As I sow, so shall you reap:

The different roles of different gestures in knowledge construction

Seokmin Kang

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
2012

ABSTRACT

As I sow, so shall you reap:

The different roles of different gestures in knowledge construction

Seokmin Kang

Gesture researchers have focused on how gestures benefit learning. For example, data have shown that the usage of gestures during conversation can enhance concept acquisition and language learning in children. An instructor's gestures also benefit students' learning. For example, by providing gestures that contain semantic value, students remember more and attain deeper understanding of a concept. However, few studies have attempted to find out the function of gestures in learning: how information in a speaker's gestures is represented and constructed in a listener's mind. The present study targets learning of STEM concepts, especially the structure and the behavior of complex systems. It was expected that certain gestures prime a specific type of knowledge. For example, iconic gestures with structure knowledge of a concept facilitate learning of structures of a given concept and action gestures facilitate learning of movements, especially causal relation of the concept. This study also explored the relation between gestures delivered by a speaker and gestures constructed by a listener; in particular, if provided gestures contribute to constructing and representing a listener's knowledge and how it is manifested by

learners’ explanations.

Participants were randomly assigned to either an action gesture group that watched an instructional video based on action gestures, or to a structure gesture group that watched an instructional video based on structure gestures. The instructional video was about how a four stroke engine works. Except for a type of gestures that a speaker used, both videos were identical in all conditions. Participants were told that after watching the video they would explain a concept in the video to a colleague coming later, therefore a video camera would record their explanation, and the colleague would learn the concept from watching the video that they created. The participants watched the instructional video, and then they were asked to answer questions that were created based on a speaker’s verbal script. This was followed by a drawing test, which asked them to draw how a four stroke engine works based on the video that they watched.

Findings showed that action gestures facilitated action knowledge of the concept and were more involved in creating a mental representation of the concept based on action. Also, the structure group represented the concept based on structure. The findings were confirmed by analyzing the participants’ gestures and speech showing that the action group used more action gestures and action information units in their explanation and the structure group delivered reliably more structure gestures and structure information units. It was assumed that the mental model of the concept that the action group was harboring was based on action and the structure group was harboring was based on structure of the concept. The knowledge representations that the participants showed corresponded to the type of knowledge within the speaker's gestures in the instructional video that they watched. The results imply that listeners' knowledge is grounded in a speaker's gestures and this relationship depends on gesture type. More specifically, information in gestures is processed and becomes listeners' knowledge based on an attribute that the speaker's gesture has, and speech and gesture work together to manifest this phenomenon.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I am deeply grateful to many communities of people who have made this process possible. First, I thank the members of my dissertation committee. Thank you, Professor John Black for everything - it is impossible to list all for which I am indebted to you. I am grateful to Professor Barbara Tversky for the invaluable guidance; to Professor Stephen Peverly for the advice during the various stages of the doctoral program; to Professor Matthew Johnson for providing a chance to look at the data from a different perspective; and to Professor Lisa Son for the valuable comments provided both at the defense and everyday life.

Wherever you guys are now, I am greatly indebted to my peers and friends from the Coglab who helped me both with academic and emotional support; Gregory Hallman, Julie Youm, Jared Katz, Susan Jang, MingTsan Lu, and Saadia Khan. I thank Dr. Fenot Aklog for the careful considerations at work and Michelle Morales for giving me energy with a beautiful smile. I also thank Daniel Deihle and Chun-Hao Chang for your help. I especially thank to Minsok Lee and Sung Young Lee for always supporting me.

Finally, thank you, Father, Mother, and Seokkyun. Thank you, Father in law; Mother in law; and my nieces, Yoonbin and Curi for growing healthy. Thank you, Miyoung and Ddol-ddori for being part of my life.

CHAPTER I

INTRODUCTION

When an instructor explains a concept, he often uses hands with words. Using his hand(s), he may refer to an object, show movements of an object, and even draw out a causal relation of the movements of each part. Therefore, when an instructor explains a concept, learners not only learn from speech, but also learn from the instructor's hands. In this sense, for learners considering both an instructor's speech and gesture is crucial in understanding of a concept.

Gesture studies have focused on the role of gestures in speakers and listeners (for example, if and how gestures benefit listeners and speakers). The usage of gestures during conversation can enhance concept acquisition in children (Church, Ayman-Nolley, & Mahootian, 2004). Also, teachers' gestures can affect student learning (e. g., Alibali & Nathan, 2007; McGregor, Rohlfinig, Bean, & Marchner, 2009). It has been known that providing gestures that contain semantic values help students remember more and attain deeper understanding of a concept.

However, compared to the mechanism of speech information process, it is rarely investigated how learners represent information and construct knowledge from a speaker's gestures and what relationship exists between knowledge in gestures and knowledge in learners' minds. Therefore, the purpose of the current study is to explore how information in a speaker's gestures is transferred to listeners and becomes their knowledge. Especially, this study explores the function of a speaker's gestures in listeners' knowledge representation and construction.

Research Hypotheses

The goal of this study is to explore the mechanism of knowledge transfer within gestures from a speaker to a listener. Two types of knowledge are considered in the experiment: action knowledge, and structure knowledge.

Research Questions

The role of a speaker's gesture on a listener's knowledge construction

1. Does a specific type of gesture prime listeners' construction of a specific type of knowledge and mental representation, which leads listeners to learn a concept?
2. Do listeners actually understand a concept based on the representation that is grounded in the speaker's gestures?
3. Does the usage of information in participants' speech correspond to the results from the posttests and gesture analysis?

Hypotheses

Based on the review of the literature as presented in the next chapter, the following hypotheses are made with respect to the research questions posed for this study:

The action group will show strength in action knowledge on posttests and represent a given concept based on action while the structure group will show better performances on structure knowledge and represent a given concept based on structure. This will be confirmed by analyzing participants' gestures and speech.

Overview of Dissertation

This dissertation is organized into four chapters. Chapter 2 provides a review of the relevant literature and has five main sections. This chapter begins with studies showing knowledge embedded in an explainer’s hands, followed by introduction of gesture type, the role of gestures in revealing one’s thinking and in helping one’s thinking, and the effect of a speaker’s gesture on a listener’s knowledge construction.

Chapter 3 describes the method used to investigate the research questions for this study. The selection of participants, the design of the study, the materials and the measures are explained. In addition, the methods by which data was coded and analyzed are described.

Chapter 4 presents the analysis and results of the study in the context of the hypotheses set forth in Chapter 1.

Finally, chapter 5 provides an interpretation and a general discussion of the study's collective findings, limitations, and it concludes with suggestions for future research and implications for how this study may impact theories of how people represent and construct knowledge in educational settings.

CHAPTER II

LITERATURE REVIEW

Information embedded in one's hands

In most educational settings, teachers use visual aids – (e.g., charts, pictures, maps, or diagrams) - to help students better understand the concepts that are being taught. Here, we focus on a particular type of visual aid that is seldom noticed – the teachers' hand gestures. During all conversation, but perhaps particularly in explanation, hand gestures contribute to the message. Instructors often point to objects as they describe them. They may point at two objects with one hand or with both hands at the same time to suggest a relationship between them, trace a virtual line between them or trace a movement that is hidden. While teaching, instructors may be aware of their gestures but they may not know the exact information being conveyed by their gestures. There are different types of gesture depending on the referent, information value, stage that the gesture is shown, and types of knowledge that the gesture conveys. Depending on its referent, gesture can be classified as deictic, iconic, metaphoric, and beat gesture. Given that deictic, iconic, and metaphoric have a semantic value, they are also called representational gestures (Holler & Beattie, 2003). When it comes to gesturing with a diagram, gestures can be divided based on where the gestures are delivered. It can be either over that diagram, on-diagram, or away from the diagram, off-diagram. Gestures can also be divided into action gestures and structure gestures depending on information that is contained in the gesture. Action gesture delivers action information of an object or process of a system. Structure gesture represents location or shape information of an object.

A speaker's gestures work together with accompanying speech to depict or represent a referent or a given situation (Streeck, 1993). Even though they do not always represent the same semantic information, gestures can either complement or conform to meaning that is delivered in speech. Bavelas (1994) argued that gestures help convey meaning to a listener at its particular moment in the conversation. In the study, she transcribed part of her students' conversation with accompanied gestures and showed that gesture had a meaning that was not expressed in speech.

Therefore, speech and gestures complement each other. Researchers have argued that this relationship between gesture and speech is similar to the interaction of speech with a diagram (Cassell et al., 1994; Tversky, Heiser, Lee, & Daniel, 2009). Just as a diagram fills out the gap resulting from a lack of words or phrases, gestures can portray information that is hard to convey with speech, such as the structure of an object, the respective location of objects, and the form of an action.

Studies that have investigated the relation between speech and gestures tell us about what types of information are delivered via hands while depicting a referent or some other situations. For example, in storytelling, descriptions of action were delivered with iconic gestures, introductions of new characters were delivered with deictic gestures, and the boundaries of episodes were delivered with beat gestures (Cassell & McNeill, 1991). Schwartz and Black (1996) found that in explaining the operation of gears participants used hands to show the movements and interlocking of gears. In this case, information of both directional and action information was delivered via iconic gestures. Iconic gestures also contain location information and an environmental model. In describing complex environments, participants used deictic gestures to place landmarks on a map in the air (Emmorey, Tversky, & Taylor, 2000). In explaining how a lock works, participants used deictic gestures in a diagram to locate parts and iconic gestures in the air to show action of parts (Engle, 1998). Even though the contexts tested differed throughout the studies, each gesture played a different role in depicting a referent or a situation. Therefore, explainers use different kinds of gestures to convey different types of knowledge. Also, we need to pay attention to not only what they say, but also to the information that is embedded in a speaker's hands.

Types of gestures

Gestures are categorized differently depending on their referent, form and convention (McNeill, 1992). Depending on its referent, gesture type can be divided into four types. First, *beat* gestures are defined as “movements that do not present a discernible meaning” (McNeill, 1992). Beat gestures, as a unit of non-narrative units, are motorically simple, rhythmic gestures that do not depict semantic content related to speech (Alibali, Heath, & Myers, 2001). This type of gesture serves to structure and advance the discourse (e.g., raising and lowering the right and left hands for “on the one hand,” “on the other hands”). They are also used to stress certain points the speaker would like to make. The following three types of gestures that contribute to the semantic content of a message are distinguished from beat gestures. The first two are of specially importance. *Deictic* comes from the Greek word for finger and indeed, the prototypical deictic gesture is a point, for example, to a reference a person, object, or place. Deictic gestures can also be subcategorized, as *spatial deictic* and *literal deictic* gestures. *Spatial deictic* gestures convey direction of movement. *Literal deictic* gestures indicate concrete objects in order to refer to those objects or to similar ones. *Iconic* gestures resemble their references for example in form, in feature, and in action. In describing the proverbial “big fish,” a speaker may use iconic gestures to indicate the length of the fish, its shape, or its flailing, as it tried to get away.

Metaphoric gestures depict abstract referents or ideas metaphorically, where the hand movement does not directly correlate with physical information. They are like iconic gestures in that they exhibit some meaning relevant to concurrent linguistic meaning, but the relation to the linguistic meaning is indirect. This type of gesture presents an image of the invisible. For example, one can raise a hand to describe inflation in the economy, or point back over one's shoulder with a thumb to indicate an event in the past. An *emblematic* gesture has a specific, culturally differed meaning, such as the "o.k" sign.

However, these categories do not distinguish gestures that convey and indicate knowledge of structure from those that convey action. This distinction is crucial in many knowledge domains, from the study of science and engineering to history and literature. Structure is fundamentally spatial, the relations among parts. Action is temporal, changes overtime, often of structures and their parts. In science, we distinguish the parts of the heart from the behavior or action of the heart.

In engineering, the parts of an engine form its behavior. Similarly in drama and literature, there are characters and their interrelations. These can be distinguished from their actions and behaviors. Iconic gestures, and to some extents deictic gestures, can also be separated into those that convey structure knowledge and those that convey action or action knowledge. Structure gestures could be used to illustrate the shape of a heart and its four chambers. Action gestures could illustrate the operation or behavior of the heart similarly for an engine. To summarize, deictic and iconic gestures whether structure or action, can be used to represent reference or events by means of spatial and dynamic features.

Gesture reveals one's thinking

There are quite a few studies supporting that gesturing while people solve problems reveals how they represent knowledge. One of the studies showing how people represent their knowledge from observing their hands' movements came from the imaginary study by Schwartz and Black (1996). In their study, they wanted to observe how people reason about a system's behavior. They assumed that people usually use imagery to trace the system's phenomenology, not reason to trace the phenomenology. In the study, this phenomenon could be detected by observing gestures that people used during problem solving. In a step which was assumed to be a depictive model stage, participants splayed and moved their hands to mimic rotation of gears when they were asked to indicate the direction in which a designated gear would turn. However, when they could induce a rule based on the depictive model, they simply counted the number of gears in the problem. By observing their hand gestures, the researcher could infer how participants' gestures represented a series of rotating gears in their minds.

Another source of gesture's role in representing one's harboring knowledge is studies of speech-gesture mismatch. Church and Goldin-Meadow (1986), and Perry, Church, and Goldin-Meadow (1988) have shown that gesture-speech mismatch can be used as an index of children's transitional knowledge. They found that when children were asked to explain their judgments of a conceptual task, some children spontaneously produced gestures that convey the same hypothesis found in their spoken explanations, whereas other children conveyed one hypothesis in gesture and a different hypothesis in speech. They termed this gesture-speech discordance. In other words, the gestures produced by the children did not always convey the same procedure as the speech that accompanied that gesture. This finding suggests that children who produced gesture-speech mismatch simultaneously dealt with more than one procedure for approaching a conceptual task. The gesture-speech "mismatch" studies raised a possibility that gesture is another pathway through which mental representation is expressed, and representation in gesture is different than speech. Therefore, gestures can be thought of as a part of a speaker's mental representation of a given concept (Alibali et al., 1999).

Finally, through gestures, we can observe one's idea which was not expressed in speech (Church & Goldin-Meadow, 1986; Melinger & Levelt, 2004). For example, Hegarty, Mayer, Kriz, and Keehner (2005) observed that information in participants' spontaneous gestures during problem solving was not in speech. In the result, some of participant's thought could be only traced in their gesture, not in speech. This implies that we know what people know more completely by considering both speech and gesture.

Gestures help one's thinking

In addition to the role of gesture in revealing one's thinking, evidence from a series of studies have shown that gesturing helps speakers think and learn. These studies range from the benefit of simply pointing and touching numbers in counting (Alibali & DiRusso, 1999) to hands' movement in describing visual images (Wesp, Hesse, Keutmann, & Wheaton, 2001; Morsella & Krauss, 2004) and solving problems (Alibali, Spencer, Knox, & Kita, 2011; Cook & Goldin-Meadow, 2006; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001).

Alibali and DiRusso (1999) were interested in the effects of preschoolers' active gestures in counting, especially touching and pointing gestures. In counting a series of objects correctly, they assumed that two procedures would be involved. One is to "keep track of the objects that have been counted and those that are yet to be counted," and the other is to "coordinate reciting the string number words with tagging each object." They wondered whether the benefit of gesturing came from counting itself or children's active involvement of counting, which includes touching an object to be counted. To observe this, they had children follow three conditions: allowing active gesture, a puppet gesturing as the children counted, and prohibiting gesture. They found that both gesturing groups facilitated counting accuracy. In further analysis, children's active involvement in counting played more of a role than just keeping track of the counted objects. It helped children keep track of the objects, coordinate reciting number words and tag each of the items. This phenomenon was found regardless of whether children spontaneously gestured or were instructed to gesture. In this case, active gestures contributed to children's development of procedural competence at counting.

Gestures help one's thinking by sustaining spatial memory. Wesp, Hesse, Keutmann, and Wheaton (2001) were interested in whether or not motor activities like hand gestures help hold spatial stimuli, just as repetitive vocalizations or sub-vocalizations help one hold verbal information in an articulatory loop. In their study, participants were asked to describe a seascape water color painting either with the painting present or absent, and the number of gestures that the participants used was compared. It was found that participants used more gestures per second when the painting was absent while they explained. They concluded that gestures help maintain and facilitate spatial images.

The following two studies show the role gestures play on working memory and they suggests that using gestures would help one's thinking by lightening mental load. Goldin-Meadow, Nusbaum, Kelly, and Wagner (2001) examined if using gesture helps thinking by observing participants' mental capacity. In their study, participants were asked to remember word lists while explaining how they solved given math problems. Compared to participants who were not allowed to gesture, those allowed to gesture remembered more world lists during the explanation. It was assumed from the results that using gestures saves one's cognitive resources on the explanation task, allowing one to allocate more mental resources on the explanation task.

Gestures also help one's thinking by sustaining spatial representation in working memory (Morsella & Krauss, 2004). In this study, there were two between-subjects variables, body movement restricted or unrestricted and visual stimulus present or absent. There was also a within-subjects variable: codable versus uncodable stimuli. After viewing a visual stimulus, participants were told to describe a given stimulus under each condition. The researchers observed that participants gestured more when describing visual stimuli which were absent and when the stimuli were difficult to remember and describe verbally. In addition, they found that participants whose gestures were restricted showed lower speech rate than those whose gesture were unrestricted. The researchers explained that gesture restriction affected speech fluency. From the results, it was concluded that using gestures facilitates maintaining spatial information and this was mediated by its effect on spatial working memory.

Using gestures especially encourages children to generalize knowledge and it leads to learning. For example, in math equivalent problem instruction, when children were given an explanation that included a correct problem solving strategy in gestures, they were more likely to produce that strategy in their own gestures than those who were not given the strategy in gestures. Also, these children were more likely to succeed on a posttest than those who did not produce strategies in gestures (Cook & Goldin-Meadow, 2006). Finally, a recent study showed that even making children gesture helped elicit unexpressed implicit knowledge and facilitated learning (Broader, Cook, Mitchell, & Goldin-Meadow, 2007). The studies above suggest that gestures help one’s thinking both directly and indirectly.

Gestures help a listener's thinking

As gesture users can receive cognitive benefits from their own gesture use, can gesture viewers also benefit from observing gestures when constructing knowledge? With the idea that gestures can be a tool to uncover the unconscious, Riseborough (1981) attempted to explore the function of gestures in communication. In a study, she created videos to explain objects based on the presentation mode. One was “Sound Mode” where participants were only given verbal information without visual information. The other was “Face Mode” in which only a speaker’s face could be seen, and the third one was “Body Mode” where the speaker’s whole body with the face was seen. The results showed that participants with viewing gestures (Body Mode), as compared to others, performed better at objects guessing, recall, and recounting a story test.

Woodall and Folger (1985) considered gestures as one type of contextual cue in explanation. They investigated whether nonverbal cues, like hand gestures, played a role in retrieving conversational messages. In a study, they created three types of videos as stimulus material. One video had no gestures; the other contained emphasizing gestures such as pounding, pointing or chopping movements towards or away from the body; the third, called representational gestures, were semantically related to the co-occurring verbal information. For example, the sentence, “I can’t keep up with his eating, he just sits there and shovels it in.” was accompanied by several sweeping motions with one hand toward the mouth. Participants in the study were asked to watch one of the three videos, and after watching it, were administered a cued recall test. Results showed that participants in the representational gesture cue condition recalled more from utterances than the other two conditions. But, there was no significant difference between the emphasizing gesture group and the no cues group.

In another series of studies, it has been verified that an important role of gestures is that they contribute to a listener’s comprehension of spoken language or help a listener understand a concept (Beattie & Shovelton, 1999; Goldin-Meadow, Kim, & Singer, 1999; Singer & Goldin-Meadow, 2005; Thompson, Driscoll, & Markson, 1998; Valenzeno, Alibali, & Klatzky, 2003). An underlying idea in these studies is that by using hands a speaker or a teacher provides his listeners with more than one mode of information. For example, Thompson, Driscoll, and Markson (1998) found out that children’s memory was more influenced by representational gestures that appeared along with predicate terms than by gestures that co-occurred with nouns.

Researchers have also gained interest in the types of gesture from which learners benefit the most. Church, Ayman-Nolley, Estrada, Glover, and Dullum (2001) showed that using representational gestures with speech instruction has greater impact in children’s understanding of a given concept than using *emphasizing* or *no gesture* accompanied speech. In a follow-up study, Church, Ayman-Nolley, and Mahootian (2004) showed that children who watched a video with *representational* gestures and speech outperformed those who watched a video either with *emphasizing* gestures or *no gestures* in understanding the Piagetian conservation concept.

More specifically, it has been observed that participants take advantage of each other's gesture as a source to better understand a conversational topic or a concept (Goodwin, 1986; Heath, 1992) and even young children (1;8 – 2;0) benefit from a speaker's gestures (Mcgregor, Rohlfing, Bean, & Marschner, 2009). In a study, Mcgregor et al. (2009) investigated the role of gestures in children's comprehension of a spatial term, *under*, where an experimenter presented children item pairs' spatial relation with speech using gestures, photo, or no aided material. It was found that gestures promoted robust knowledge of the meaning of *under*, and the learning effect in the gesture group lasted longer than other groups.

Showing gestures without to-be-explained objects also helps children's understanding of a concept of the object to which the gestures are referring. Ping and Goldin-Meadow (2008) explored conditions under which iconic gestures promote learning. They were especially interested in whether children would be more likely to learn when given instruction containing speech and ungrounded iconic gestures than when given instruction containing speech alone when teaching conservation quantity. They mentioned that previous studies had mostly focused on the benefits of gestures referring to an object that is present from the visible context. It was observed that using iconic gestures made instruction more effective and gestures could help children learn even when they did not direct attention to concrete objects. A series of these studies show that the gestures complemented speech information and a speaker's gesture is helpful in learners' understanding of a concept.

What remains unanswered? Although we know that different gestures are derived from different knowledge representations (Kita & Özyürek, 2003) and speakers' gestures affect listeners' actions (Cook & Tanenhaus, 2009), we still do not know if different representations within gestures actually lead to learning.

According to Kita and Özyürek (2003), representational gestures originate from spatio-motoric representations which contain spatial information and information about actions. In other words, when a listener views representational gestures that have structural and spatial information of an object, it means that structural and spatial representation of the object is delivered throughout an explainer's hands. Therefore, the listener goes through different information processing processes when viewing a certain type of gesture than when viewing anther type of gesture.

This corresponds to the situated cognition theory that the environment plays a central role in shaping cognitive mechanisms (Gibson, 1979), and studies of priming or imitating other's actions in the social cognition area (Brass, Bekkering, & Prinz, 1999; Sebanz, Knoblich, & Prinz, 2003). For example, viewing specific actions of others can affect viewers' actions (Bach & Tipper, 2006; Ferguson & Bargh, 2004) and perceiving events produced by other's actions activates the same representation that controls one's own action (Jordan & Knoblich, 2004). It was also argued that actions are coded based on the perceptual events resulting from the actions (Sebanz, Knoblich, & Prinz, 2003).

Research Questions for the Present Study

From the embodiment theories above it is possible to reason that semantic value in a gesture helps represent a concept in certain ways and this finally leads to developing knowledge based on the semantic value in the gesture. Since it is still unknown how each gesture differently contributes to a listener's knowledge construction, this study explores: (1) if a specific type of gesture from a speaker is involved in a listener's constructing a specific type of knowledge, and (2) if it does, whether the knowledge constructed corresponds to mental representation that a listener has.

More specifically, it is hypothesized that participants in the action group will show higher performance at action questions than structure questions, which would, in turn, be reflected in participants' drawings. For example, participants who watch the action gesture video will use more action-related components such as arrows or action effect in the drawing test, since arrows in a diagram are related to functional organizations such as temporal, dynamic, and causal processes (Heiser & Tversky, 2006). Therefore, action gestures would play arrows-like role in a visual explanation task to express action or movement traces of a mechanical concept. Also, participants would use more action and process-related gestures when they explain the concept to a peer.

In the same way, structure gestures are expected to lead to listeners' constructing structure knowledge of the concept. Therefore, participants in the structure group will show better performance at structure questions than action questions. This will be represented through the drawing test. Participants' drawing will focus more on the structure of the engine.

In addition to analyzing the participants' drawings as a way of observing mental representation, to verify if attributes in drawings actually come from listeners' mental representation of the concept, participants' speech and gestures are analyzed. Since gestures are connected to a user's implicit knowledge (Garber, Alibali, & Goldin-Meadow, 1998), by looking at gesture patterns that the participants use we are able to trace if and how each type of gesture affects listeners' knowledge construction differently. It is expected that the action group will use more action gestures and the structure group will use more structure gestures when they explain the concept.

Again, the proposal in the present study is that action gestures better facilitate constructing action knowledge, and structure gestures better facilitate structure knowledge, since different gestures contain different symbolic information. By observing different gestures' different role in learning, better insight into the role of gesture on knowledge construction and representation can be obtained.

CHAPTER III

METHOD

Participants. 59 (15 male) students participated in the study from a university. They were all native English speakers and had not learned about a concept of a four stroke engine. Participants ranged in age from 20 to 36 with an average age of 26 (*SD* = 3.50).

Materials. To address the role of gesture in knowledge representation and construction, instructional videos were created, where the topic was about how a four stroke engine works. This learning material was chosen from science, technology, engineering and mathematics (STEM), especially because it delivers both structure and behavior information of a complex system. First, a verbal script of the concept was created (see Appendix A and B for a script). In this topic, structure of each part, the part's shape, a relative position of each part and its accompanying movements such as causal relations between each part's movements in a cylinder are key factors in understanding the concept. Considering these attributes in understanding the concept, I categorized nouns and verbs in the script into two types based on information embedded in the words: action information word, structure information word. I then created gestures which corresponded to a semantic value within each word. Information in each gesture which accompanied a word is listed in Appendix C and D. With these gestures and the script, two instructional videos were created: action gesture video, and structure gesture video. In each video, explaining a concept, a speaker delivered 22 gestures in front of him. The running time was slightly different in the two videos because the time needed to perform each gesture varied. The running time was 209 seconds for the action gesture video and 212 seconds for the structure gesture video.

The instructional video was composed of three parts: introduction, process explanation, and wrap up session. A different type of gesture was used in the “process” and “wrap up” session. In the “introduction” part, to provide viewers with a structural overview of a four stroke engine, relative positions of each part in a cylinder, was explained based on a diagram by using only deictic gestures, pointing to an object with an index finger.

In the “process explanation” part, the processes of four steps of stroke were explained. In the structure gesture video, a speaker explained structure of objects in a four stroke engine and showed static images of the objects using his hands without a diagram. These structure gestures delivered the shape of each object and its relative position in a cylinder. On the other hand, the speaker in the action gesture video showed movements of the objects and flows of a mixture of air and fuel. Finally, in the “wrap up” session, it was explained how these steps led to a rotational motion of a car’s wheels.

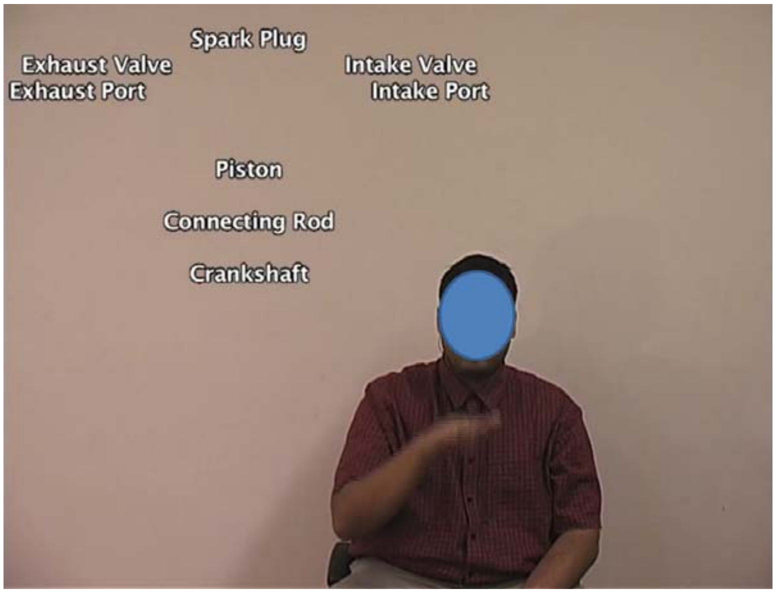


Figure 1. Action gesture video (209 sec)

For the action gesture video (Figure 1), a speaker showed the rotational motion of a crankshaft, the direction of piston movement, the flow of fuel and air, the movement of an intake and an exhaust valve, the flow of byproducts, and etc. with hands. In his action gestures, the speaker only showed each part's action or movement and avoided making any shape or positional information of each part to get rid of possibilities of delivering another type of information like structural information which was not intended to deliver. For example, for a part in the verbal script, "the first half rotation of the crankshaft pulls the piston downward inside the cylinder...," the speaker first drew a half circle with his index finger in front of his torso to show the rotational motion of the crankshaft. Then, reading a word "downward", he quickly placed his right hand horizontally with his palm downward at his chin level and lowered it to his chest level. This delivered information of the movement of the piston. Therefore, the action video was designed to have the participants more exposed to action information of a four stroke engine's working.

For the structure gesture video (Figure 2), using his hand(s), the speaker made a shape of a crankshaft, a piston and a cylinder, and showed positions of a piston, a crankshaft, a spark plug, an intake port, an intake value, an exhaust port, an exhaust valve and mixture of fuel and air. For example, for a part in the script, "the first half rotation of the crankshaft pulls the piston downward inside the cylinder...," the speaker first made a round shape with his two hands by touching two thumbs together upwards and the rest of the fingers together downwards, placing it towards the lowest part of his torso. This delivered information of the shape and the position of the crankshaft. Then, speaking a word "piston," the speaker made a rectangular shape with his two hands, placing it near the lower part of his torso but a bit higher than the position at which the crankshaft was placed. This gesture delivered information of a piston's shape and its relative position against the crankshaft. Finally, with a word "cylinder," the speaker placed his two lower arms with hands spread vertically. This showed the relative position of the crankshaft and the piston inside the cylinder. Throughout these series of structure gestures with a script, it was expected that the participants would be more exposed to the relative position and the shape of each part of a four stroke engine.

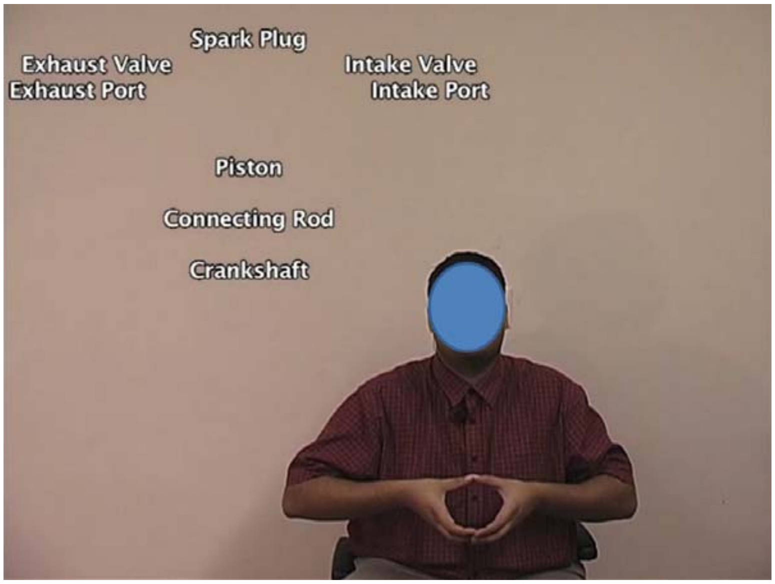


Figure 2. Structure gesture video (212 sec)

A schematic diagram was added by a speaker during his explanation. In the schematic diagram, names of each part were given. Given that participants had to process names related to parts in the cylinder, structure of each part, and its function, providing names of parts would help them focus more on exploring the functions and the structure of a four stroke engine system, instead of splitting their mental resources into memorizing various names in the system. On the other hand, it was expected that this treatment would contribute to developing relatively more structure knowledge of the concept than action knowledge for the participants in both groups.

Also, to get rid of the effect of lexical stress on information processing, the speaker tried to give the same lexical stress on all words that accompanied gestures. If each instructional video has different stressed words, these stressed words would play a role of verbal cues in the viewers' information processing and may intervene in the gestures' role in the instructional video. According to studies on the role of lexical stress in speech recognition, lexical stress information may affect intelligibility and facilitate word recognition (Field, 2005; Heuven, 1988). To minimize this possibility, a speaker practiced a verbal script before recording the instructional videos and then gave the same lexical accent to all words accompanied. With this manipulation, it was expected that the speaker's verbal explanation would not sound monotonous like reading a text. Therefore, the speaker practiced the script several times, making sure to stress the actions and the parts for both videos.

Posttest. To see if a different type of gesture contributed to developing a different type of knowledge, posttests were administered right after participants watched a video. The posttests were a knowledge test and a drawing test. The knowledge test was composed of 16 True/False and 4 multiple choice questions (see Appendix E). Among 16 True/False questions, 8 questions queried the participants' action knowledge and the other 8 questions queried their structure knowledge of the concept. Action knowledge queried knowledge of the action, causal relations between movements of each part or its subsequent results in a four stroke engine's working. For example, in the four stroke engine the understanding of the movements of the piston, the crankshaft, the valves in a cylinder, the flows of a mixture of air and fuel, and its accompanying movement were regarded as action knowledge. Structure knowledge refers to knowledge about shape of each object and its relative position in a cylinder. Finally, 4 multiple choice questions queried general knowledge of the given concept as a neutral condition to see if, regardless of knowledge type, there was a difference in knowledge benefits from a different gesture type. Therefore, the maximum score was 20 for the knowledge test, which was provided in random order.

The second posttest was the drawing test, which was also called a visual explanation test, where the participants were asked to draw how a four stroke engine worked based on what they learned from the instructional video. From the drawing test, as a way of observing how the participants represented a four stroke engine's working, the participants' usage of visual components in delivering the mechanism of the concept was analyzed. Five visual components were considered: action word, action arrow, action effect, labeling arrow, and labeling line. Action word refers to a verb, a participle, or a to-infinitive that has a meaning of action or movement. For example, *"piston moves up pushing fuel & air mixture out…"* is regarded as having 2 action words. Action arrow refers to an arrow that shows movement of a part or flow of a mixture of air and fuel. Action effect refers to any expression showing effect of explosion, density of air and fuel, or flow of air and fuel. For instance, if a participant expressed a star-like shape under the spark plug which showed ignition, it was regarded as action effect. A labeling arrow was regarded as an arrow which labeled what it referred to. Therefore, if a participant drew an arrow indicating a part within a cylinder, the arrow was regarded as a labeling arrow. Finally, labeling line referred to a line to label a part. It had the same function as the labeling arrow except there was no arrow shape at the tip of the line.

Procedure. Once a participant arrived, the participant was seated at a table with a laptop computer that had a 15.4 inch screen. The participant was randomly assigned to either one of two video groups: action gesture video, or structure gesture video group. The participant was then told that “Today, your job is to watch a video of how a four stroke engine works four times[1] in a row and explain the concept in the video to a peer coming later. However, since you are not directly explaining a concept, your explanation will be videotaped and showed later either to him or her. He or she will learn about the concept from your explanation.” The participants were not allowed to pause or stop, and to take notes while he was watching the instructional video. Also, the participants were asked to pay attention to the instructional video, even if the instructional

[1] The instructional video was shown through the laptop computer four times. This number was obtained based on an empirical study. In the empirical study, participants watched a similar instructional video twice and reached at most chance level in a posttest. Therefore, in the dissertation study, participants were given enough time to construct knowledge so that knowledge representation and construction can be properly measured as an index to see if they represented knowledge in a specific way based on a different gesture type.

video was boring. He was then told that an experimenter would be outside while the participant watched the instructional video to make him comfortable and concentrate on the video.

After participants watched the video, they were given a posttest packet which had a knowledge test and a drawing test in which the participants were asked to draw how a four stroke engine worked based on the video that they watched.

After the posttests, they were asked to explain how a four stroke engine worked in front of a video camera. They were told that another colleague would come and learn how a four stroke engine worked from their explanation. The video camera was set 3 meters away from the participants at the opposite side. The participants were asked to stand up against a wall and explain how a four stroke engine works in front of the video camera, and were encouraged to spend as much time as they wanted for the explanation. Then the experimenter went out of the room. When the explanation was done, the experimenter was called. And then the experimenter entered the room and debriefed the participant.

Gesture coding. To confirm the origin of the knowledge constructed, the participants' gestures were coded and analyzed. A gesture unit in this study was defined as "the period of time between successive rests of the limbs (McNeill, 1992)." Hand(s) movements starting from a resting position and returning to a resting position was regarded as one gesture. If the hands did not return to a resting position between two gestures, the boundary was defined by a pause in motion and an obvious change in shape or trajectory. When a participant used both hands simultaneously to describe one object, concept, or part, it was regarded as one gesture. If a participant used both hands and one described an object, a concept, or a part and the other hand a different concept, the gestures were coded as two different gestures.

In this study, the participants' gestures were categorized into two types depending on semantic value: action gesture, and structure gesture; and into another two types depending on its shape: invented gesture, and imitated gesture. Analyzing participants' gestures involved observing the function of a speaker's gestures and what type of representation they were harboring based on the instructional video; therefore only structure and action gestures referring to the engine and its movements were coded.

Action gestures deliver action information about an object or a process of a system, such as hands' movements showing action or movement of each part. Structure gestures deliver positional or structural information of an object, such as hands' movements showing static image of an object or pointing an object to show position or relative position of each part. For a gesture shape, if the participants' gesture was not shown in the instructional video that they watched, it was coded as *invented* and if a participants' gesture was the one that was shown in the instructional video, it was coded as *imitated*.

Inter-rater reliability was assessed on randomly selected 240 gestures (18%) of the data by a second coder who was trained and blind to the experimental design. Agreement for identifying gestures was 87.8%. Agreement for categorizing gestures was 99.6% for action and structure gesture and 90.0% for invented and imitated gesture.

Speech coding. Participants' verbal descriptions were segmented into propositions (i.e., the unit of meaning in a sentence). The information units were coded as either *action, structure*, or *other*.

A proposition with an 'is-a' or a 'has-a' is basically regarded as a structure information unit. For example, *"…on each side there are two valves…"* is one structure information unit. For an action information unit, if a proposition contains a predicate that is a verb indicating an argument's action or an argument being acted, it would be regarded as an action information unit.

Also, for a proposition that has a 'has-a' or an 'is-a,' if the meaning of argument(s) in the proposition is about action or movement, then the proposition is regarded as an action information. So, for example, *"...that's one half cycle rotation..."* is regarded as one action information. As another example, *"...there is the compression phase..."* is regarded as one structure information unit. Even though there is an argument, *"compression,"* which has a meaning of action, *"compression"* is not regarded as an action argument - because the word *"compression phase"* is the name of the stage in the instructional video. Therefore, if arguments that have a meaning of action were used to refer to a part or a stage in the instructional video, they would not be considered in deciding action or structure information.

The *other* information included greetings, such as "*Good evening,*" introductory information such as "*I'm going to explain how a four stroke engine works*," and meta-comments such as "*…let me tell you a little bit more about each stage…*"

CHAPTER IV

RESULTS

Knowledge test analysis

The test scores from 29 participants in the action group and 30 in the structure group were compared. Table 1 shows average scores of each group in the knowledge test.

Table 1. Mean scores in the knowledge test (SD)

	Action question [8]	Structure question [8]	General question [4]	Total [20]
Action group	6.93 (1.22)	5.76 (1.60)	2.69 (0.60)	15.38 (2.44)
Structure group	6.33 (1.32)	5.83 (1.44)	2.60 (0.97)	14.77 (2.82)

From UNIVARIATE ANOVA, no interaction was found in the scores for action and structure questions between group and question type, $F(1, 114) = 1.70$, $p = .20$. Also, in ANOVA test, there were no group differences in the total mean score of the knowledge test ($p = .38$). Also, there was no difference in both action question ($p = .08$), structure question ($p = .85$), and general question ($p = .67$) between the two groups. However, in within group comparisons, the action group received higher scores on the action question than the structure question, $t(28) = 3.56$, p

< .01, *d* = 0.82, while there was no score difference between action questions and structure questions within the structure group (*p* = .11).

Table 2. Independent Chi-square in each item. Numerator is the number of participants who got correct answer and denominator is a total number of participants.

Question	Group		P-value
	Action	Structure	
Q1	27/29	30/30	.24
Q2	28/29	29/30	1.00
Q3	25/29	27/30	.71
Q4	25/29	19/30	.07
Q5	28/29	29/30	1.00
Q6	25/29	28/30	.42
Q7	21/29	14/30	.06
Q8	22/29	14/30	.03
Q9	20/29	25/30	.23
Q10	26/29	28/30	.67
Q11	24/29	20/30	.23
Q12	23/29	22/29	.76
Q13	16/29	24/30	.05
Q14	15/29	15/30	1.00
Q15	23/29	25/30	.75
Q16	20/29	16/30	.29
Q17	19/29	14/30	.19
Q18	26/29	25/30	.71
Q19	6/29	12/30	.16
Q20	27/29	27/30	1.00

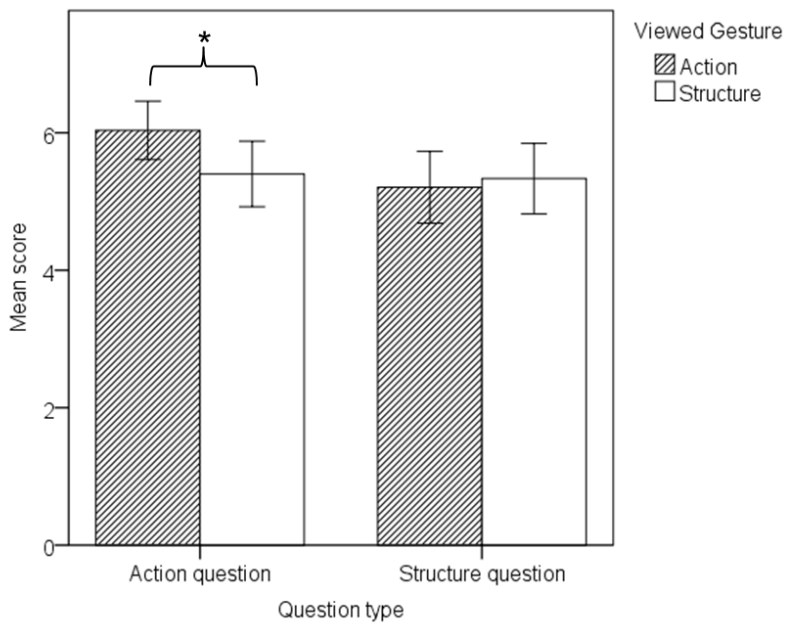
To observe the characteristics of a particular question and to ensure the questions in the knowledge test were of an appropriate standard, independent Chi-square test (Table 2) and item analysis were performed (See Appendix F for item analysis). Table 2 shows numbers of participants who got correct answer in each question item. There were significant group differences in the item 8 and marginally significant in the item 4, 7, and 13.

Also, it was found that among the question items, one from 8 action questions and one from 8 structure questions undermined the test reliability the most. The questions were *"A byproduct of air and fuel is pushed by a piston and go out through an exhaust port. (   )"* from action knowledge items and *"The piston is located closer to the crankshaft in the combustion phase than in the exhaust phase. (   )"* from structure knowledge items.

Table 3. Mean scores in the knowledge test after deleting two question items (SD)

	Action question [7]	Structure question [7]	General question [4]	Total [18]
Action group	6.03 (1.12)	5.21 (1.37)	2.69 (0.60)	13.93 (2.28)
Structure group	5.40 (1.28)	5.33 (1.37)	2.60 (0.97)	13.33 (2.80)

Therefore, after deleting those two question items from the scoring of the knowledge test, the two groups' performance was reanalyzed. In other words, the participants' mean scores for 7 action questions, 7 structure questions, and 4 general questions were compared. Table 3 shows mean scores that each group received in the action and structure questions.



* $p < .05$

Figure 3. Mean scores in the action and structure questions by the two groups. Error bars represent standard errors of the means.

No interaction was observed by group and question type between the two groups, $F(1,114) = 2.57$, $p = .11$. However, there was a group difference in the action question scores. The action group performed better on action question than the structure group, $F(1,57) = 4.12$, $MSE = 1.44$, $p < .05$, while there was no group difference in the structure questions (p = 73). Also, there was no group difference in total scores ($p = .37$). Figure 3 shows mean scores of the action and structure questions by the two groups in a bar graph.

Drawing analysis

In the drawing test, visual components expressed by the participants (29 in the action group, 30 in the structure group) were counted and compared. An inter-rater reliability analysis using the Kappa statistic was performed to determine consistency among the raters. The reliability for each visual component was found to be *Kappa* = .56 ($p < .001$) for action word, *Kappa* = .63 ($p < .001$) for action arrow, *Kappa* = .65 ($p < .001$) for action effect, *Kappa* = .60 ($p < .001$) for labeling arrow, and *Kappa* = .73 ($p < .001$) for labeling line. Therefore, substantial agreements were obtained in the scoring of the visual components.

For the data analysis, a Poisson regression analysis was administered to model count variables with the assumption that the conditional means equal the conditional variances[2]. Mean numbers of the visual components that the participants drew are shown in Table 4. First of all, there was a significant difference in the total number of the visual components drawn by the two groups. Overall, the action group drew more visual components than did the structure group on the drawing test (χ^2 (1, N = 59) = 7.26, $p < .05$).

---

[2] For reasons mentioned here, Poisson regression analyses were administered for count variables in gesture and speech analysis.

Table 4. Mean number of visual components used in the drawing test (SD)

Group	Visual component Action word	Action arrow*	Action effect*	Labeling arrow	Labeling line**	Total*
Action	5.34 (5.44)	7.48 (7.51)	2.28 (2.05)	1.76 (2.32)	1.52 (3.81)	18.38 (11.35)
Structure	4.37 (4.79)	5.77 (6.02)	1.37 (1.71)	1.50 (2.22)	2.77 (3.56)	15.77 (10.87)

* $p < .05$, ** $p < .01$

Also, there were significant differences in the number of action arrow (χ^2 (1, N = 59) = 6.90, $p < .05$), action effect (χ^2 (1, N = 59) = 1.99, $p < .05$) and labeling line (χ^2 (1, N = 59) = 7.04, $p < .01$) between the two groups. No differences were observed in the number of action word ($p = .09$) and labeling arrow ($p = .44$) between the two groups. The action group used more action arrows and action effects than the structure group. The structure group expressed more labeling lines than the action group. Figure 4-8 are histograms showing the numbers of the visual components used by the action group (n = 29) and the structure group (n = 30).

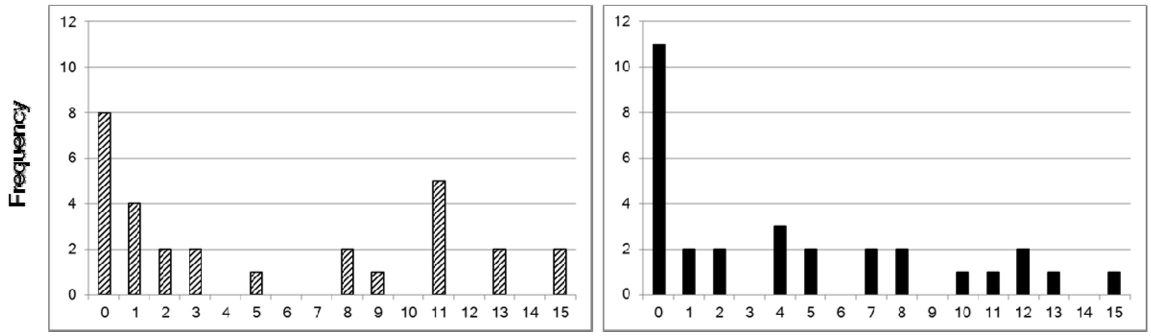


Figure 4. Number of participants that used action word: the left panel is for the action group and the right panel is for the structure group.

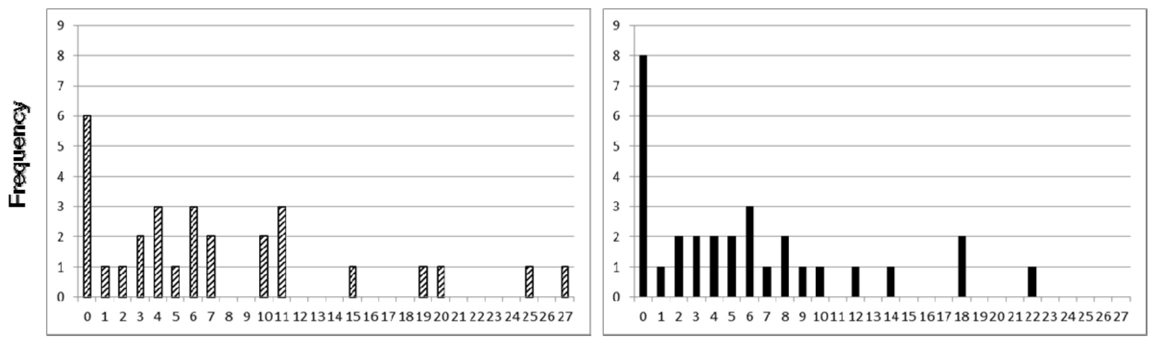


Figure 5. Number of participants that used action arrow: the left panel is for the action group and the right panel is for the structure group.

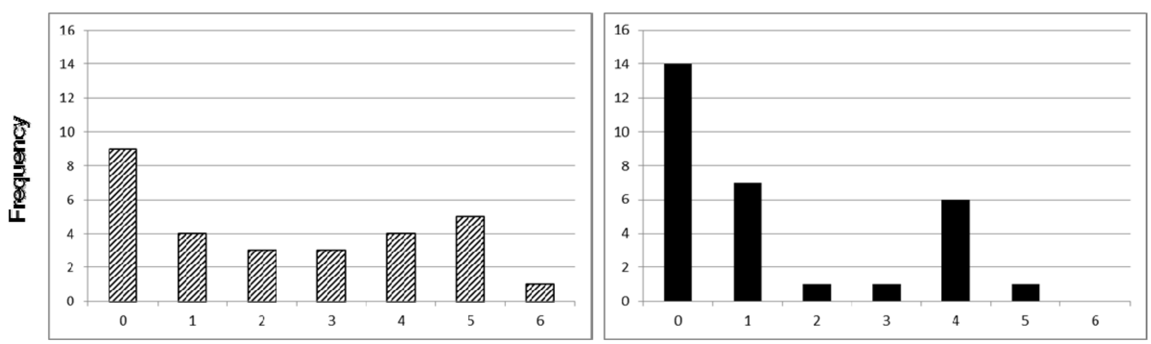


Figure 6. Number of participants that used action effect: the left panel is for the action group and the right panel is for the structure group.

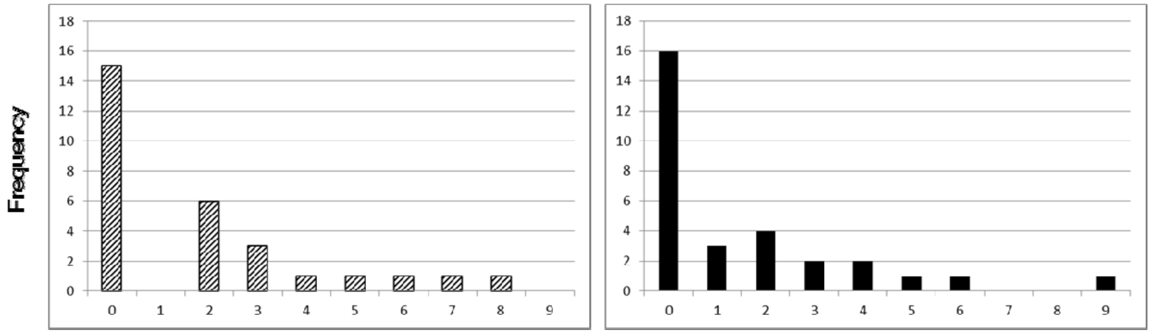


Figure 7. Number of participants that used labeling arrow: the left panel is for the action group and the right panel is for the structure group.

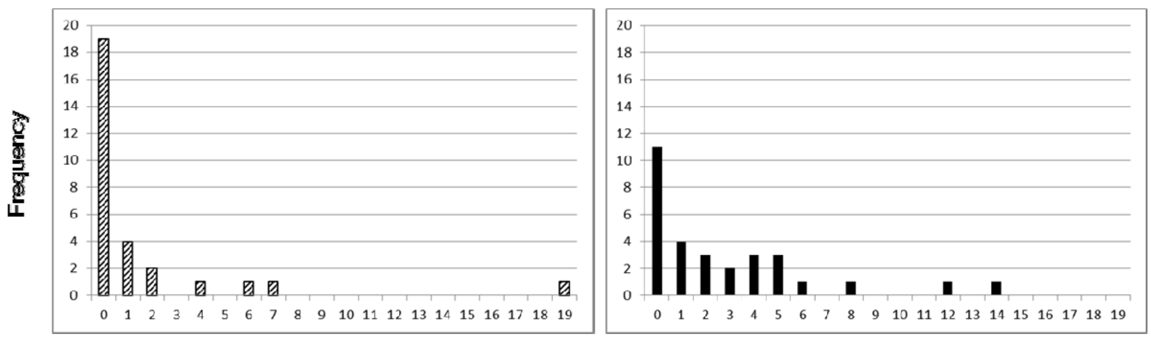


Figure 8. Number of participants that used labeling line: the left panel is for the action group and the right panel is for the structure group.

As an example, drawings from a participant in the action group and in the structure group are shown in Figure 9 and 10. As seen in Figure 9, a lot of action arrows and action effects were expressed, such as information of compression and inflation of a mixture of air and fuel, explosion from a spark plug and etc. On the other hand, Figure 10 shows the concept of how a four stroke engine works where the focus is on its structure with the use of labeling arrows. The results showed that action gestures were more involved in constructing action knowledge and mental models based on action rather than on structure. As a following analysis, the participants' gestures and speech used while they had explained were analyzed to verify if the participants' knowledge measured in the knowledge test and the mental model represented in the drawing test corresponds to what the participants were actually harboring.

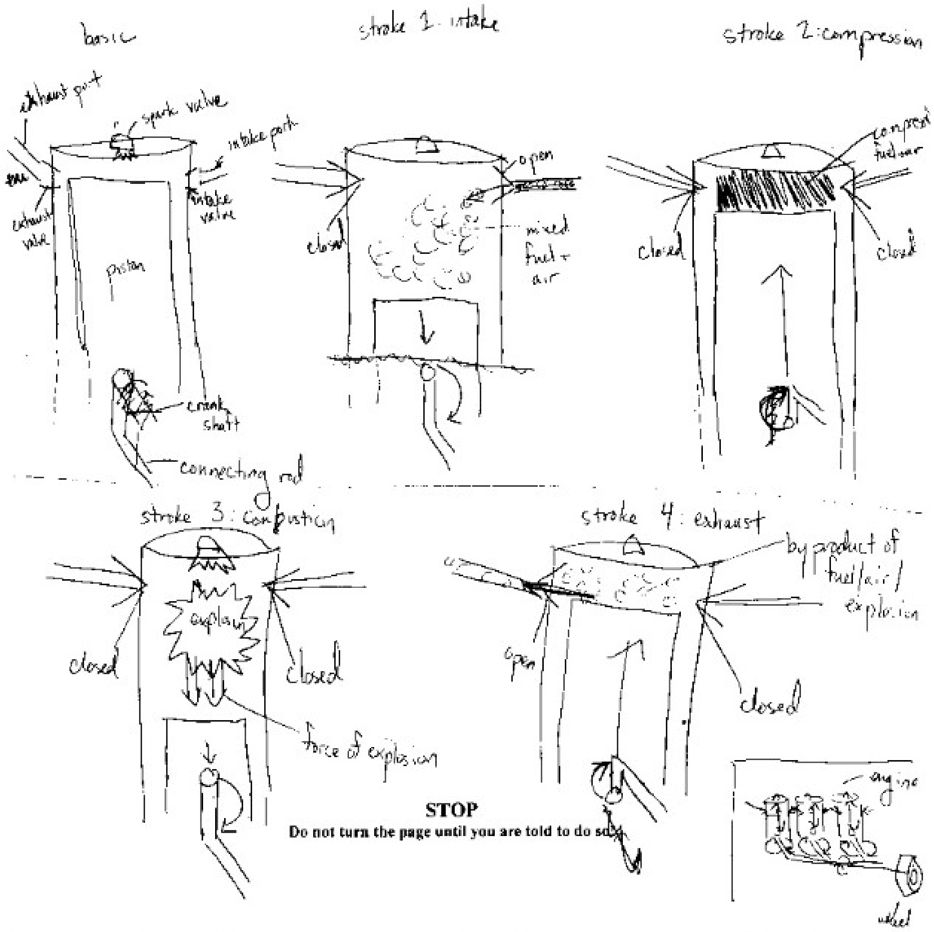


Figure 9. Drawing example from a participant in the action group

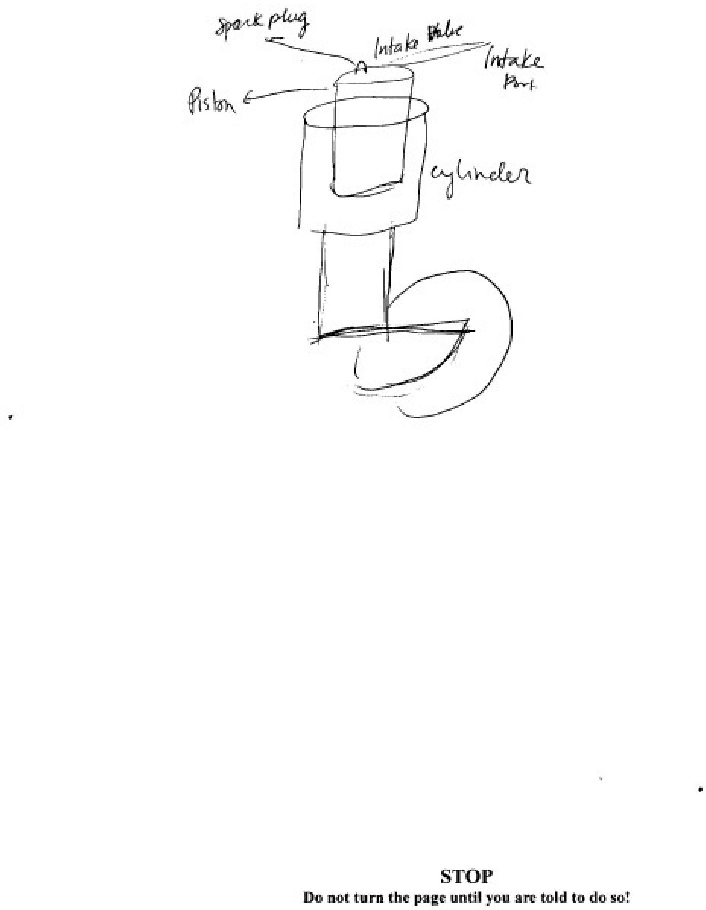


Figure 10. Drawing example from a participant in the structure group

Complexity in the drawings

The participants' drawings were further analyzed to see if there was a difference in complexity of knowledge that they acquired of the given concept. The instructional video contained the explanation of one complete cycle of the four stroke engine that was composed of 4 steps that finally make a car move. In this analysis, with either text or diagram, if a participant did not show all four engine steps in the participant's drawing it was coded as an incomplete step, and if a participant delineated a complete four stroke process it was coded as a complete step. Also, just writing the names of the four steps was not regarded as delineating complete four steps. Therefore, if a participant explained or drew all four steps of an engine work, he is regarded as having complex knowledge of the concept. Table 5 shows Crosstabulation between group and step of the complexity in the participants' drawings.

Table 5. Crosstabulation between group and step used in the drawing

		Step		Total
		Incomplete	Complete	
Group	Action	4	25	29
	Structure	11	19	30
	Total	15	44	59

In the action group, 4 participants delivered the given concept in a single step while 25 participants delivered the given concept in a complete step. In the structure group 11 participants delivered the given concept in a single step and 19 participants delivered the concept in a complete step. The percentage of participants that showed complexity in their drawing differed by group, χ^2 (1, N = 59) = 4.07, $p < .05$.

Gesture analysis

The participants' video recordings were analyzed[3]. Two participants from the action group and three from the structure group never used their hands. However, they were included in calculating an average gesture use, because using no gestures was also one of behavioral patterns that the participants showed. The participants in the both groups delivered a total of 1,350 gestures while they were explaining the concept to their peer: the action group delivered a total of 770 gestures while the structure group delivered a total of 580 gestures. There was no difference in the explanation time (p = .10). Poisson regression was used for a group comparison on total gesture use. The action group used more gestures than the structure group in the mean numbers of a total gesture usage (χ^2 (1, N = 58) = 13.34, p < .01). Table 6 shows the averages of explanation time for the two groups and the mean numbers of gestures that the participants used in each group.

Table 6. Mean number of explanation time and gesture usage by type of viewed gesture (SD)

Group	Average of Explanation time	Mean number of action gesture	Mean number of structure gesture	Mean number of total gesture
Action	177.14 (56.84)	21.62 (15.21)	4.93 (5.28)	26.55 (19.09)
Structure	152.34 (55.94)	12.90 (11.79)	7.10 (6.67)	20.00 (16.09)

[3] One participant's explanation in the structure group was not recorded because of equipment malfunctioning. Therefore, a total of 58 participants' videos (29 in the action group, 29 in the structure group) were analyzed.

Poisson regression analysis was administered to observe an interaction between gestures viewed and gestures used. It was found that there was an interaction between group and gesture type in the number of gesture use, χ^2 (1, *N* = 58) = 2.63, $p < .01$. The action group used action gestures more frequently and the structure group used relatively more structure gestures.

In within group comparison, a paired sample t-test was performed. Even though the participants in the both groups delivered more action gestures than structure gestures, the action group ($t(28) = 7.25$, $p < .0001$, $d = 1.47$, $r = .59$) reliably used more action gestures than structure gestures when compared to the structure group ($t(28) = 3.00$, $p = .006 < .01$, $d = .61$, $r = .29$). This was observed through a comparison of the significance level and the effect size between the two groups. The action group showed a much higher significance level when comparing the use of action and structure gestures. Also, the effect size in the action group is much stronger than the one in the structure group.

It is also possible that the number and pattern of gestures differed by the length of an explanation. Although there were no significant differences in the explanation time and in the overall gesture usage, explanations in the action group were longer arithmetically. To avoid any influence from the explanation time and to make the data analysis more reliable, the same data were analyzed based on the rate of gestures per minute. In other words, further analyses were proceeded to present a more detailed analysis of the results by gesture rate. By analyzing gestures based on a time unit, it was assumed that more close observation was possible of how a speaker's specific gesture influenced listeners' knowledge representation. Table 7 shows mean numbers of gesture per minute in the two groups.

Table 7. Mean number of gesture use per minute by type of viewed gesture

Group	Mean number of action gesture per min	Mean number of structure gesture per min
Action	7.00 (4.15)	1.51 (1.41)
Structure	4.87 (3.55)	2.62 (2.41)

For comparison of the two groups' gestures use per minute, General Linear Model analysis was administered. The same pattern of gesture use was observed even when a time factor was considered. It was found that there was an interaction between the two groups in gesture use such that the action group used more action gestures and the structure group used relatively more structure gestures, $F(1,56) = 11.13$, $MSE = 6.81$, $p = .002$, $< .01$, $\delta = .91$. Figure 11 shows gesture usage by the two groups in a bar graph.

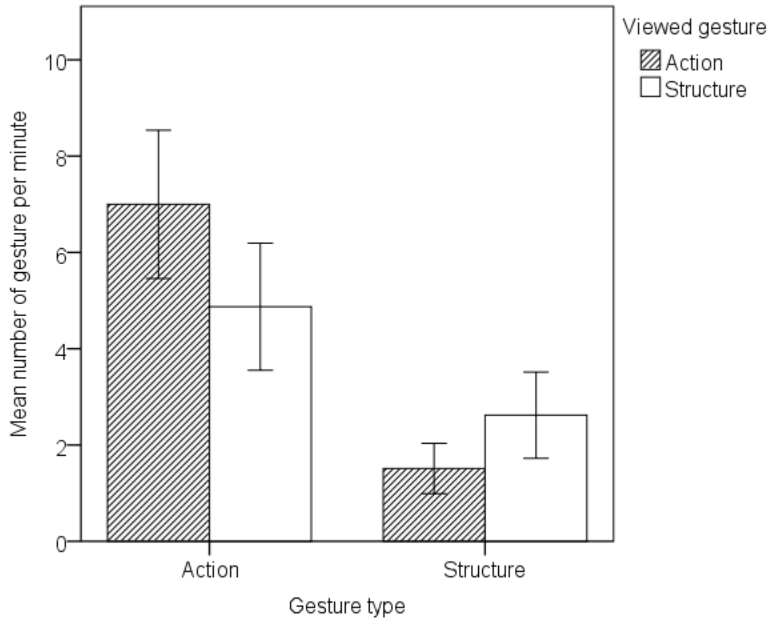


Figure 11. Mean number of gesture per minute by type of viewed gesture. Error bars represent standard errors of the means.

In within group comparison, when compared to the structure group (t(28) = 3.27, p = .003 < .01, d = .74, r = .35), the action group (t(28) = 8.14, p < .0001, d = 1.77, r = .66) reliably used more action gestures than structure gestures. This was also true from the significance level and effect size: the action group showed a much higher significance level and a bigger effect size when compared to the structure group of action gestures and structure gestures usage. Figure 12 is snapshots of participants' gestures during their explanation.

Figure 12. Video snapshots of participants' explanation. The left panel is from the action gesture group and the right panel is from the structure gesture group.

Gesture shape analysis

Participants' gestures were further analyzed to find out if a speaker's gestures differently affected the participants' knowledge representation. In this analysis, I focused on whether the participants' gestures were invented or imitated based on the instructional video that they had watched (see Appendix G and H for gestures used in the instructional videos). If a participant's gesture was not shown in each instructional video, it was coded as *invented*. In other words, any structural gestures used by the action group and any action gestures used by the structure group were coded as *invented*. Also, any additional action gestures used by the group that viewed action or additional structural gestures viewed by the group that viewed structure were regarded as *invented*. If a participants' gesture was one that was shown in the instructional video, it was coded as *imitated*. Participants' gesture uses by its shape are shown in Table 8.

Table 8. Participants' gesture use in its shape

Group	Mean number of invented gesture	Mean number of imitated gesture	Mean number of invented gesture per min	Mean umber of imitated gesture per min
Action	19.97 (14.96)	6.76 (5.74)	6.43 (4.18)	2.14 (1.74)
Structure	15.76 (14.58)	4.38 (5.93)	5.80 (4.10)	1.76 (2.52)

No interaction was found between group and gesture type in gesture use (p = .62). However, across the groups the participants used more *invented* gestures than *imitated* gestures, χ^2 (1, N = 58) = 9.26, p < .01. Also, the action group used more *invented* gestures (χ^2 (1, N = 58) = 12.35, p < .01) and *imitated* gesture (χ^2 (1, N = 58) = 6.45, p < .01) than the structure group.

The similar patterns were observed when a time factor was considered. There was no interaction between group and gesture type in gesture use per minute (p = .83). However, across the groups the participants used more *invented* gestures than *imitated* gestures, F(1, 114) = 46.64, MSE =10.79, p < .01. But, there was no group difference in *invented* gesture use per minute (p = .57) and *imitated* gesture use per minute (p = .51).

Speech analysis

The participants delivered a total of 2,550 information units in their speech. Among them, 1,607 were *action* information, 737 were *structure* information, and 206 *other* information. The action group delivered a total of 1,425 information units. Among them, 929 were *action* information, 387 were *structure* information, and 109 were *other* information. The structure group delivered a total of 1,125 information units. Among them, 678 were *action* information, 350 were *structure* information, and 97 were *other* information. Figure 13 shows the mean number of information units delivered by the two groups.

Poisson regression analysis was performed. There was an interaction between group and information type in the number of information unit, χ^2 (1, N = 58) = 5.27, $p < .01$. There was a difference in the number of information units among the information types, χ^2 (1, N = 58) = 4.75, $p = .0001 < .01$. Also, it was found that more *action* information units were used than the *structure* and *other* information units ($p < .0001$), and more *structure* information units were used than the *other* information units ($p < .0001$).

ANOVA tests were administered for a group comparison. Overall, the action group (M = 49.14, SD = 20.81) delivered more information units than the structure group (M = 38.79, SD = 17.22), $F(1, 56) = 4.25$, MSE = 364.75 , $p = .04 < .05$. Also, the action group (M = 32.03, SD = 12.89) delivered more *action* information than the structure group (M = 23.38, SD = 12.24), $F(1, 56) = 6.87$, MSE = 158.03, $p = .01 < .05$. There was no group difference in the number of *structure* information units between the action group (M = 13.34, SD = 7.82) and the structure group (*M* = 12.07, *SD* = 7.16), *p* = .52 and there was no group difference in the number of *other* information used between the action group (*M* = 3.76, *SD* = 4.09) and the structure group (*M* = 3.34, *SD* = 3.00), *p* = .66.

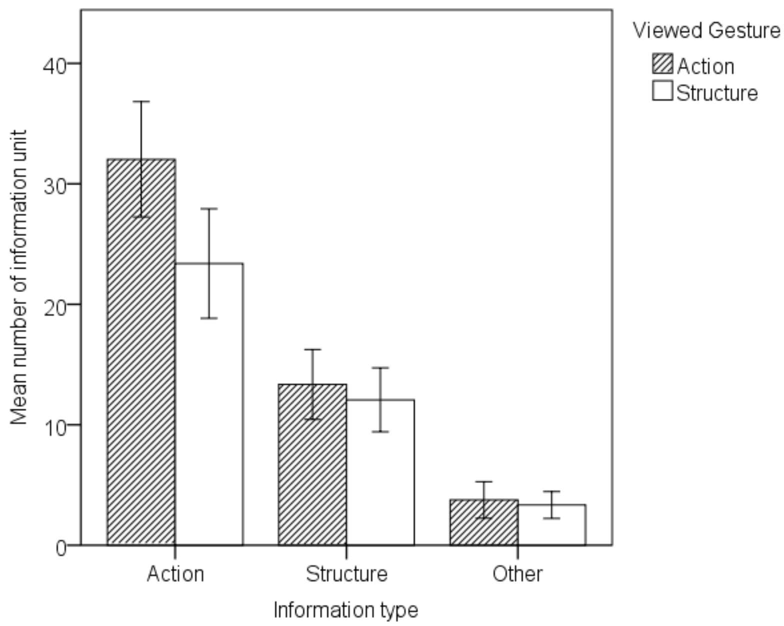


Figure 13. Mean number of information units by information type in the two groups. Error bars represent standard errors of the means.

To closely look into the relationship between the gestures' role and the information type, UNIVARIATE ANOVA test was administered without *other* information units, that is, in this analysis, only the numbers of the action and structure information units were considered, since this study is to observe the relationship between gesture' role and its representation of the action and structure knowledge. No interaction was found in the number of speech information units between group and information type (p = .06).

*Proportion of information type in speech***.** Remind that although there was no group difference in the explanation time between the two groups, overall the action group delivered more information units than the structure group. Also, no interaction was found in the number of information units between group and information type, only when the action and structure information were considered. To closely look at the relation between information type and gesture's role, the ratios of each information type in the participants' speech were calculated and compared, that is, the proportion of *action* information, *structure* information and *other* information. Figure 14 shows the mean percentages of each information type used by the two groups. For the action group, *action* information accounted for an average of 66.62% (*SD* = 10.30), *structure* information accounted for an average of 25.76% (*SD* = 10.13), and *other* information accounted for an average of 7.61% (*SD* = 7.30) of their total explanation.

For the structure group, among their explanation, an average of 59.28% (*SD* = 15.89) was *action* information, 31.59% (*SD* = 13.28) was *structure* information, and 9.14% (*SD* = 8.34) was *other* information.

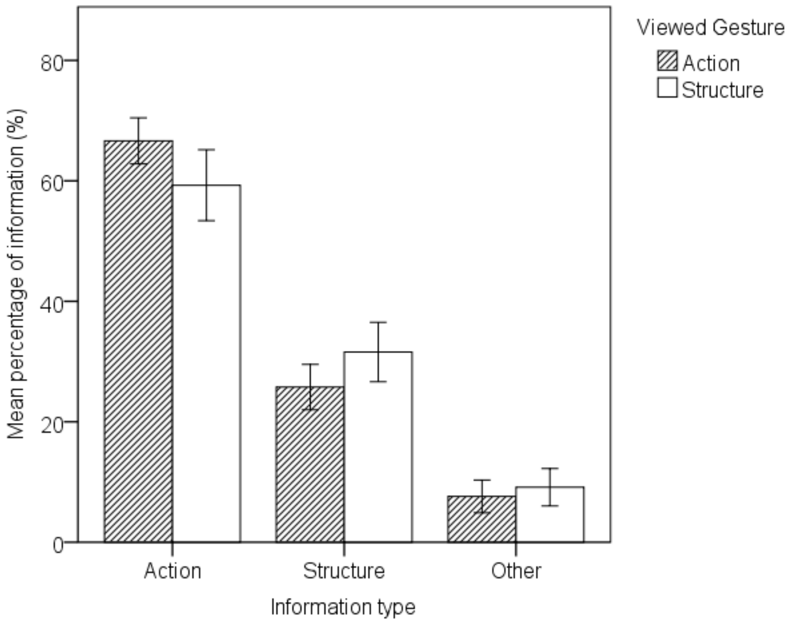


Figure 14. Mean percentage of information units by the two groups. Error bars represent standard errors of the means.

In UNIVARIATE ANOVA test, there was an interaction between group and information type in the mean percentage of information unit, $F(2, 168) = 5.16$, $MSE = 126.74$, $p = .007 < .01$, $\delta = .82$. Also, there was a difference on mean percentages depending on the information type, $F(2, 168) = 348.23$, $p = .0001 < .01$. Also, the same patterns were observed in post hoc tests (*Tukey HSD*); more *action* information units were used than the *structure* and *other* information units ($p < .0001$), and more *structure* information units were used than the *other* information units ($p < .0001$).

The action group assigned a larger portion on *action* information while they were explaining than was the structure group, $F(1, 56) = 4.37$, $P = .04 < .05$. However, there were no differences on the information type between the two groups in the proportions of the *structure* information units ($p = .07$) and the *other* information units ($p = .46$).

For the same reason mentioned earlier, UNIVARIATE ANOVA test was administered again considering only the action and structure knowledge. There was an interaction in the proportion of information units between group and information type, $F(1, 112) = 7.89$, $MSE = 159.40$, $p = .006 < .01$.

Summary of results

Although no interaction and no group difference were found in the knowledge test, based on item analysis, it was found that there were two items, one from action and the other from structure question items, that undermined the test reliability the most. After deleting these two items, the two groups' test scores were compared again. It was found that there was a significant difference in action knowledge scores between the action group and the structure group. The action group received a higher score on action questions than the structure group. In the drawing test, the visual components that the participants expressed were counted and compared. The action group showed more action arrows and more action effects than the structure group. On the other hand, the structure group drew more labeling lines than the action group. In addition, the participants' gestures and speech during their explanation were analyzed. There was an interaction between the two groups in gesture use. The action group used more action gestures and the structure group used relatively more structure gestures. In speech analyses, there was an interaction between the two groups in the usage of the information unit. The action group delivered relatively more *action* information than the structure group while the structure group delivered relatively more *structure* information than the action group.

CHAPTER V

DISCUSSION

This study was to investigate if and how a specific type of gesture influences knowledge construction and representation. There are abundant evidence that gestures provide a rich source of information, including information about structure and process (Beattie et al., 1999; Becvar, Hollan, & Hutchins, 2008). Here, I asked if gestures can convey action or structure, over and above verbal and diagrammatic explanations, simply and abstractly at a pace that allows comprehension.

Did a specific type of gesture prime listeners' construction of a specific type of knowledge and mental representation, and which led the listeners to actual learning of a concept?

In the knowledge test, no interaction was found in the scores of action and structure question between group and question type. Also, there were no group differences in total mean scores of the knowledge test and no group differences in action question, structure question and general question. However, in within group comparisons, the action group received a higher score on action question than structure question. There was no score difference between action questions and structure questions within the structure group. It makes sense that there was no difference dependent on viewed gesture because the knowledge test was created solely based on the verbal script. Also, because the diagram showing the structure was always in view and because the introduction to both explanations overviewed the system structure, it was assumed that the gestures emphasizing structure served as a control and did not affect performance on the questions.

To closely look into the characteristics of a particular question, item analysis was performed. It was found that there was one item from each of the action questions and structure questions which undermined the reliability of the knowledge test the most. The question was *"A byproduct of air and fuel is pushed by a piston and go out through an exhaust port. (　　)"* from action question items and *"The piston is located closer to the crankshaft in the combustion phase than in the exhaust phase. (　　)"* from structure question items. First, it seems that the item from the action questions not only asked the movement pattern of the piston, but also required positional knowledge of the exhaust port which was related to the result of the piston's movement. Therefore, to answer this question, participants had to retrieve both the structure knowledge of an exhaust port and action knowledge and the causal relation between the piston's movement and its result. It was assumed that this type of question required the participants to activate both action knowledge and structure knowledge of the given concept, which would undermine the test reliability.

The participants would feel difficulty for a similar reason on the question, *"The piston is located closer to the crankshaft in the combustion phase than in the exhaust phase. (   )"* To answer this question, the participants had to retrieve a relative position of each part from two different stages and compare them. Even though the question was created based on the speaker's verbal script, targeted on checking structure knowledge, the participants probably were required to construct action information of the positional change of the piston. It is possible that the participants had to retrieve two different relative positions of the piston to the crankshaft and compare them, which requires knowledge of the change of the piston over time. This process involves simulated action. Since the movements and the effect of the piston were under an action schema, it would be possible that the participants were required to fill out the gap in between the movement of the piston and its effect in the cylinder, which required both action and structure knowledge.

Based on the assumption above, those two question items from the knowledge test were deleted and the two groups' scores in the knowledge test were compared again. Even though no interaction was found by group and question type between the two groups and there were no group difference in structure question, the action group performed better on action question than the structure group.

Also, in the analyses of participants' drawings, it was found that the action group expressed more visual components overall than the structure group. Since a schematic diagram was placed by the speaker in both instructional videos and the overview of the structure of the cylinder was explained in the introduction part in both videos, it was reasonable that the action group drew more visual components than the structure group. The action group could encode both action and structure related components while watching the video and retrieved those in the drawing test. More specifically, the action group drew significantly more action arrows and action effects. Since action arrows not only deliver information of movement, but also represent the direction of the parts and the size of the strength of the movement, it is assumed that participants in the action group were primed by the speaker's action gestures and represented a given concept based on action knowledge. In other words, action gestures activated action representation of a given concept. Also, the action group drew more action effects. For example, they drew a star-shaped figure placed below the spark plug expressing the explosion of a mixture of air and fuel by ignition of the spark plug or several wavy lines between the spark plug and the piston to show compressed air and fuel within the cylinder. It was assumed that these effects were more activated when action-based representation of the concept was based.

On the other hand, the structure group drew more labeling lines than the action group. Labeling lines were used to name each part in a four stroke engine. This analysis confirmed that the action knowledge that the action group built was actually derived from action-based mental representation and the structure knowledge that the structure group built was derived from structure-based mental representation.

In addition to the group differences in the visual components, more participants in the action group showed that they had complex knowledge of the given concept compared to the participants in the structure group. More participants in the action group delineated their drawing of a four stroke engine process in a complete four step while a relatively small number of participants in the structure group showed a complete cycle of the given concept. The drawings by the action group were richer than the ones by the structure group. It was assumed that the action group developed more complex and detailed knowledge of the given concept from action gestures.

If so, did the listeners actually construct knowledge based on the representation that is grounded on the speaker's gestures?

To see if the specific type of knowledge and its representation were derived from the same prototype to what the participants had, the participants' gestures were analyzed. Overall, the action group used more action gestures and the structure group used relatively more structure gestures. Although there were no significant differences in the participants' explanation time and in their overall gesture use, the explanations in the action group were longer than the ones in the structure group. To avoid any influence from explanation time and to make the data analysis more reliable, the same data were analyzed based on the rate of gestures per minute. In the result, the same pattern of gesture use was observed even when a time factor was considered. There was an interaction between the two groups in gesture use. The action group used more action gestures and the structure group used relatively more structure gestures. Again, the study results showed that a speaker's action gestures were more dedicated to activating action knowledge and a speaker's structure gestures were engaged more in activating structure-based knowledge of the concept. Also, it was observed that the knowledge expressed through the drawings and delivered through gestures was based on same mental representation that was activated by the speaker's gestures in the instructional video.

The participants' gestures were further analyzed to see if the gestures they used were invented or imitated compared to the speaker's gestures in the instructional videos. There was no group difference in the number of invented gestures and imitated gestures, and the both groups used more invented gestures than imitated gestures. This result implies that a speaker's gestures affect listeners' way of thinking, not producing a same gesture shape.

Did the usage of information in participants' speech correspond to the results from the posttests and gesture analysis?

There was an interaction between group and information type in the participants' speech. The action group used relatively more action information units than the structure group and the structure group delivered more structure information units than the action group. The representation of the concept that the participants expressed in speech again corresponded to the type of knowledge that was delivered in gestures: the action group used more action gestures than the structure group and the structure group used relatively more structure gestures than the action group. This finally implies that gesture and speech correspond to each other in delivering certain concepts and they share the same cognitive architecture in knowledge representation.

In terms of the relation between gestures and the representation embedded in them, it was argued that representational gestures originate from spatio-motoric representations which contain spatial information and information about actions (Kita & Özyürek, 2003). Given that both action and structure gestures are included in a category of representational gesture (Holler & Beattie, 2003), and the current study showed the different roles of action gesture and structure gesture in knowledge representation, their argument can be further specified and modified.

In sum, action gestures involved in listeners' constructing knowledge and mental representations are based on action, and structure gestures engaged in listeners' constructing knowledge and mental representation are based on structure. Learners embodied an attribute in the speaker's gestures and used this for representing and constructing knowledge of the concept.

This implies that listeners' knowledge is grounded in a speaker's gestures and this relation is specialized depending on gesture type.

Knowledge transfer from a speaker to a listener

The study results showed that knowledge transferred from a speaker is represented differently depending on its semantic properties. In other words, action knowledge coming from a speaker's hands and speech primes a listener's representation of action while knowledge of image or shape primes a listener's representation of structure. Based on the results from this study, Figure 15 shows schematic flows of knowledge from a speaker's hand(s) and speech to a listener whose knowledge again is expressed via the listener's hands and speech.

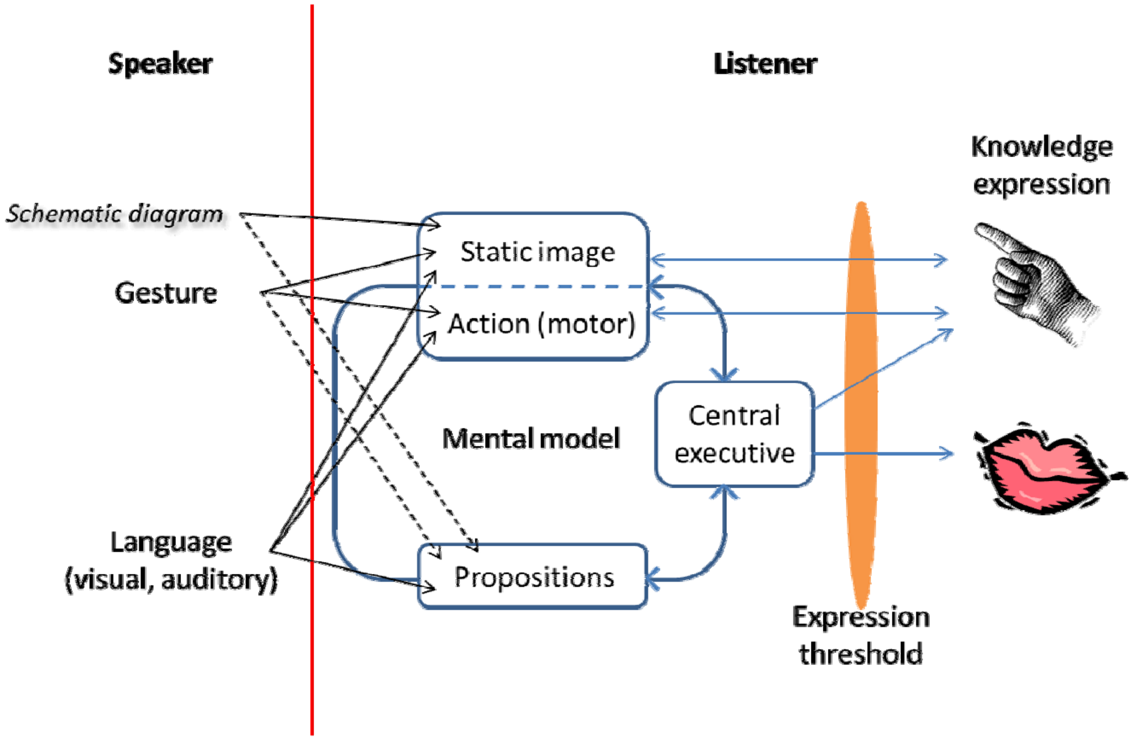


Figure 15. Knowledge transfer from a speaker to a listener in the explanation situation (modified based on Hostetter and Alibali, 2008)

Information from language (text and a speaker's voice) is processed and encoded into either perceptual symbols (Glenberg & Robertson, 2000) or propositions (Bansford & Franks, 1971; Kintsch, 1974). At the same time, information from a speaker's gestures is mostly perceptually symbolized, since viewing a particular concept can trigger simulating a perceptual and motor representation (Barsalou, 1999). In addition, this study does not exclude the possibility that information from a speaker's gestures is encoded into propositions, because even though it is not as clearly as language is processed, information in gestures can also be represented as propositions consisting of relations and arguments. Then the information from speech and gesture is controlled by the central executive which communicates with perception-based representations (modality based) and meaning-based knowledge representations (propositional networks). This finally comprises a listener's mental model, and during this process gestures can lighten cognitive load which is imposed by propositional representation (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001).

Once representation is activated strongly enough to be lifted up above the threshold level, it comes out and is expressed through hands or an articulator. Also, implicit knowledge which is not processed at a conscious level also can be expressed through hands. According to studies investigating information within gestures, information conveyed by learners' gesture cannot be found anywhere in their speech (Bavelas, 1994; Church & Goldin-Meadow, 1986; Garber, Alibali, & Goldin-Meadow, 1998), and it is assumed that spontaneous gestures reveal implicit knowledge which is not exposed in speech (Goldin-Meadow & Alibali, 1999). At the same time, action induced by hands directly primes or activates mental representations through which learners can also embody its corresponding knowledge (Beilock & Goldin-Meadow, 2010).

In the Figure 15, dotted lines in the square box imply that information from gesture and speech is represented differently based on its semantic property. If information has action property, it is encoded as action or motoric representation while if information has spatial but no action property it is encoded as a visuo-spatial representation. However, this does not mean there is a clear distinction between image representation and action information since both types of information share a common feature (i.e., spatial property). Therefore, the static image and the action-motoric representation communicate at some point. Throughout these processes, a listener encodes a speaker's verbal and nonverbal information and constructs knowledge corresponding to the semantic property that incoming information bears.

Significance and future direction

Action and structure distinction generalizes to many areas, from biology to dynamic within an organization to mechanics. Both types of knowledge were closely intertwined to complete knowledge of a complex system. However, compared to structural information, it is hard to deliver action information purely with speech. It could be a movement pattern that is hard to delineate only with speech. Animation can be used but it is not common in most conversational and instructional settings. Even though it has been known that a speaker's gestures help learning, we do not know what relation exists between a speaker's gestures and the type of knowledge listeners develop. Through this study, it was observed that a different gesture promotes or at least primes a specific type of knowledge and its representation. This implies that we can transfer information more efficiently by appropriately using action or structure gestures.

Action gestures promote action understanding and structure gestures affect reliably promoting structural knowledge. Teachers may strategically use a certain type of gesture that is more appropriate to derive a certain type of knowledge in teaching a concept in a classroom setting. In this sense, the result would provide a theoretical background for leading students to more effective learning. Also, this study shows that by using hands a speaker can tailor listeners' way of thinking.

As a future study, it would be interesting to investigate the possibility of developmental difference in constructing and representing knowledge from a speaker's different gestures. It has been found that children differently perceive and organize given information than adults. For example, children organize information less on the basis of a semantic theme and more on the basis of a perceptual organization, such as spatial position (Bruner, Oliver, & Greenfield, 1966). Also, children have less domain and relevant knowledge than adults (Chi, 1978; Stein & Trabasso, 1981) and children younger than twelve may perceive the visual world differently because the integration of perceptual information is not fully developed yet by that time (Nardini, M., Bedford, R., & Mareschal, D., 2010). Therefore, the role of gesture in children's learning of a certain concept may show a different pattern from the results of this study.

Limitations of the study

This study tried to show a different role of a different gesture in knowledge representation and construction. I chose a concept of the mechanism of a four stroke engine as a learning material and the concept has both action knowledge and structure knowledge embedded. Therefore, the learning material was assumed to be appropriate to trace the knowledge transfer from a speaker to a listener. Accordingly, two different types of gestures were created and embedded in the instructional videos. However, in creating video instructions of the mechanical system, it was almost impossible to make two different instructional videos mutually exclusive; one video solely about functional knowledge, the other solely about structural information. That can be a reason that no interaction was observed between the two groups in the knowledge test score or no group differences were found on the score of structure questions. Therefore, the conclusions obtained from this study will become stronger if this problem can be tested under a condition where two different types of knowledge are to be delivered separately throughout an instructional video.

It turned out from item analysis that the reliability of the knowledge test was not high enough to say that each question in the knowledge test measured the participants' knowledge of the given concept. Even though the knowledge test used in this study is not a standardized test, it is expected that the knowledge test would attain a certain level of reliability, so then we can say that the test properly measured listeners' knowledge of how a four stroke engine works.

Finally, in this study a gesture's effect was investigated in a mechanical concept: how a four stroke engine works. Therefore, to generalize the role of a speaker's gestures in listeners' constructing knowledge and mental representation, a follow up study should be examined with a different subject in other domains.

REFERENCES

Alibali, M., Bassok, M., Solomon, K., Syc, S., & 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., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language, 44*, 169-188.

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

Bach, P., & Tipper, S. P. (2006). Bend it like Beckham: Embodying the motor skills of famous athletes. *The Quarterly Journal of Experimental Psychology, 59,* 2033-2039.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences, 22*, 577-660.

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

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

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

Beilock, S. L., & Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *PsychologicalScience, 21*, 1605-1611.

Brass, M., Bekkering, H., & Prinz, W. (1999). Movement observation affects movement execution in a simple response task. *Acta Psychologica, 106*, 3-22.

Bruner, J. S., Olver, R. R., & Greenfield, P. (1966). Studies in cognitive growth. New York: Wiley & Sons.

Cassell, J., & McNeill, D. (1991). Gesture and the poetics of prose. *Poetics Today, 12*, 375-404.

Cassell, J., Stone, M., Douville, B., Prevost, S., Achorn, B., Steedman, M., Badler, N., & Pelachaud, C. (1994). "Modeling the Interaction between Speech and Gesture." in Ashwin Ram and Kurt Eiselt (eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, pp. 153-158. Lawrence Erlbaum Associates, Publishers.

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develop?* (pp.76-93). Hillsdale, NJ: Erlbaum.

Church, R. B., Ayman-Nolley, S., Estrada, J., Glover, D., & Dullum, T. (2001). What role can gesture play in teaching mathematical concepts? Paper presented at the Society for Research in Child Development, Minneapolis, MN.

Church, R. B., Ayman-Nolley, S., Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *Bilingual Education and Bilingualism, 7,* 303-319.

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

Cook, S. W., & Goldin-Meadow, S. (2006). The Role of Gesture in Learning: Do Children Use Their Hands to Change Their Minds? *Journal of Cognition and Development, 7,* 211-232.

Cook. S. W., & Tanenhaus, M. K. (2009). Embodied communication: Speaker' gestures affect listeners' actions. *Cognition, 133,* 98-104.

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

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

Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trends in Cognitive Science, 8*, 33-39.

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

Garber, P., Alibali M. W., & Goldin-Meadow, S. (1998). Knowledge conveyed in gesture is not tied to the hands. *Child Development, 69*, 75-84.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. New York : Houghton Mifflin.

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

Goldin-Meadow, S., & Alibali, M. W. (1999). Does the hand reflect implicit knowledge? Yes and no. *Behavioral and Brain Sciences, 22*, 766-767.

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

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

Goodwin, C. (1986). Gesture as a resource for the organization of mutual orientation. *Semiotica, 62*, 29-49.

Heath, C. C. (1992). Gesture's discrete tasks: Multiple relevancies in visual conduct in the contextualization of language. In P. Auer & A. di Luzio (Eds.), *The contextualization of language* (pp. 102-127). Amsterdam and Philadelphia: Benjamins.

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

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

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

Hostetter, A. B., & Alibali, M. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review, 15*, 459-514.

Jordan, J. S., & Knoblich, G. (2004). Spatial perception and control. *Psychonomic Bulletin and Review, 11*, 54-59.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillside, NJ: Erlbaum.

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

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

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

Melinger, A., & Levelt, W. J. M. (2004). Gesture and communicative intention of the speaker. *Gesture*, *4*, 119-141.

Morsella, E., & Krauss, R. (2004). The role of gestures in spatial working memory and speech. *American Journal of Psychology, 117*, 411-424.

Nardini, M., Bedford, R., & Mareschal, D. (2010). Fusion of visual cues is not mandatory in children. *Proceedings of the National Academy of Sciences, 107,* 17041–17046.

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

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

Riseborough, M. G. (1981). Physiographic gestures as decoding facilitators: Three experiments exploring a neglected facet of communication. *Journal of Nonverbal Behavior, 5*, 172-183.

Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules.

Cognitive Science, *20*, 457-497.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88*, B11-B21.

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

Stein, N. L., & Trabasso, T. (1981). What's in a story? Critical issues in comprehension and instruction. In R. Glaser (Ed.), *Advances in the psychology of instruction* (Vol. 2, pp.213-268). Hillsdale, NJ: Erlbaum.

Streeck, J. (1993). Gesture as communication I: Its coordination with gaze and speech. *Communication Monographs, 60*, 275-299.

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

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

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

Wesp, R., Hesse, J., Keutmann, D., & Wheaton, K. (2001). Gestures maintain spatial imagery.

American Journal of Psychology, 114, 591-600.

Woodall, W. G., & Folger, J. P. (1985). Nonverbal cue context and episodic memory: On the availability and endurance of nonverbal behaviors as retrieval cues. *Communication Monographs*, *52*, 319-333.

Appendix A: Script for an action gesture video

Action gesture video scrip

Today, I am going to explain how 4 stroke engine works. Almost all cars currently use what is called a four-stroke engine to convert gasoline into motion. A four stroke engine to the cycle refers to a series of processes including intake, compression, combustion and exhaust. The series of processes happen inside the cylinder, which includes the intake valve, the intake port, the spark plug, the exhaust valve, the exhaust port, the piston, the connecting rod, and the crankshaft.

Each cycle entails two rotations of the crankshaft for engines fueled by diesel or gasoline. Understanding the cycle's 4 strokes is a key to understanding how the engine type works. "4-strokes" to the cycle takes place continuously as the engine runs. In practice, this cycle happens one after the next in every cylinder of the engine.

Intake
The beginning of the cycle starts at what is known as top dead center. The first half **rotation** of the crankshaft pulls the piston **downward** inside the cylinder, **reducing** pressure inside. As the piston **descends**, the intake valve is **pulled open**, **letting in** a mixture of forced fuel and air.

Compression
The second half **rotation** of the crankshaft pushes the piston **back up** again inside the cylinder, **compressing** the fuel and air mixture as the intake valve **closes**.

Combustion
The third half **rotation** of the crankshaft is known as the combustion stroke. At the end of the compression stroke, a spark plug **ignites** the combustible mixture of fuel and air. This small **explosion** pushes the piston **downward** again in the cylinder through its power stroke.

Exhaust
The final stroke is known as the exhaust stroke. After the power stroke, the last half **rotation** of the crankshaft pushes the piston **upwards** in the cylinder for a second time, **expelling** the byproduct of the fuel and air combustion. As the crankshaft pushes the piston **up**, the exhaust valve **opens**, allowing the byproduct **to go out**.

In this process, the **linear motion** of the piston is converted into **rotational** motion by the crankshaft and this rotational motion is then used to rotate the car's wheels.

Appendix B: Script for a structure gesture video

Structural gesture video scrip

Today, I am going to explain how 4 stroke engine works. Almost all cars currently use what is called a four-stroke engine to convert gasoline into motion. A four stroke engine to the cycle refers to a series of processes including intake, compression, combustion and exhaust. The series of processes happen inside the cylinder, which includes the intake valve, the intake port, the spark plug, the exhaust valve, the exhaust port, the piston, the connecting rod, and the crankshaft.

Each cycle entails two rotations of the crankshaft for engines fueled by diesel or gasoline. Understanding the cycle's 4 strokes is a key to understanding how the engine type works. "4-strokes" to the cycle takes place continuously as the engine runs. In practice, this cycle happens one after the next in every cylinder of the engine.

Intake
The beginning of the cycle starts at what is known as top dead center. The first half rotation of the **crankshaft** pulls the **piston** downward inside the **cylinder**, reducing pressure inside. As the **piston** descends, the **intake valve** is pulled open, letting in **a mixture** of forced fuel and air.

Compression
The second half rotation of the **crankshaft** pushes the **piston** back up again inside the **cylinder**, compressing the fuel and air mixture as the **intake valve** closes.

Combustion
The third half rotation of the **crankshaft** is known as the combustion stroke. At the end of the compression stroke, a **spark plug** ignites the combustible **mixture** of fuel and air. This small explosion pushes the **piston** downward again in the **cylinder** through its power stroke.

Exhaust
The final stroke is known as the exhaust stroke. After the power stroke, the last half rotation of the **crankshaft** pushes the **piston** upwards in the **cylinder** for a second time, expelling the byproduct of the fuel and air combustion. As the crankshaft pushes the **piston** up, the **exhaust valve** opens, allowing the byproduct to go out.

In this process, the linear motion of the **piston** is converted into rotational motion by the **crankshaft** and this rotational motion is then used to rotate the car's wheels.

Appendix C: Information contained in action gestures

	Action gesture	
	accompanied word	**Information in gesture**
1	rotation	rotational movement of a crankshaft
2	downward	direction of piston movement
3	reducing	reduced pressure caused by piston movement
4	descends	direction of piston movement
5	pulled open	movement of an intake valve
6	letting in	flow of fuel & air
7	rotation	rotational movement of a crankshaft
8	back up	direction of piston movement
9	compressing	compressed fuel & air caused by piston movement
10	closes	movement of an intake valve
11	rotation	rotational movement of a crankshaft
12	ingnites	ignition of a spark plug
13	explosion	cause of pistion movement
14	downward	direction of piston movement
15	rotation	rotational movement of a crankshaft
16	upwards	direction of piston movement
17	expelling	flow of byproduct
18	up	direction of piston movement
19	opens	movement of an exhaust valve
20	to go out	flow of byproduct
21	linear motion	directional information of piston movement
22	rotational	rotational movement of a crankshaft

Appendix D: Information contained in structure gestures

	Structural gesture	
	accompanied word	**Information in gesture**
1	crankshaft	position and shape of a crankshaft
2	piston	position and shape of a piston
3	cylinder	relative poisition of piston & shape of a cylinder
4	piston	position and shape of a piston
5	intake valve	position of an intake valve & an intake port
6	mixture	place information which fuel & air come in
7	crankshaft	position and shape of a crankshaft
8	piston	position and shape of a piston
9	cylinder	relative poisition of piston & shape of a cylinder
10	intake valve	position of an intake value & an intake port
11	crankshaft	position and shape of a crankshaft
12	spark plug	position of a spark plug
13	mixture	relative poisition of mixture of fuel & air in a cylinder
14	piston	position and shape of a piston
15	cylinder	relative position of a piston & shape of a cylinder
16	crankshaft	position and shape of a crankshaft
17	piston	position and shape of a piston
18	cylinder	relative position of a piston & shape of a cylinder
19	piston	position and shape of a piston
20	exhaust valve	relative position of an exhaust valve & an exhaust port
21	piston	position and shape of a piston
22	crankshaft	position and shape of a crankshaft

Appendix E: Posttest questionnaire

*** Please choose a correct answer, or mark each of these statements are true (T) or false (F).**

When intake valve is pulled open, air and fuel move inside the cylinder. ( )

The crankshaft is attached to the wall of the cylinder. ( )

The piston is above the cylinder. ( )

The byproduct enters the cylinder, when the crankshaft pushes the piston downward. ( )

The exhaust valve is located between the spark plug and the intake valve. ( )

In the exhaust phase, the piston moves upwards by rotation of the crankshaft. ( )

In the compression phase the piston is located closer to the intake valve than in the intake phase. ( )

From where does a mixture of air and fuel enter the cylinder? ( )
1) intake valve 2) intake port 3) cylinder 4) exhaust valve

In the combustion phase, the mixture of fuel and air expands when the piston is pushed up within the cylinder. ( )

How many spark plugs does the cylinder have in one cycle? ( )
1) one 2) two 3) three 4) four

The piston is located closer to the crankshaft in the combustion phase than in the exhaust phase. ( )

How many times does the crankshaft rotate in each cycle? ( )
1) one 2) two 3) three 4) four

The piston is attached to the wall of the cylinder. ( )

A byproduct of air and fuel is pushed by a piston and go out through a exhaust port. ( )

The linear motion of the crankshaft makes the car's wheels rotate. ( )

A mixture of air and fuel is located above the piston. ( )

In the intake phase, the rotational motion of the crankshaft pulls the piston downward. ( )

The piston is below the crankshaft. ( )

In which stage does the engine get most of its power? ( )
1) intake 2) combustion 3) compression 4) exhaust

In the compression phase, the piston is pushed back up again inside the cylinder. ( )

STOP

Do not turn the page until you are told to do so!

*** Based on the video you watched, draw a diagram showing how a 4 stroke engine works as detailed as possible.**

STOP

Do not turn the page until you are told to do so!

Gender:

Age:

Have you ever learned or studied about 4 stroke engine either at school or via media?
(Yes / No)

If yes, when was the last time?

Are you a native English speaker? (Yes / No)

STOP
Do not turn the page until you are told to do so!

Appendix F: Item analysis result table

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
q1	14.10	6.886	.043	.577
q2	14.10	6.748	.189	.567
q3	14.19	6.430	.257	.555
q4	14.32	5.636	.543	.500
q5	14.10	6.817	.116	.572
q6	14.17	7.005	-.085	.594
q7	14.47	6.219	.201	.561
q8	14.46	5.701	.433	.516
q9	14.31	5.974	.383	.530
q10	14.15	6.821	.041	.579
q11	14.32	6.291	.216	.558
q12	14.31	6.595	.082	.580
q13	14.39	6.001	.320	.539
q14	14.56	6.596	.043	.591
q15	14.25	6.745	.030	.586
q16	14.46	6.011	.293	.544
q17	14.51	5.840	.359	.530
q18	14.20	6.372	.270	.552
q19	14.76	7.081	-.136	.617
q20	14.15	6.683	.137	.569

Appendix G: Snapshots of gestures in the action gesture video

close

compression

descend

expelling

explosion

ignite

letting in

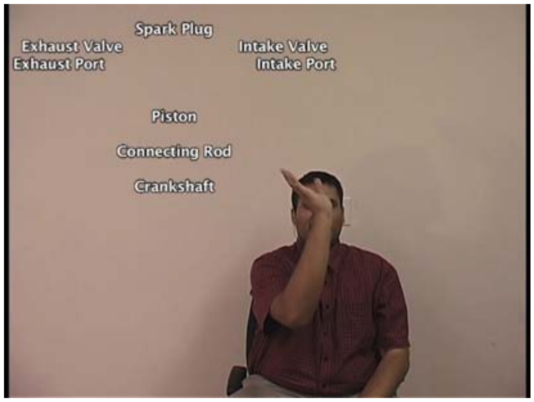
linear motion

open

pointing each part

reducing

rotation


Spark Plug
Exhaust Valve
Exhaust Port
Intake Valve
Intake Port
Piston
Connecting Rod
Crankshaft

go out


Spark Plug
Exhaust Valve
Exhaust Port
Intake Valve
Intake Port
Piston
Connecting Rod
Crankshaft

up

Appendix H: Snapshots of gestures in the structure gesture video

crankshaft

cylinder

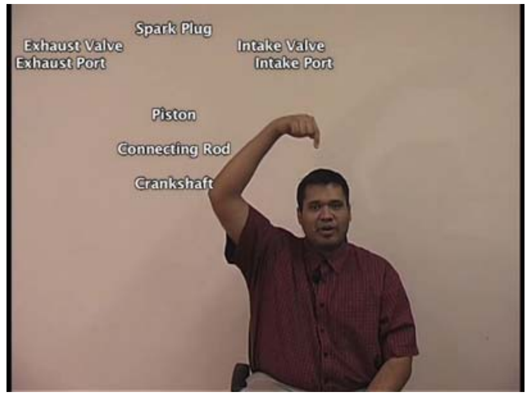
exhause valve

intake valve

mixture of air and fuel

piston

pointing each part



spark plug