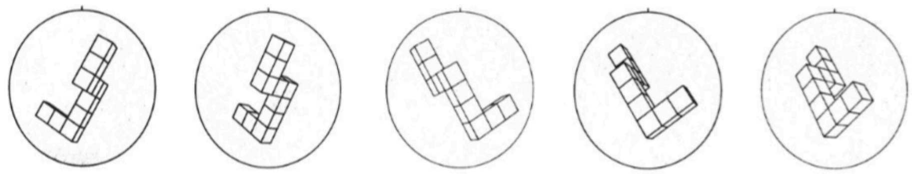
1. Your gender

   a) Male
   b) Female

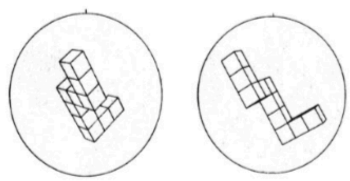
1

**Mental Rotation Test**

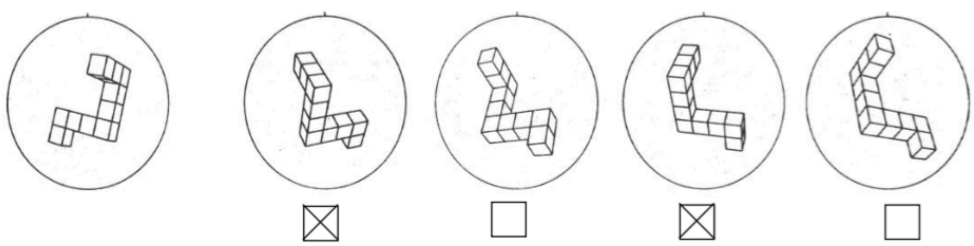
This is a test of your ability to look at a drawing of a given object and find the same object within a set of dissimilar objects. The only difference between the original object and the chosen object will be that they are presented at different angles. An illustration of this principle is given below where the same single object is given in five different positions. Look at each of them to satisfy yourself that they are only presented at different angles from one another.



Below are two drawings of new objects. They cannot be made to match the above five drawings. Please note that you may not turn over the objects. Satisfy yourself that they are different from the above.

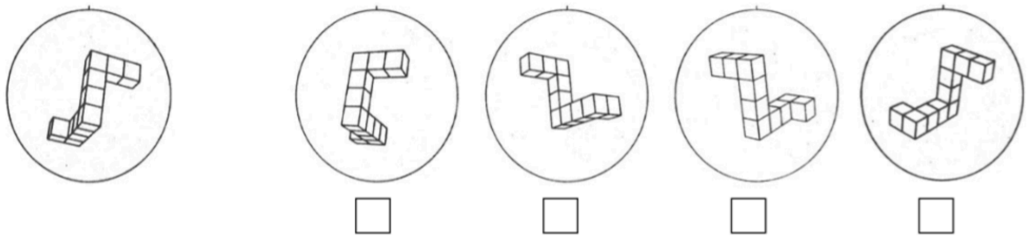


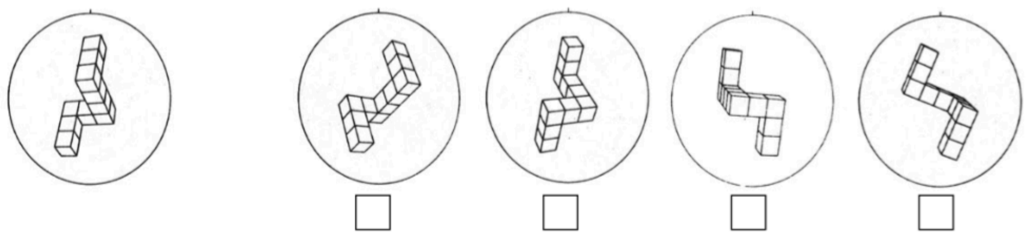
Now let's do some sample problems. For each problem there is a primary object on the far left. You are to determine which two of four objects to the right are the same object given on the far left. In each problem always <u>two</u> of the four drawings are the same object as the one on the left. You are to put Xs in the boxes below the correct ones, and leave the incorrect ones blank. The first sample problem is done for you.

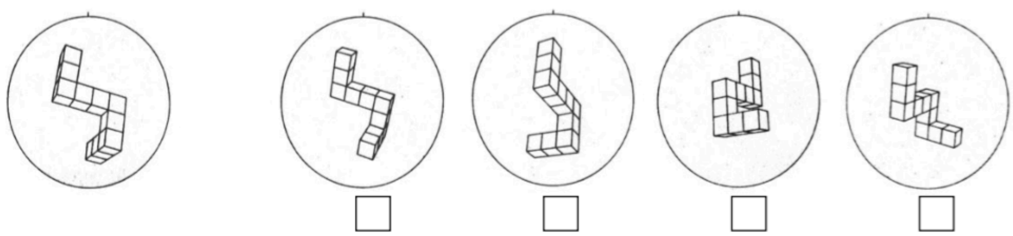


*Proceed to the next page*

Do the rest of the sample problems yourself. Which two drawings of the four on the right show the same object as the one on the left? There are always two and only two correct answers for each problem. Put an X under the two correct drawings.







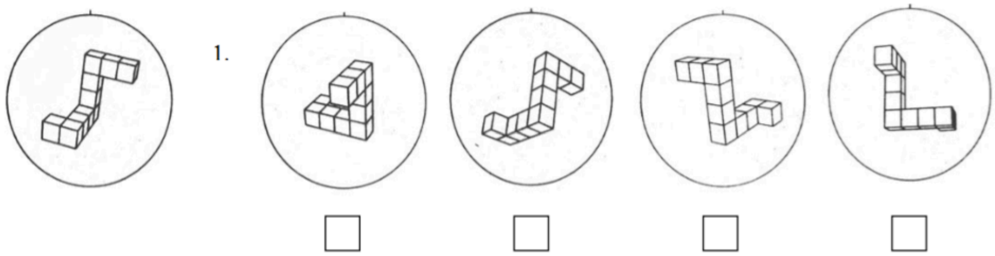
Answers: (1) first and second drawings are correct
(2) first and third drawings are correct
(3) second and third drawings are correct

This test has two parts. You will have 3 minutes for each of the two parts. Each part has two pages. When you have finished Part I, STOP. Please do not go on to Part II until you are asked to do so. Remember: There are always two and only two correct answers for each item.

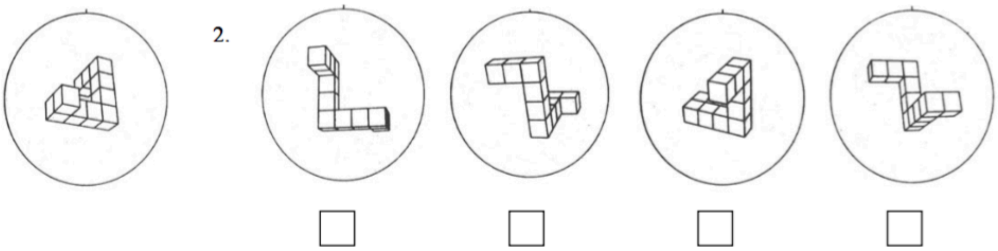
Work as quickly as you can without sacrificing accuracy. Your score on this test will reflect both the correct and incorrect responses. Therefore, it will not be to your advantage to guess unless you have some idea which choice is correct.

<u>DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO</u>

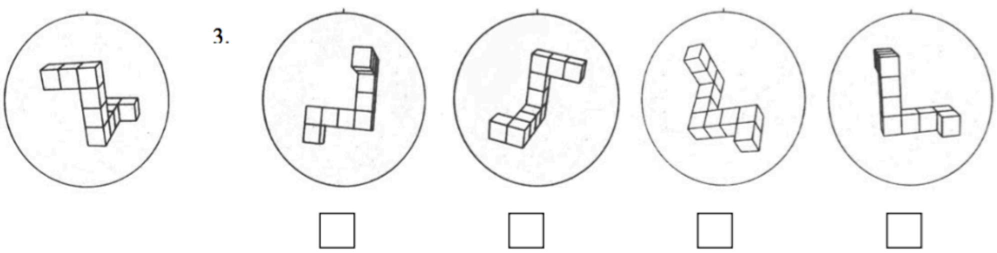
1.



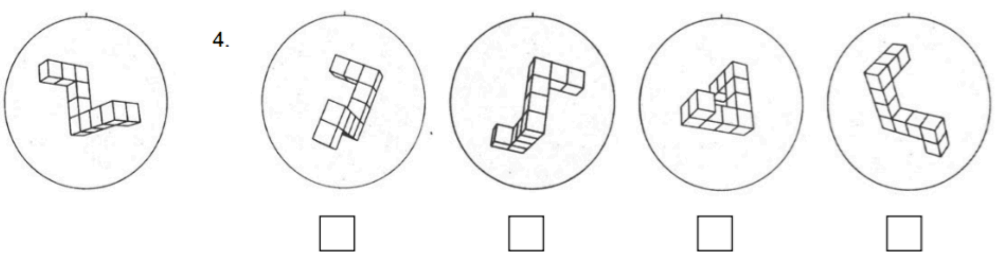
2.

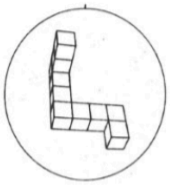


3.

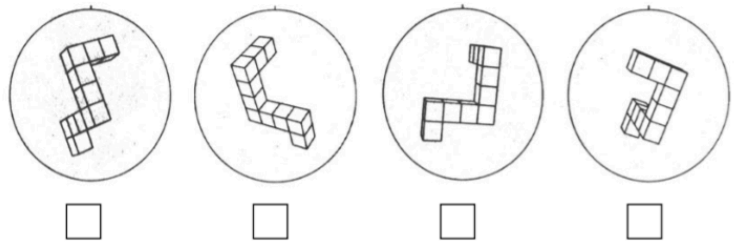


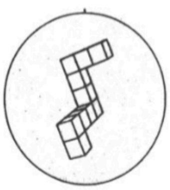
4.



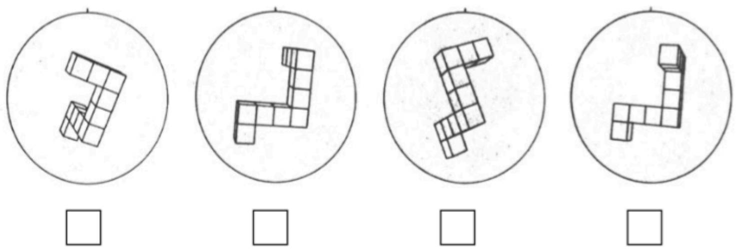


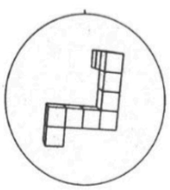
5.



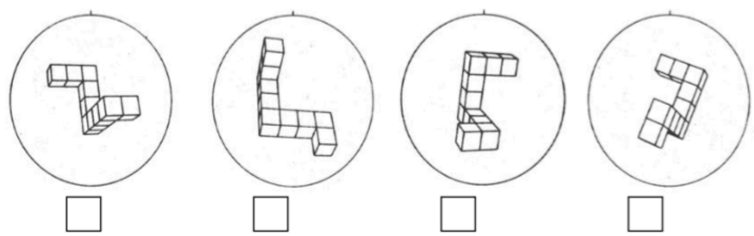


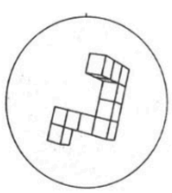
6.



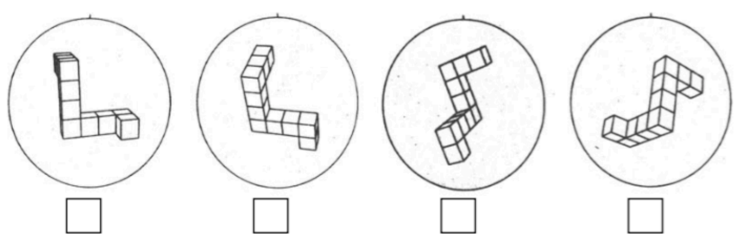


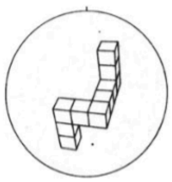
7.



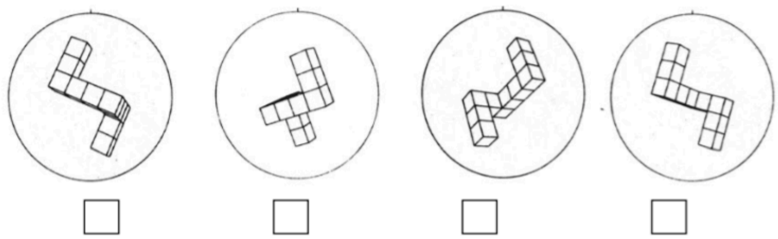


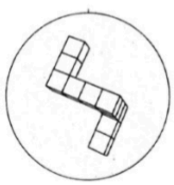
8.



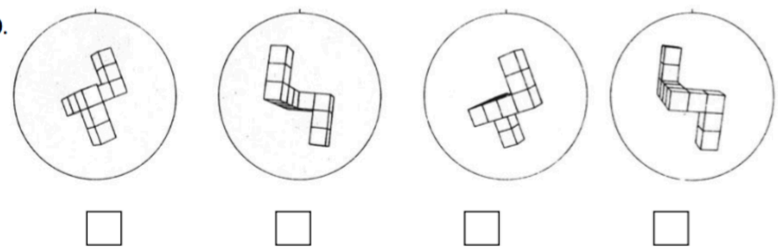


9.





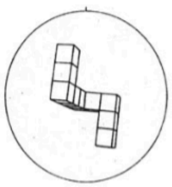
10.



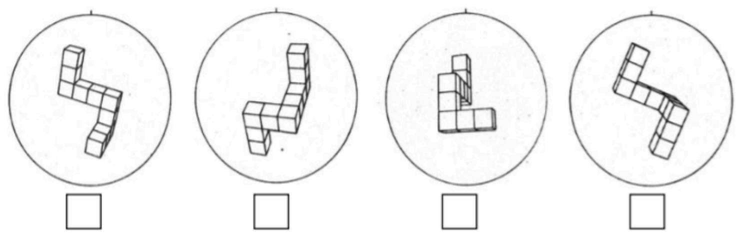
DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.

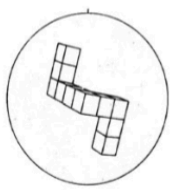
**STOP**

## PART II

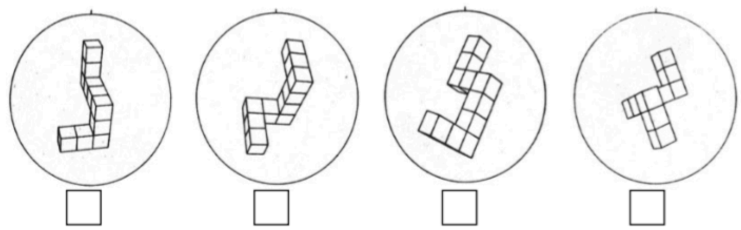


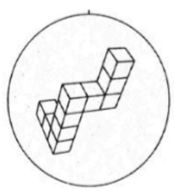
11.



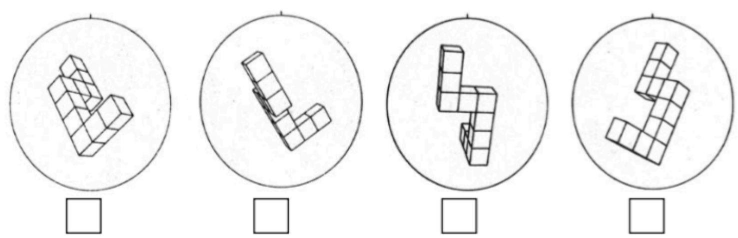


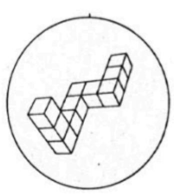
12.



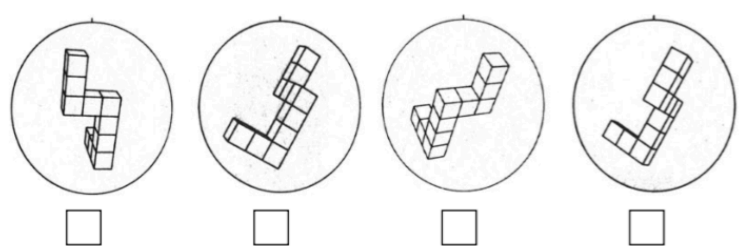


13.



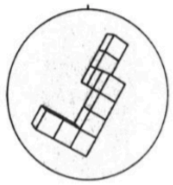


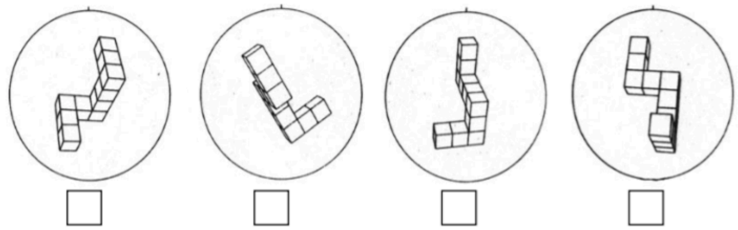
14.



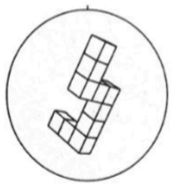
*Proceed to the next page*

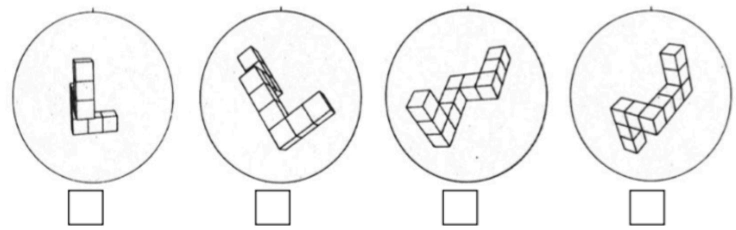
15.



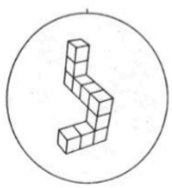


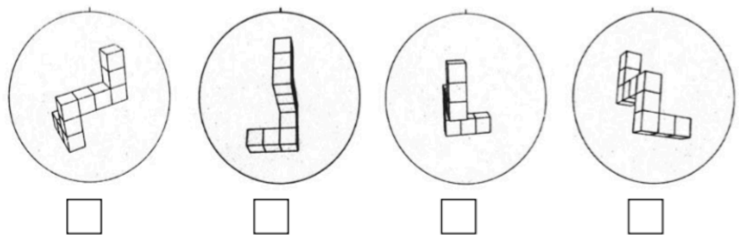
16



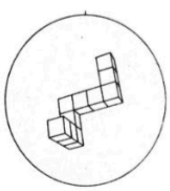


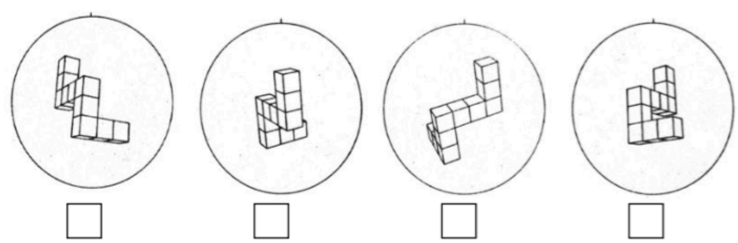
17.



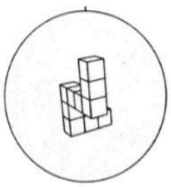


18

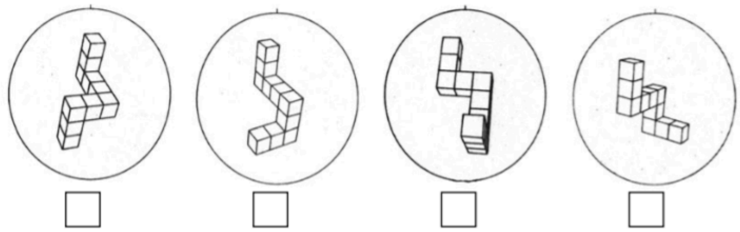


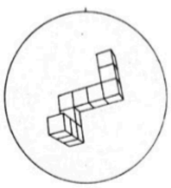


*Proceed to the next page*

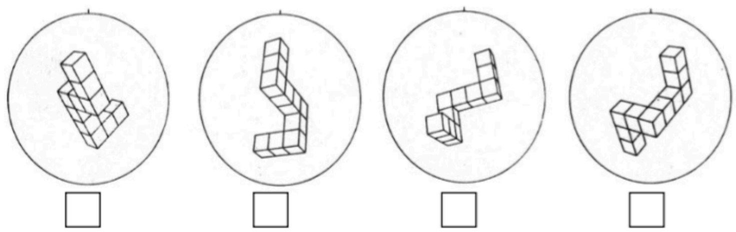


19





20.

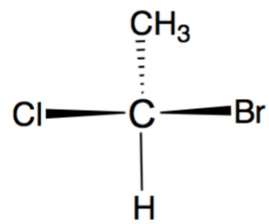


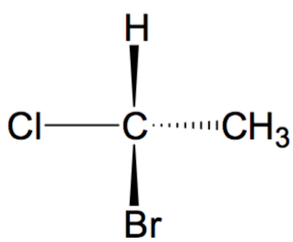
END

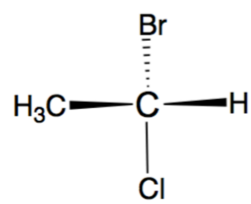
Appendix B: Primary Task

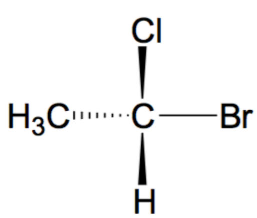
**Mental Rotation Task**

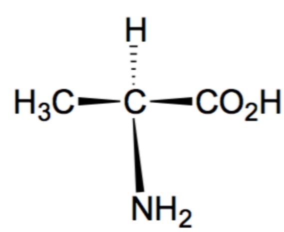
Please decide if the following molecules are R- or S-configuration according to what you just learned.

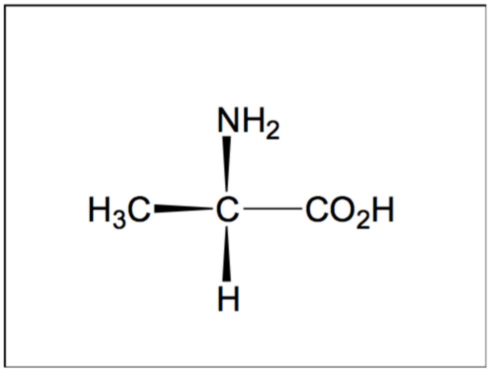


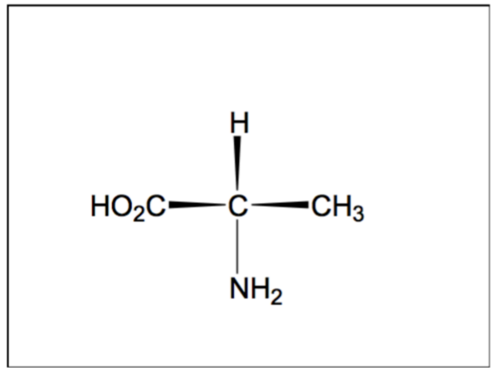



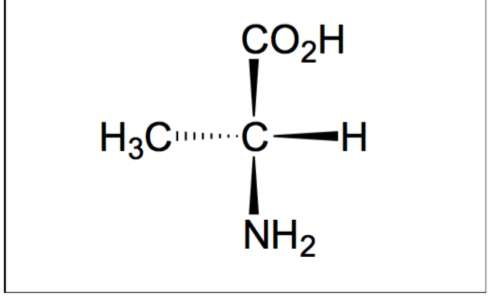





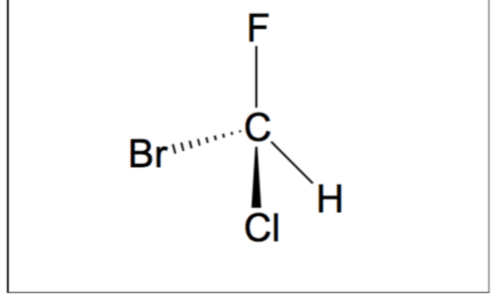
NH2
H3C
C
CO2H
H

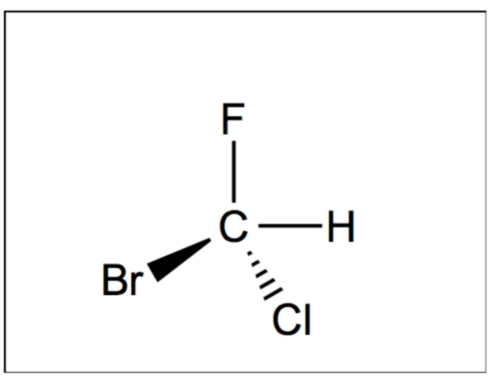
H
HO2C
C
CH3
NH2

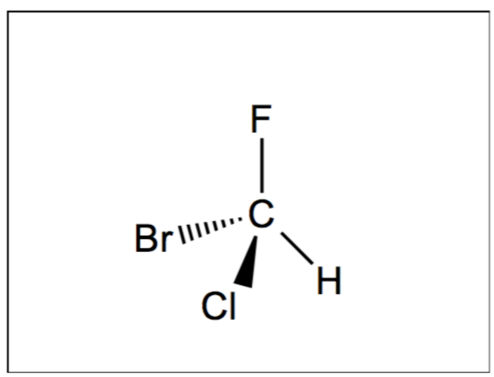
CO2H
H3C
C
H
NH2

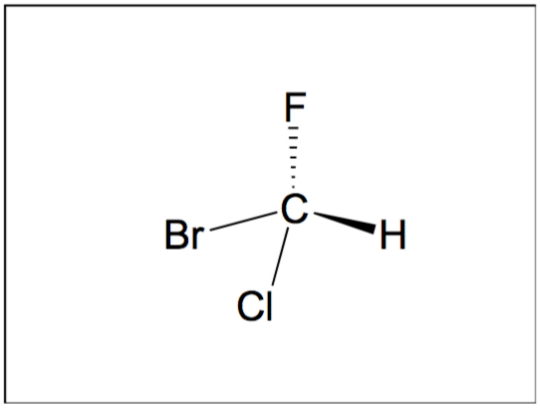
F
Br
C
H
Cl

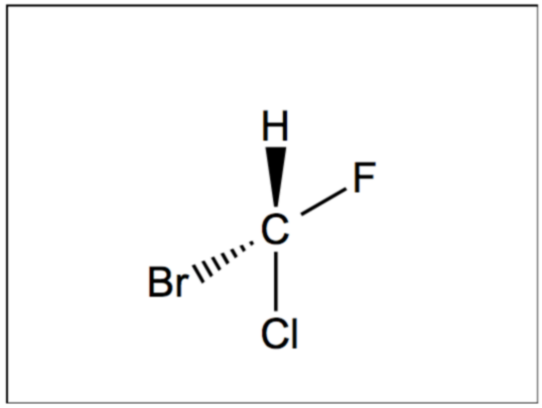



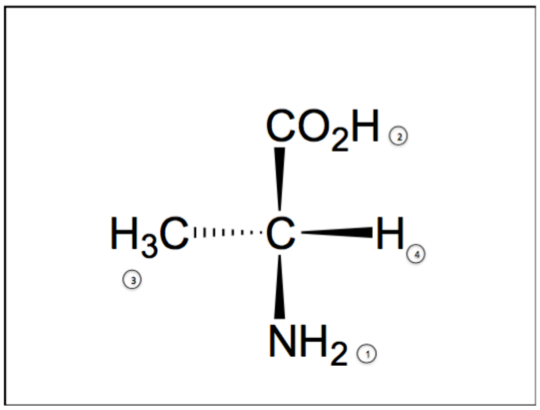
F
C
H
Br
Cl

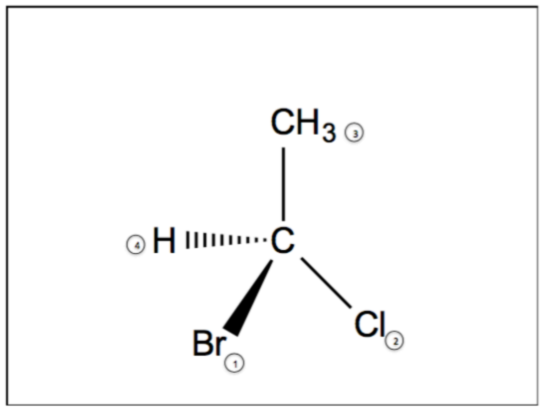



F
Br
C
H
Cl


F
Br
C
H
Cl


H
F
C
Br
Cl


CO2H 2
H3C
3
C
H 4
NH2 1


CH3 3
4 H
C
Br 1
Cl 2

Appendix C: NASA Task

Participant ID____________ Date____________

According to your experience on the molecule problems, please circle the location to indicate your response.

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

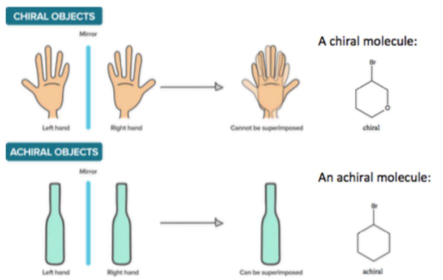
Very Low — Very High

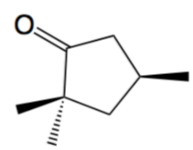
Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

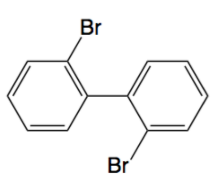
Very Low — Very High

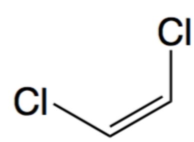
# Appendix D: Transfer Task

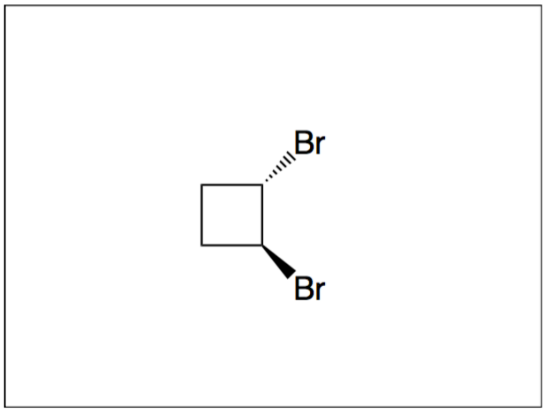
Check these examples, and decide if the follow six molecules are chiral (not identical to its mirror image) or achiral (identical to its mirror image).



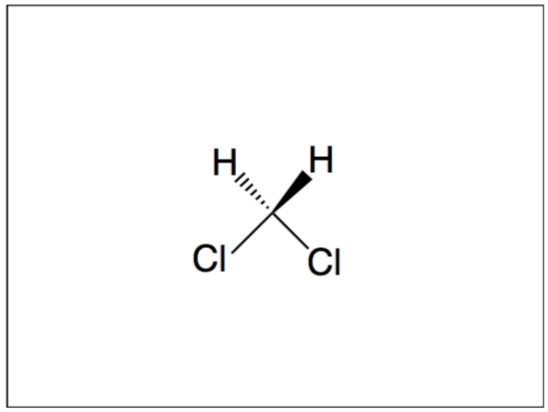



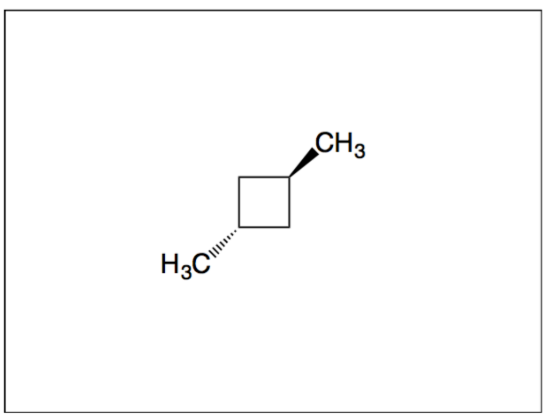





Br
Br

H
H
Cl
Cl

CH3
H3C