
Electronic Theses and Dissertations, 2004-2019

2015

Gestures and mental models: A triple coding hypothesis

Maura Austin
University of Central Florida

 Part of the [Education Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Austin, Maura, "Gestures and mental models: A triple coding hypothesis" (2015). *Electronic Theses and Dissertations, 2004-2019*. 649.

<https://stars.library.ucf.edu/etd/649>

GESTURES AND MENTAL MODELS:
A TRIPLE-CODING HYPOTHESIS

by

MAURA AUSTIN
B.S, University of Central Florida, 2012

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Applied Learning and Instruction
in the Department of Teaching, Learning and Leadership
in the College of Education and Human Performance
at the University of Central Florida
Orlando, Florida

Summer Term
2015

Major Professor: Michele Gill

© 2015 Maura Austin

ABSTRACT

Gestures and speech have been intertwined since the beginning of human communication. Recently the role of gestures in cognition and learning has become a topic of interest in both cognitive and educational psychology. Some researchers have speculated that gestures inherently communicate information that is not provided in purely verbal communication, and that this supplemental information can lead to more thorough mental models in the receiver by acting on a physical/motor modality in addition to the two modalities proposed in the dual code hypothesis. To further understand this issue, in this study, I will examine the effects of watching a gesturing or a non-gesturing lecturer on the learner's cognitive load and mental model development. The results will have implications for cognitive psychology as well as educational psychology, particularly in multimedia learning.

Keywords: learning, multimedia, gestures, mental models, cognitive load, curriculum design, dual-code hypothesis.

ACKNOWLEDGMENTS

I'd like to thank Dr. Michele Gill for her patience, guidance, and wisdom – all of which were important inspirations during my graduate career. Without her this thesis would not have been possible. I'd also like to thank Dr. Valerie Sims and the Applied Cognition and Technology Lab for welcoming me and providing me with valuable guidance and experience in research. Dr. Sims was an inspirational role model in her prowess for research design and analysis throughout my graduate career.

TABLE OF CONTENTS

LIST OF TABLES.....	vii
CHAPTER 1: INTRODUCTION.....	1
Problem Statement.....	1
Theoretical Framework.....	1
Gesture.....	2
Mental Models and Learning.....	3
Cognitive Load.....	4
Theory of Cognitive Multimedia Learning and Gestures - A Third Modality.....	5
Purpose of the Study.....	7
Research Questions.....	8
Hypotheses.....	8
Significance and Implications.....	9
CHAPTER 2: REVIEW OF THE LITERATURE.....	11
Gesture and Speech.....	11
Gesture Production and Cognitive Load.....	12
Gesture Production and Learning.....	13
Gesture and Mirror Neurons.....	15
Watching Gesture.....	16
Gesture Represented in a Third Channel of Processing.....	17
Current Study.....	17
CHAPTER 3: METHODS.....	22

Participants.....	22
Stimuli.....	22
Measures.....	23
Procedure.....	25
CHAPTER 4: RESULTS.....	28
Secondary Vigilance Task Performance.....	28
Post-test Performance.....	28
Feedback Survey.....	30
NASA-TLX.....	31
CHAPTER 5: DISCUSSION.....	34
Mental Demand.....	34
Representative versus Meaningless Gestures.....	35
Gender Differences.....	37
Limitations.....	38
Conclusion.....	39
APPENDIX A: POST-TEST.....	41
APPENDIX B: SCREENSHOTS OF EXPERIMENTAL VIDEOS.....	50
APPENDIX C: DEMOGRAPHICS AND FEEDBACK SURVEY.....	52
APPENDIX D: DIGIT SPAN ANSWER SHEET.....	55
APPENDIX E: NASA-TLX.....	57
APPENDIX F: IRB PERMISSION LETTER.....	59
REFERENCES.....	61

LIST OF TABLES

Table 1 Gender Differences in Post-Test Performance.....	30
Table 2 Gender Differences in Reports of Having Liked the Experimental Video.....	31
Table 3 Condition Differences on the NASA-TLX: Mental Demand.....	32
Table 4 Gender Differences on the NASA-TLX.....	32

CHAPTER ONE: INTRODUCTION

Problem Statement

Recently, online and virtual classes have become ubiquitous among universities. Distance education classes provide a larger audience of students with convenient access to college-level courses than traditional face-to-face classes, allowing non-traditional students to complete bachelor's degrees. However, some argue that online courses do not provide the same level of rigor and meaningful learning experiences as face-to-face courses. In this study, I used principles from Paivio's (1986) dual-coding hypothesis, Wickens' (2002) multiple resource theory, and Mayer's (2005) cognitive multimedia learning theory to propose the possibility that videos of gesturing lecturers may improve online courses by promoting more thorough learning experiences.

While researchers have found that producing gesture can alleviate cognitive load (Cook et al., 2012; Goldin-Meadow, Nusbaum, Kelly & Wagner, 2001), none have examined the effects of watching gesture on cognitive load. As such, in order to contribute to the literature and possibly provide an opportunity for curriculum designers to improve learning experiences online, I will study the effects of watching gesture on cognitive load and learning.

Theoretical Framework

In his seminal book on gestures, Clark (1996) posited that gestures and speech comprise one single system of language processing. He stated that gesture and language complement each other in order to create a complete message, and, as such, communication that lacks either gesture or speech is linguistically incomplete. Since then, gesture's relationship with speech has been thoroughly researched (for a review, see Hostetter, 2011). Considering that many online

classes lack the visual and auditory components of lecture in a traditional face-to-face class, Clark's (1996) theory would imply that online learning classes provide incomplete models of communication of the class material, possibly affecting students' ability to absorb, comprehend, and remember the content.

Clark's theory is further supported by Paivio's (1986) dual-coding theory, which posits that learners can better comprehend incoming information when the material is presented in two modalities. Both Paivio (1986) and Wickens (2002) in two separate seminal papers theorized that processing requires less cognitive load when it takes advantage of both the auditory and visual channels. As such, gesture could support cognitive processing and help alleviate cognitive load by spreading information across multiple modalities. Further, I hypothesize that there may be a third channel - a motor channel - through which gestures may be processed and work to further alleviate cognitive load during learning. I cover the details of this hypothesis below.

Gesture

Recently, there has been growing interest in gesture's previously unrecognized role in cognition and learning, with evidence indicating that gestures play a larger role in cognition than was previously understood. For example, gestures have been found to lighten cognitive load during mathematical problem solving tasks (Cook, Yip, & Goldin-Meadow, 2012), facilitate lexical access (Cook, Yip, & Goldin-Meadow, 2012), and promote transfer of newly learned material more so than working with concrete objects (Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014). Gesture's facilitative role in learning continues to be examined, yielding results with far-reaching implications for both educational psychology and cognitive psychology.

Currently, the literature on gestures tends to refer to a set of definitions first proposed by McNeill (1992) that identifies four kinds of gestures:

1. *Iconic gestures* - Gestures that represent concrete semantic elements of speech
e.g. stroking the air to represent the act of petting a dog
2. *Metaphoric gestures* - Gestures that represent more abstract elements of speech
e.g. putting hands together to communicate “thank you”
3. *Deictic gestures* - Gestures that generally involve pointing either with the purpose
of identifying an object or sometimes used to refer to “conversational space” (p.
422, Goldin-Meadow, 1999)
4. *Beat gestures* - Least representative of all gestures, beat gestures accompany
speech rhythmically and at times to stress certain parts of speech (e.g., tapping a
finger).

Mental Models and Learning

Some researchers have recently proposed that even meaningless gesture, such as metaphoric and beat gestures, support learning by helping to develop a more thorough mental model during communication (Cutica & Bucciarelli, 2008; Noice & Noice, 2001). This supports the theory first posited by Clark (1996) that language without gesture is communicatively incomplete and could possibly determine the mechanisms by which gesture supports communication.

The theory that more well-developed mental models correlate with more meaningful learning is well established in the educational psychology literature. Of particular note is Mayer’s (2001, 2005) cognitive theory of multimedia learning, which outlines a theoretical relationship between the dual coding hypothesis (Paivio, 1986) and mental models. The dual code hypothesis (Paivio, 1986), along with Baddeley’s model of working memory (1986, 1999) and Wickens’s multiple resource theory (2002), suggest that learners have limited capacity to

process incoming information, and that this capacity is divided into resource pools drawing from two modalities - visual and auditory. As such, information that is presented in two modalities simultaneously capitalizes on both resource pools, resulting in minimized cognitive load and maximized processing capacity (Paivio, 1986; Wickens, 2002). Mayer (2001, 2005) builds on the dual code hypothesis, positing, as a result of more thorough processing, learners also develop more thorough mental models when incoming information is presented in two modalities. As a result, learners who receive information in two modalities tend to retain the information better and perform better on transfer tasks, suggesting that they understand the information more thoroughly (Baggett, 1984; Mayer, 1989; Mayer & Anderson, 1991; Mayer & Gallini, 1990).

Cognitive Load

Cognitive load plays a large part in Mayer's (2005) cognitive theory of multimedia learning in that minimizing cognitive load is a major goal in designing efficient multimedia learning curriculum, and reduced cognitive load is generally correlated with better learning outcomes. In his work, Mayer (2001, 2005; DeLeeuw & Mayer, 2008) refers to Sweller's (1999, 2006) triarchic theory of cognitive load, which defines three different types of cognitive processing with various effects on learning:

1. *Intrinsic load* - Results from processing and comprehending incoming material.
2. *Extraneous load* - Results from processing additional information that is not essential to comprehending the incoming material.
3. *Germane load* - Results from deep processing of incoming material, including developing mental models.

Ideal learning activities tend to minimize extraneous load so as to allow for maximum germane processing by reducing redundant and otherwise unnecessary information (Mayer,

2005). For example, in multimedia learning, information that is presented verbally in two modalities (e.g. words written on a screen and spoken simultaneously) results in extraneous load as the information is being conveyed in a redundant fashion (Mayer, 2005). On the other hand, material that is presented in two modalities while avoiding redundancy by communicating the material in qualitatively different forms (e.g. spoken words accompanied by a visual picture that conveys the message) maximizes cognitive resources and leads to more detailed mental models and better learning outcomes (Mayer, 2005).

Theory of Cognitive Multimedia Learning and Gestures - A Third Modality

I propose that Mayer's (2001, 2005) theory of cognitive multimedia learning can potentially support Cutica and Bucciarelli (2008) and Noice and Noice's (2001) hypotheses that learners who watch gesturing lecturers or perform gestures themselves develop more thorough mental models of the learned material. Gestures provide a third modality through which learners can process incoming material, resulting in lowered extraneous cognitive load and more thoroughly developed mental models. A third modality - physical/motor - that involves the use of gestures has not been proposed in the literature. I hypothesize that a third modality may be an important addition to the dual coding hypothesis (Paivio, 1986) and Mayer's (2005) cognitive theory of multimedia learning.

In fact, a coding theory involving three processing channels - an auditory, visual, and physical/motor - would parsimoniously explain the majority of the current findings on gestures' facilitative role in learning. For example, Cook et al. (2012) may have found that gesture production reduced cognitive load because gestures may draw from a third modality, which acts to lower cognitive load in the same way that involving both visual and auditory modalities in a learning task does (Mayer, 1989; Mayer & Anderson, 1991; Paivio, 1986). Further, studies

which found that gesture production facilitates retention may be able to be explained in the same way that the dual coding theory explains how learning tasks that involve multiple modalities also tend to result in better retention (Mayer, 1989; Mayer & Anderson, 1991).

It may seem intuitive to consider gesture processing as being visual in nature; thus any learning that involves both speech and gesture would make use of the visual and verbal channels and Paivio (1986) and Wickens (2002) first described. However, in a study measuring gesture's effect on cognitive load while explaining math problems, researchers have found that not all gestures draw from the visuospatial processing channel and hypothesized that gestures may be represented in a motor modality rather than a verbal or visual one (Wagner, Nusbaum, & Goldin-Meadow, 2004). In their study, Wagner et al. (2004) compared participants' performance on a dual-task paradigm in which they either memorized visuospatial information or verbal information before explaining math problems either with or without gesture. Per Paivio's (1986) and Wickens's (2002) theories, participants should perform worse on the memorization task that requires the same type of processing as gestures. As such, if gestures were processed in the verbal channel, participants would perform worse on the verbal memorization task as both tasks are imposing cognitive loads in the same modality. However, Wagner et al. (2004) found that gesture promoted better performance on the memorization tasks in both conditions, suggesting that gesture processing is not solely verbal or visuospatial in nature. Based on their results, Wagner et al. (2004) suggested that there is a possibility that gestures may be represented in another modality: a motor modality. They called for more research to study whether this may be the case.

Purpose of the Study

Cutica and Bucciarelli (2008) and Noice and Noice's (2001) theories that gestures promote the development of more thorough mental models during learning have yet to be empirically examined. Because these hypotheses could possibly further our understanding of the mechanisms by which gesture tends to support communication (Clark, 1996), I feel it warrants further investigation. As outlined above, I propose examining this hypothesis through the lenses of the dual code hypothesis (Paivio, 1986) along with Mayer's (2001, 2005) theory of cognitive multimedia learning, which suggest that learners develop more thorough mental models of incoming material and experience minimized cognitive load when the material is presented efficiently in two modalities. Considering Clark's (1996) theory that gestures supplement speech communication to create a more complete message, I hypothesize that gestures facilitate the development of thorough mental models by drawing from a third modality - physical/motor. Building on Paivio's (1986), Wickens's (2002) and Mayer's (2001, 2005) theories, gestures could facilitate language processing by providing another channel through which communication can be processed. Per Mayer's (2001, 2005) theory, providing an additional processing channel would lower cognitive load and effort required to understand the incoming content, allowing more cognitive resources to be allocated to developing a more thorough understanding of the content. As such, it may be through this mechanism that gestures supplement pure speech during communication, as was theorized by Clark (1996) and others.

This study will also aim to test Noice and Noice's (2001) theory that gestures inherently communicate supplemental information that is not conveyed in pure verbal communication. If Noice and Noice's (2001) hypothesis that gestures communicate additional information that is not communicated verbally, then watching a gesturing lecturer should result in lowered

extraneous load, as the information is not redundant and unnecessary in its presentation.

However, if gestures simply reinforce the information communicated verbally, then they act as redundant information and will result in higher extraneous load (Mayer, 2005).

Research Questions

1. Do learners who watch a gesturing lecturer develop a more thorough mental model of presented material than learners who watch a non-gesturing lecturer?
 - a. Do learners who watch a gesturing lecturer perform better on a problem-solving task after watching the lecture than learners who watch a non-gesturing lecturer?
2. Do learners who watch a gesturing lecturer experience less extraneous cognitive load than learners who watch a non-gesturing lecturer?
 - a. Do learners who watch a gesturing lecturer react more quickly to a secondary vigilance task while watching the video than learners who watch a non-gesturing lecturer?

Hypotheses

1. I hypothesize that learners who watch a gesturing lecturer will answer more questions on the problem solving task correctly than learners who watch a non-gesturing lecturer, suggesting that they developed a more thorough mental model of the material and that gestures draw from a motor/physical modality that supports my proposed theory of triple coding - as opposed to the dual coding hypothesis (Paivio, 1986).
2. I hypothesize that learners who watch a gesturing lecturer will react more quickly to a secondary vigilance task than learners who watch a non-gesturing lecturer,

suggesting that they experienced less extraneous cognitive load and that the information communicated by the gestures is not redundant to the information communicated verbally.

Significance and Implications

Should the outlined hypotheses be supported, such results could support a new theory of triple - rather than dual (Paivio, 1986) - coding, in which processing and mental representation of incoming information occur in visual, auditory, and motor/physical modalities. To the best of my knowledge, such a theory has been neither proposed nor supported in empirical studies. A triple-coding theory would have implications for cognitive psychology in providing a more thorough model of coding processes. The results of this study carry significant theoretical relevance, as I will simultaneously build on Paivio's (1986) dual-coding hypothesis and integrate Mayer's (2001, 2005) theory of cognitive multimedia learning, thus proposing a theoretical integration that has not yet been explored.

Further, such a theory would be relevant to educational psychology, in that it could inform the development of curricula and lesson plans that more thoroughly integrate all three modalities of processing and mental representation, thus possibly promoting learning outcomes and minimizing cognitive load. Currently, online distance education classes are becoming more widely used, as they are more easily accessible for students who do not live near a campus and for non-traditional students whose schedules may be more demanding. With the growing use of online classes, researchers have been examining differences in learning and classroom experiences between face-to-face classes and online classes. However, much research on instructional design and developing curricula for distance-learning classes focuses on the lack of classroom community in online classes and how to promote interpersonal communication that

promotes positive learning outcomes, or what kinds of visuals and presentations are best suited to maintain learners' attention, facilitate meaningful understanding, and promote retention (see Means, Toyama, Murphy, Bakia, & Jones, 2009).

This study examined learning that can be applied to an online course and could provide a much simpler way to promote more meaningful learning in online classes. Should I find that watching lectures involving gestures results in more thoroughly developed mental models and successful application of the lecture material than watching lectures without gesture, instructional designers will have reason to believe that simply including a naturally speaking and gesturing lecturer in online classes can improve learning outcomes.

CHAPTER TWO: REVIEW OF THE LITERATURE

Gesture's affect on cognition has been extensively studied especially in the past decade. Researchers have examined, among other topics, the effects of producing meaningless compared to meaningful gestures on working memory (Cook, Yip, & Goldin-Meadow, 2012), the effects of beat gestures on language (Graham & Heywood, 1975; Rauscher, Krauss & Chen, 1996; Ravizza, 2003), the relationship between watching gesture and activation of the mirror neuron system (Brucker et al., 2014, Mainieri et al., 2013; Molenberghs, Cunnington & Mattingley, 2011), and the effects of producing gesture on memory and learning tasks (Cook, Mitchell & Goldin-Meadow, 2008; Noice & Noice, 2001; Novack, Congdon, Hemani-Lopez & Goldin-Meadow, 2014). The literature on the cognitive effects of gesture is robust and continues to expand today. However, some gaps in the literature still exist. For example, very little work has been done investigating the cognitive effects of watching gesture. The following chapter will unpack the current relevant literature on the cognitive effects of gesture and will address the questions that still remain unanswered.

Gesture and Speech

Gesture has been found to facilitate fluent speech production in a variety of ways. Researchers have found that speech accompanied by gesture and even meaningless movements has fewer pauses and is more quick and fluent than speech that is not accompanied by gestures (Graham & Heywood, 1975; Rauscher, Krauss & Chen, 1996). Further, Ravizza (2003) found that having participants produce meaningless movements - particularly rhythmically tapping their fingers - resulted in quicker resolution of tip of the tongue states. He speculated that the act of movement caused increased activity in motor production areas of the brain, which facilitated speech production (Ravizza, 2003).

Gesture has also been found to facilitate speech comprehension. In her meta-analysis on gesture research, Hostetter (2011) found that, across studies, watching gesture accompany speech tended to facilitate improved comprehension of the message. Further, gestures that were meaningful and representative, such as iconic or deictic gestures, tended to be more supportive of comprehension than meaningless gestures, such as beat gestures.

Gesture Production and Cognitive Load

Meaningful and representative gestures have also been found to be qualitatively different from meaningless movements in the nature of their facilitative roles on cognition. In a study comparing the use of meaningless movements and gestures while explaining math, Cook et al. (2012) found that meaningless movements facilitated lexical access but did not offset cognitive load as meaningful gesturing did. They suggested a few different reasons this may have occurred. Meaningful gestures may offset cognitive load by providing additional structure to the speakers' math explanations or by conveying information in a second modality, which has been found to ease cognitive load (Cook et al., 2012; Mayer, 2005; Paivio, 1986; Wickens, 2002). Alternatively, gestures may have simply helped the speakers focus their attention on the task at hand, thereby easing cognitive load (Cook et al., 2012). Finally, Cook et al. (2012) suggested that meaningful gesturing may have allowed the speakers to externalize their abstract ideas, thus offsetting cognitive load in the same way that working with concrete objects requires less mental effort than manipulating abstract ideas or representations, such as in a mental rotation task.

Similarly, researchers found that both adults and children who were allowed to spontaneously gesture while explaining their math problem solving strategies experienced less cognitive load than those that were constrained and thus prohibited from gesturing (Goldin-Meadow, Nusbaum, Kelly & Wagner, 2001). Goldin-Meadow et. al (2001) measured cognitive

demand by administering a working memory task in the form of a word list. Participants were given a word list before providing their explanations and then recalled the words. Goldin-Meadow et al. (2001) found that those that were allowed to spontaneously gesture were able to recall more words from the list on average than those that were not, suggesting that gesturing relieved some of the cognitive demand of explaining math strategies.

Gesture Production and Learning

Gestures have also been found to promote meaningful learning, as measured by students' performance on tasks that require transfer of newly learned material. In a study comparing learning math with concrete objects and with gestures, Novack et al. (2014) found that students who were taught meaningful gestures to represent math equations more successfully completed tasks that required transfer of their learned knowledge than students who used concrete blocks to learn the same math concepts. They speculated that this phenomenon might be prevalent at a specific developmental stage - as defined by Piaget (1983) - during which students are capable of abstract thought during scaffolded instruction (Novack et al., 2014). Gestures act as a convenient midpoint between abstract thought and the physical world, thus scaffolding students' abstract math concepts in a way that allows for deeper understanding than working with concrete objects (Novack et al., 2014).

Producing gestures while learning has also been found to facilitate memory encoding and retrieval. Cook et al. (2008) found that young children who were taught to gesture specific meaningful representative movements while learning new mathematical concepts retained the newly learned content more thoroughly when tested four weeks later than children who were not taught to gesture. Cook et al. (2008) suggested that their results support recent theories in embodied cognition (Barsalou, 1998) in that using the body to represent abstract concepts can

facilitate deep and meaningful understanding. Cook et al. (2008) also noted that their evidence supports the theories that gesture facilitates learning, but that they cannot currently draw any conclusions on what specific mechanisms are responsible for their supportive role. They theorized that gestures might possibly allow for more easily constructed mental representations of the incoming information, thus freeing mental resources to be allocated to creating a more thorough understanding of the material. They also hypothesized that gestures may contribute to long-term memory encoding and retrieval in that gestures provide more action based encoding, creating “more robust memory traces” (Cook et al., 2008, p. 1055). Finally, they wondered if perhaps meaningful gestures, such as pointing, ground the new material in the physical environment, thus imposing a lower cognitive demand on the learner and perhaps allowing for more insights than if the new material were not grounded in the environment.

Similar results have been found in various studies, with gestures facilitating long-term retention of phrases in a foreign language (Allen, 1995) and of observed events (Cook, Yip & Goldin-Meadow, 2010). Further, Noice and Noice (2001) found that inexperienced actors who produced movements associated with a script while learning the script retained more of the material than those that had not. They posited that their subjects retained script material more successfully when they learned accompanying physical movements because such movements inherently hold more information than the purely verbal components of the script, thus helping the subjects develop a more thoroughly informed and organized mental model of the material they were memorizing (Noice & Noice, 2001). As such, Noice and Noice (2001) suggested that relevant gestures and other meaningful bodily movements inherently convey supplemental information that cannot be communicated verbally, and that this information provides additional structure for incoming material to be understood and retained.

The majority of the current literature in gesture research examines the effects of gesture production on cognition. However, if Noice and Noice's (2001) and Clark's (1996) theories that gesture and speech complement each other in a single system of linguistic communication hold true, then watching gesture should be just as effective in influencing cognition and communication as producing gesture. In addition, evidence from research on the mirror neuron system, covered below, also suggests that watching gestures should have the same neurological effects as producing gestures.

Gesture and Mirror Neurons

Mirror neurons - which activate both while intentionally producing a motor act and when passively watching another produce the same act - were first discovered in Macaque monkeys (see Gallese, Fadiga, Fogassi & Rizzolatti, 1996) and since then have been well established in human neurobiology (e.g Molenberghs et al., 2011). Following their discovery, cognitive psychologists and neuropsychologists have researched their various roles across many aspects of human cognition, with studies examining their possible roles in theory of mind processes, Autism, imitation, understanding facial expressions, and learning (see Molenberghs et al., 2011).

The mirror neuron system (MNS) consists of neurons that activate identically both when a person is carrying out a specific motor act and when they are watching someone else carry out the same motor act (Gallese, Fadiga, Fogassi & Rizzolatti, 1996; Molenberghs et al., 2011). Researchers have found that the MNS activates not only during passive observation of an intentional motor act, but also during observation of facial expressions (Montgomery et al., 2007) and communicative gestures (Brucker et al., 2014; Mainieri et al., 2013; Molenberghs et al., 2011). Following this evidence, I should expect similar neurological responses in learners during both the production and observation of gestures - assuming that the learners had

previously produced the gestures themselves at some point in their past. As such, I should expect that previous results indicating the beneficial role of producing gestures during learning tasks (Allen, 1995; Cook et al., 2008; Cook et al., 2010; Cook et al., 2012; Noice & Noice, 2001; Novack et al., 2012) should be equally relevant and applicable to the act of observing gestures.

Further, research from literature on the mirror neuron system (MNS) suggests that the act of watching gestures supports information processing in both the visual and the physical/motor modalities (Brucker, Ehlis, Häußinger, Fallgatter, Gerjets & 2014; Mainieri, Heim, Straube, Binkofski, & Tircher, 2013; Molenberghs, Cunnington, & Mattingley, 2011; Montgomery, Isenber, & Haxby, 2007) and, as such, should not communicate redundant information or overload the visual modality. Thus, contributing gestures to a visual lecture should result in lower cognitive load, per Mayer's (2005) cognitive theory of multimedia learning and Paivio's (1986) dual-code hypothesis. Indeed, as outlined above, evidence for gesture's minimizing effect on cognitive load has been found in various studies (Cook et al., 2012; Goldin-Meadow et al., 2001).

Watching Gesture

While little work has been done investigating the effects of observing gesture, the little evidence found on the topic suggests that watching lecture can provide the same facilitative cognitive effects as producing gesture, as would be expected per the literature on the mirror neuron system. Cutica and Bucciarelli (2008) found that students who viewed a video of a gesturing speaker retained more information and made more correct inferences on the content of the speech than those that viewed the same speech by an actor who was not gesturing. Their results support past findings that have suggested that watching gesture can facilitate comprehension during communication (see Hostetter, 2011). In the same vein as Noice and

Noice's (2001) theory, Cutica and Bucciarelli (2008) suggested that the students who viewed a gesturing professor had developed a more thorough mental model of the lecture content.

Although both Cutica and Bucciarelli (2008) and Noice and Noice (2001) posited that gestures facilitate the development of a more thorough mental model, neither study was able to provide convincing support for their theory. Both studies indeed found that participants who produced or watched gestures retained more information from their learning activities; however, neither author was able to definitively conclude that said results were caused by more thoroughly developed mental models.

Gesture Represented in a Third Channel of Processing

In an attempt to provide a more detailed theoretical framework for Cutica and Bucciarelli's (2008) and Noice and Noice's (2001) theories, I posit that there may be a third channel in addition to the two that Paivio (1986) and Wickens (2002) proposed that provides an additional opportunity for information to be processed in a different modality. A third-channel of processing can support Cutica and Bucciarelli's (2008) and Noice and Noice's (2001) theories that producing and watching gestures promote learning by encouraging more thorough mental model development of incoming information. Per Mayer's (2001, 2005) theory of cognitive multimedia learning, providing an opportunity for information to be processed in multiple channels promotes more thorough mental model development during learning. As such, it may be possible that gestures support the development of more thorough mental models during learning by providing a third channel – a motor channel.

Current Study

The current study examines the theory that gestures have been found to facilitate comprehension and memory by supporting the development of a more thorough mental model,

as posited by both Cutica and Bucciarelli (2008) and Noice and Noice (2001). While the facilitative cognitive effects of gesture production have been well documented, similar effects from watching gesture have not been thoroughly researched. As Noice and Noice (2001) hypothesized that gestures help develop more thorough mental models by communicating supplemental content that is not conveyed in pure speech, it should follow that watching gesture should produce the same facilitative effects as producing gesture. As such, the current study will examine the effects of watching gesture on cognition to test their hypothesis.

As indicated above, researchers have found that producing gesture tends to minimize cognitive demand (Cook et al., 2012; Goldin-Meadow et al., 2001). In the current study, I test whether watching gesture can produce the same effect. Both studies that found this effect used a secondary task to measure cognitive load while completing a task that involved producing gesture. Similarly, to measure cognitive load in the current study, I use a secondary vigilance task, which involves responding to a change in a background screen color as quickly as possible. This and other similar secondary vigilance tasks have been used in previous studies measuring cognitive load imposed by a specific task, and have been found to be a valid indicator of task difficulty and cognitive load when compared to other verified measures of cognitive load (DeLeeuw & Mayer, 2008). According to theory, the participant's reaction time to the background color change correlates to the amount of cognitive load experienced, with higher reaction times indicating higher experiences of cognitive load (DeLeeuw & Mayer, 2008). As a supplemental measure, in the current study I also administer a paper and pencil version of the NASA-TLX to measure cognitive load, frustration, and perceived task difficulty. Finally, I have participants provide reports of how clear, informative, and likeable the videos were in order to measure any differences in perceived quality of the lecture based on the presence of gesture. As

far as I am aware, there have not been any previous studies that have sought to test whether watching gesture alleviates cognitive load.

Researchers in the past have found that gesture production can facilitate both recall and transfer of newly learned information (Cook et al., 2008; Cook et. al, 2010; Noice & Noice, 2001; Novack et al., 2014). However, a large majority of the current research on gesture focuses on its effects on learning concrete subjects, such as simple math. It makes sense that researchers studying effects of gesture within concrete subjects have found evidence for the supportive cognitive effects of gesture, as Hostetter (2011) found in her meta-analysis that representative gestures tend to facilitate comprehension more so than meaningless gestures. However, will these effects hold true for meaningless gesture in abstract topics? The current study seeks to address this question. To do so, I chose to study how gesture affects learning about logical fallacies - a topic that is abstract in nature and that cannot be expressed or supplemented with iconic, representative gestures.

As outlined previously, various researchers have found evidence that gestures facilitate memory, especially for long-term retention (Allen, 1995; Cook et al., 2008; Cook et al., 2010). However, less work has been done to test whether gesture facilitates meaningful learning and transfer of new material. As such, the current study measures gesture's effects on both memorization and ability to transfer newly learned material. To do so, the post-test that is administered after watching the video lecture contains questions that are purely memorization in nature (i.e. recalling examples used in the video or phrases spoken by the lecturer) as well as questions that require transfer of the newly learned material (i.e. problem solving task using the material from the lecture).

To my knowledge, there is only one previous study that has tested the effects of watching gesture similarly to the current study. Cutica and Bucciarelli (2008) measured differences in participants' ability to memorize and draw inferences on discourse material presented in a video speech of either a gesturing or a non-gesturing speaker. In their first experiment, participants watched a video of an actor reciting a story and were asked to recall as much as possible from the video in a free-recall task. Cutica and Bucciarelli (2008) found that participants who watched the gesturing actor as opposed to the non-gesturing actor recalled more information and made more correct discourse based inferences. In their third experiment, to more rigorously test their hypothesis that gestures facilitate the development of more thorough mental models, they used a video speech of a topic that was more abstract in nature. The participants completed a free-recall task as well as a recognition task in the third experiment to test both pure memorization and ability to draw inferences on the topic. Cutica and Bucciarelli (2008) found the same results as in the first experiment, supporting their hypothesis that gestures help to create a more thorough mental model of information even for abstract information.

Although they found differences between their gesture and non-gesture viewing groups, the videos that they used were two different videos that were filmed separately. The actor gave the same speech in both videos, but was instructed to not gesture in one of the videos. As such, the speaker may have differed in his voice, intonation, or facial expressions between the two videos. To address this issue, Cutica and Bucciarelli (2008) conducted a pilot experiment in which participants listened to the audio recordings of either the gesture or the non-gesture video and then completed a free-recall task. They found no significant differences between the groups and concluded that there were no differences in voice, intonation, or other implicit factors between the two videos.

However, in order to conclude that gesture - and gesture alone - was responsible for the differences between the two experimental groups, all other possible differences between the videos must be eliminated. As such, other than the visibility of the speaker's gestures, the videos that are used in this study are identical. This was achieved by filming one video that included the speaker's body from the waist up, then cropped a second video in which the speaker is only visible from the shoulders up. As such, there are no possible differences in any factors of speech or facial expressions between the two videos.

Cutica and Bucciarelli (2008) used free-recall tasks to measure their participants' ability to recall and draw inferences on the discourse in their video. They had two judges score their participants' responses, but did not mention whether the judges were blind to the participants' condition or the investigators' hypothesis. In the current study, I sought to eliminate any possibility of confirmation bias by using the objective measure of a multiple-choice posttest. Further, I sought to test whether gesture affects meaningful learning by requiring participants to complete problem solving and transfer questions in the posttest. Problem solving and transfer questions should be more sensitive measures of meaningful learning and mental model development than the recall task used by Cutica and Bucciarelli (2008).

As outlined above, I intended to test Cutica and Bucciarelli's (2008) theory and sought to eliminate any possibility that their results may have been affected by variables other than gesture. Further, I intended to measure meaningful learning in a more rigorous manner by testing learners' ability to problem solve using the lecture material. Per Cutica and Bucciarelli's (2008) and Noice and Noice's (2001) theories that gestures help develop more thorough mental models, I expected to find differences in ability to memorize and transfer the abstract lecture material between those that view a gesturing speaker and those that view a non-gesturing speaker.

CHAPTER THREE: METHODS

Participants

Eighty-three undergraduate students (32 males, 51 females) at the University of Central Florida (UCF) between the ages of 18 and 40 ($M = 19.98$, $SD = 2.93$) chose to take part in the study to fulfill a class requirement. The study took place in the Psychology Building at UCF and used the Applied Cognition and Technology Lab space in room 207D.

Stimuli

All stimuli were presented individually to each participant. The digit span task was given via an automatic PowerPoint presentation. Results were recorded on a paper answer sheet (see Appendix D).

SuperLab 4.5 and SuperLab 5.0 were used to present the pre-test, training video, experimental video, and post-test. Two computers in the ACAT Lab used SuperLab 5.0 and three used SuperLab 4.5. All computers are Dell computers running Windows.

The pre-test was initiated via SuperLab after an instruction screen was presented. Participants typed their answers into a text box provided on the screen.

A video of a nature scene with a runtime of 1:11 was used for the secondary vigilance task training portion. The video was presented full screen on the computer monitor and was run by SuperLab after an instruction scene was presented. The video had a solid color background around the scene frame (see Appendix B). The color of the background gradually changed from pink to black or black to pink at the following times: 00:04, 00:18, 00:35, and 00:54. Participants reported a background color change by pressing the spacebar on the keyboard. Each key-press was recorded by SuperLab.

Based on the randomized condition, one of two videos was used for the experimental lecture video: either a gesture or a non-gesture video. Both videos had a runtime 3:01 and were presented full-screen after an instruction screen was presented. The video used for the gesture condition featured the lecturer from the waist up and the video used for the non-gesture condition was cropped to feature the lecturer from the shoulders up so as to make his gestures not visible (see Appendix B). The videos were identical in every other sense. As with the training video described above, both videos had a surrounding background color that changed from black to pink or pink to black at the following times: 00:10, 00:32, 1:00, 1:33.9, 1:48.9, and 2:22.9. Participants reported a background color change by pressing the spacebar on their keyboard. Each key press was recorded by SuperLab.

Measures

A secondary vigilance task was used during the video to measure extraneous cognitive load. The task required participants to respond to a change in background color as quickly as possible by pressing the spacebar. The background color gradually changed from black to pink at random times throughout the video. This task was designed according to the design used by DeLeeuw and Mayer (2008) in their examination of the efficacy of secondary vigilance tasks as a measure of cognitive load. They determined that this specific design successfully measured cognitive load, and, as such, I found it appropriate to mimic their design as closely as possible.

The multiple-choice posttest was initiated via SuperLab after an instruction screen. Each question was presented one at a time on the screen and participants responded to each question by pressing the key corresponding to their chosen answer (A, B, C, D, etc.). Each key press was recorded by SuperLab. The posttest originally consisted of twenty questions, but two were not

included in the final data analysis as a result of technological difficulties with response recording.

The final post-test was comprised of two kinds of questions. Thirteen transfer questions required application of the lecture content in a problem-solving manner (e.g. “which of the following is an example of ‘ad hominem’ in use?”). Three of the questions from this section of transfer questions were taken from GRE writing practice tests and the others were written by the primary investigator. The remaining five questions required participants to recall facts from the lecture video and acted as a measure of memorized material from the video. As such, the post-test measured participants’ ability to transfer the learned material as well as to encode and recall the lecture content.

Two of the questions tested knowledge on the fallacy of the single cause, three on the appeal to authority fallacy, six on ad hominem, and five on the post hoc fallacy. In addition, two questions referred to both the fallacy of the single cause and the post hoc fallacy.

In the original study design, the post-test included two creativity questions as follows: 1. “List as many one-sentence examples of an attempt to “appeal to authority” as you can.” 2. “List as many one-sentence examples of an attempt to “ad hominem” as you can.” However, the technical difficulties occurred during the administration of these questions, and many participants’ answers were not recorded. As such, they were removed from the post-test and were not included in the final analysis.

The NASA-TLX (Task Load Index) is a widely used and robustly tested self-report measure used to measure perceived frustration and cognitive load on specific tasks (Hart & Staveland, 1988). It consists of six categories that are designed to measure perceived task difficulty (see Appendix E). It was administered via pencil and paper, as it has been found that

the paper and pencil version imposes less cognitive load than the computer version (Noyes & Bruneau, 2007). The original version of the NASA-TLX included a section in which participants rate the importance of each variable to overall mental workload. However, I administered a raw TLX in which this second section is not required, as it has been suggested that the second section may be unnecessary and using a raw TLX increases validity (Noyes & Bruneau, 2007).

The demographics and feedback surveys were also given via pencil and paper (see Appendix C). The demographics questionnaire was administered to record information on gender, age, GPA, and year in school. Included in the demographics questionnaire was a feedback survey, in which participants gave feedback on the effectiveness, clarity, and likeability of the both video and the speaker on a 5-point scale (see Appendix C).

Procedure

First, all participants were given a document of informed consent and agreed to participate in the study.

Next, participants completed a working memory measure in the form of a digit span task. The task required the participants to memorize strings of numbers and to report them immediately after each string was presented. There were 14 number strings in total, beginning with a three-number string and increasing to a nine-number string. Seven of the number strings were reported as they were presented and seven number strings were reported in the reverse order that they were presented (see Appendix D). Only data of those who performed within two standard deviations of the average were included in the analysis. Four participants' data were excluded from analysis per this requirement.

Participants then completed a pretest in which they defined the four logical fallacies that were to be explained later in the video lecture: ad hominem, the post hoc fallacy, appeal to

authority, and the fallacy of the single cause. Any participant that demonstrated significant prior knowledge on the subject by correctly defining at least three of the four logical fallacies was not included in the final data analysis. Eight participants' data were excluded from analysis per this requirement. One participant's data was excluded due to her visual problems related to strabismus that could affect her results.

Next, participants were trained to complete a secondary vigilance task in which they were required to report gradual changes in background color (black to pink or pink to black) throughout a video with a 1:11 runtime. To do so, participants were instructed to immediately press the spacebar when they first perceived a change in the background color. The primary investigator gave an example to each participant individually on how to complete the secondary vigilance task, then allowed them to complete the task for the remainder of the training video. Training was implemented as a result of participants' failure to successfully complete the secondary vigilance task without training during a pilot study.

Participants then randomly assigned to watch one of two videos: a gesture condition video ($n = 40$) or a non-gesture condition ($n = 43$) video. Both videos were exactly three minutes and one second long in runtime. Both videos featured an experienced male high-school teacher giving a scripted lecture on the following logical fallacies: appeal to authority, post hoc ergo propter hoc, the fallacy of the single cause, and ad hominem. The speaker defined each fallacy and gave an example of the fallacy being used in an argument. The video used for the gesture condition was framed to feature the speaker's body from the waist up so as to include his gestures. The video used for the non-gesture condition was cropped at the speaker's shoulders to make his gestures not visible (see Appendix B). As such, the video and lecture were identical for

both conditions - other than the difference in the visible frame - to control for any speech differences that might have occurred should two separate videos have been filmed.

While watching the video, participants completed a secondary vigilance task in which they were required to identify six different gradual changes in background color (pink to black and black to pink) behind the video as quickly as possible (see Appendix B). As with the training task, participants were instructed to immediately press the spacebar when they first perceived a change in the background color.

After watching the video, the participants completed a twenty-question post-test (see Appendix A). Next, participants completed a paper and pencil version of the NASA-TLX to measure overall cognitive load, perceived task difficulty and frustration (see Appendix E). Finally, the participants were debriefed and were given thorough explanations of the study.

CHAPTER FOUR: RESULTS

Secondary Vigilance Task Performance

I analyzed performance on the secondary vigilance task using the reactions times (RTs) that were automatically recorded by SuperLab. Data was recorded for RTs to each of the six background color changes as well as an average RT to all six changes. No gender differences in average RT were found, and assumptions of normality were met, so data were pooled for hypothesis testing. A one-way between subjects ANOVA was conducted to determine the effect of watching gesture on RT to the secondary vigilance task, and found no significant differences between the two conditions in average RTs [$F(1, 81) = .15, p = .71$]. As such, gesture did not appear to affect cognitive load as measured by RT.

Post-Test Performance

To measure the effect of watching gesture on learning, post-test performance between groups was analyzed. The questions were multiple-choice, and answers were recorded as either incorrect or correct. All analyses for post-test performance were done using a one-way between subjects ANOVA, and assumptions of normality were met unless otherwise stated.

No significant differences between conditions in average post-test performance were found [$F(1, 79) = 1.19, p = .28$], but significant gender differences were found [$F(1, 79) = 6.97, p = .01, \eta_p^2 = 0.072$] with males averaging 73.89% of questions correct ($SD = 17.60\%$) and females averaging 64.63% ($SD = 15.50\%$). No interaction effect was found [$F(1, 79) = 2.48, p = .12$].

Performance differences were analyzed between groups on questions that only tested memorization of the lecture material. There were no significant differences between groups on memorization questions [$F(1, 79) = .46, p = .50$]. While there were no gender differences in memory question performance, an interaction effect approaching significance between gender and condition was found [$F(1, 79) = 3.76, p = .056, \eta_p^2 = 0.045$].

Analysis was performed on the 13 questions that required transfer of the lecture material and no significant differences between conditions were found [$F(1, 79) = 1.02, p = .32$], but significant gender differences were found [$F(1, 79) = 9.37, p = .003, \eta_p^2 = 0.106$], with males performing better ($M = 73.22\%, SD = 19.24\%$) than females ($M = 60.65\%, SD = 18.51\%$). However, there was no interaction effect present [$F(1, 79) = 1.09, p = .30$].

No significant differences were found in performance on questions measuring knowledge on the post hoc fallacy, appeal to authority, or the fallacy of the single cause. A significant gender difference was found in performance on questions testing familiarity with the ad hominem fallacy [$F(1, 79) = 7.084, p = .001, \eta_p^2 = 0.139$], with males performing better ($M = 74.85\%, SD = 23.82\%$) than females ($M = 61.67\%, SD = 21.91\%$). No interaction effect or significant differences between conditions were found.

Table 1
Gender Differences in Post-Test Performance

	Overall Performance		Transfer Questions		Ad Hominem Questions	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Males	73.89%	17.60%	73.22%	19.24%	74.85%	23.82%
Females	64.63%	15.50%	60.65%	18.51%	61.67%	21.91%

Disregarding experimental condition, participants performed best on questions testing knowledge of appeal to authority ($M = 87.95\%$, $SD = 24.74\%$) and worst on questions testing knowledge on the fallacy of the single cause ($M = 52.61\%$, $SD = 26.67\%$). Participants also performed better on memorization questions ($M = 73.73\%$, $SD = 20.12\%$) than on questions that required transfer of the lecture material ($M = 65.49\%$, $SD = 19.67\%$).

Feedback Survey

Analysis was performed on the participants' perception of the likeability, effectiveness, and clarity of the video and the speaker based on their responses on the feedback survey. No significant differences were found on participants' perceptions of the speaker's clarity, the effectiveness of the video, the speaker's likeability, or the ease with which they understood the speaker.

Levene's test for equality indicated unequal variances ($F = 4.51$, $p = .04$) in the analysis of the following question: "How much did you like or dislike the video?" As such, it was appropriate to perform a Mann-Whitney U test rather than an ANOVA. No significant differences were found between conditions. As such, gesture did not affect participants' perception of the speaker's or video's clarity, effectiveness, or likeability. However, significant

gender differences were found (Mann–Whitney $U = 627$, $n_1 = 51$, $n_2 = 32$, $p < 0.05$ two-tailed), with males ($M = 3.56$, $SD = .67$) reporting having liked the video more than females ($M = 3.25$, $SD = .56$).

Table 2
Gender Differences in Reports of Having Liked the Experimental Video

	<i>M</i>	<i>SD</i>
Males	3.56	0.67
Females	3.25	0.56

Question was presented as follows: “How much did you like or dislike the video?” Responses are on a scale of 1-5, with 1 labeled as “Disliked the video a lot”, 3 labeled as “Neutral” and 5 labeled as “Liked the video a lot.” See Appendix C for an example of the survey.

NASA-TLX

One-way ANOVAs were performed on the data for each of the NASA-TLX questions.

Assumptions of normality were met unless otherwise stated.

Significant differences for both gender [$F(1, 78) = 12.75$, $p = .001$, $\eta_p^2 = 0.140$] and condition were found [$F(1, 78) = 4.062$, $p = .047$, $\eta_p^2 = 0.049$] for the participants’ reported amount of mental demand experienced while watching the video and completing the post-test. Males perceived less mental demand during the task ($M = 9.75$, $SD = .70$) than females ($M = 12.92$, $SD = .55$). Further, participants in the gesture condition perceived more mental demand ($M = 12.23$, $SD = .63$) than those in the non-gesture condition ($M = 10.44$, $SD = .63$). There was no significant interaction effect.

Table 3
Condition Differences on the NASA-TLX: Mental Demand

	<i>M</i>	<i>SD</i>
Gesture	12.23	0.63
Non-gesture	10.44	0.63

Responses are on a scale of 1-20

Significant gender differences were found in participants' perception of their performance on the post-test [$F(1, 79) = 3.94, p = .05, \eta_p^2 = 0.048$] with males reporting higher perceived rates of success ($M = 7.94, SD = .72$) than females ($M = 9.77, SD = .57$). There was no interaction effect or significant difference between conditions.

Table 4
Gender Differences on the NASA-TLX

	Mental Demand		Perception of Performance		Frustration	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Males	9.75	0.7	7.94	0.72	5.76	0.84
Females	12.92	0.55	9.77	0.57	8.11	0.64

Responses are on a scale of 1-20. For mental demand and frustration, there is a positive relationship between scores and perceived mental demand/frustration. For perception of performance, there is an inverse relationship: lower scores indicate higher rates of perceived success.

Significant gender differences were found in participants' reports of frustration experienced while watching the video and completing the post-test [$F(1, 79) = 4.89, p = .03, \eta_p^2$

= 0.058], with males reporting less frustration ($M = 5.76$, $SD = .84$) than females ($M = 8.11$, $SD = .663$). There were no significant interaction or condition effects.

No significant differences were found on participants' reports of physical demand, temporal demand, or effort required to complete the post-test. As such, gesture only affected participants' perceived mental demand, with gesture seeming to impose more mental demand on learners.

CHAPTER FIVE: DISCUSSION

With no significant differences between conditions in post-test performance or secondary vigilance task performance, the two hypotheses were not supported. Gesture did not seem to affect mental model development as measured by post-test performance. In addition, gesture did not affect cognitive load experienced while watching the lecture, as measured by reaction time in the secondary vigilance task. As such, no evidence was found to support a theory of triple-coding involving a motor processing channel.

Mental Demand

Significant differences were found between conditions on the reported amount of mental demand experienced while watching the lecture video and completing the post-test. Interestingly, participants in the gesture condition reported having experienced more mental demand than those in the non-gesture condition. This finding was unexpected and did not support the hypotheses. Further, this contradicts past findings, which indicated that gesture production during learning tends to alleviate cognitive load (Cook et al., 2012; Goldin-Meadow et al., 2001). According to these findings, it seems that watching gesture does not alleviate cognitive load, and may even increase it – at least in the context of learning abstract concepts.

These findings may have been a result of the added mental demand of the secondary vigilance task while watching the video, during which participants were required to actively pay attention to multiple visual points. During the secondary vigilance task – which is not a component of natural learning settings – participants' attention was split between the speaker in the video and the background color changes. In the gesture condition, participants may have experienced more mental demand because they were paying attention to an additional

component: the speaker's gestures. However, the added mental demand of the secondary vigilance task and the availability of gestural visual cues did not affect post-test performance or RTs in a negative way. As such, it seems that the participants actively watched the speaker's gestures even when the secondary vigilance task was imposing additional mental demand, although they participants did not experience any benefit to their learning as a result of watching the speaker's gestures.

It is important to note that while gestures imposed additional mental demand in the presence of a secondary vigilance task, they might have also worked to offset some of the mental demand that they imposed. That is, gestures may have acted to support learning and offset mental workload by providing additional structure to the verbal material being presented in the video, but because of the additional mental demand imposed by the secondary vigilance task, this effect was not evident in the results. Instead, it is possible that the secondary vigilance task imposed more mental demand than was evident in the results, but the facilitative nature of the gestures acted to offset some of the mental demand. Future studies should examine whether these differences in mental demand persist even in the absence of the secondary vigilance task.

Representative versus Meaningless Gestures

A large majority of the literature on gestures has found that representative gestures facilitate learning and memory. In the current study, I tested whether more abstract and meaningless gestures can facilitate learning in the same way. While Cutica and Bucciarelli (2008) found that meaningless gestures facilitation recall and learning, no evidence was found to suggest that abstract and meaningless gestures facilitated learning or alleviated cognitive load in any way. As such, these findings support those of Woodwall and Folger (1985), who compared representative and emphasizing gestures' effects on learning. They found that representative,

meaningful gestures facilitated absorption and longer term recall of verbal material significantly more than gestures that acted only to emphasize the speech. They concluded that co-speech gesture can facilitate language recall, but particularly more so when the gestures meaningfully represent components of speech rather than simply emphasize them.

The above finding contradicts Noice and Noice's (2001) theory that gestures inherently facilitate and support communication, as well as Cutica and Bucciarelli's (2008) findings that abstract gestures facilitated a more thorough understanding of newly learned content. The current study may have produced different results from those of Cutica and Bucciarelli (2008) for various reasons. As mentioned before, Cutica and Bucciarelli (2008) used two different videos for their gesture and non-gesture condition: one in which the speaker was told to gesture and one in which the speaker was told not to gesture. The videos used in the current study were identical other than a framing difference. As such, Cutica and Bucciarelli (2008) may have found significant differences between their conditions as a result of the differences in their two videos. If this is the case, then these results suggest that the act of gesture-watching was not solely responsible for Cutica and Bucciarelli's (2008) findings as they theorized, and that meaningless gestures do not affect learning. Instead, they may have found significant differences between their conditions because the speaker in their video may have conveyed the lecture material slightly differently - either in intonation, speed, volume, or other speech variables - when told not to gesture than when he was gesturing naturally.

Another possibility is that gesture's effects were tested using a topic and a post-test that were too difficult for the population being studied. This is evident in the participants' low average post-test performance ($M = 68.20\%$; $SD = 16.86\%$). As such, the current study design may have not been sensitive enough to detect any effects that gesture may have had on learning.

To determine whether the lack of significant findings was due to a lack of sensitivity, future studies should test meaningless gesture's effect on learning using a topic that is more appropriate for the population of undergraduate students.

It is important to note that although there was no evidence to suggest that gestures facilitate learning in the context studied here, gestures did not cause any deficit to learning. Even when gestures imposed additional mental demand in the presence of a secondary vigilance task, performance on the post-test did not suffer as a result. As such, while meaningless gestures did not support abstract learning in the current study, they also did not work to impede learning – even when the content and post-test were difficult.

Gender Differences

Results indicated interesting gender differences in multiple areas, with males performing better on transfer questions, questions that tested knowledge of ad hominem, and on the post-test in general. Further, males reported having liked the video more, experienced less mental demand, and felt that they performed better on the post-test. However, males did not benefit more from gestures than females did, as there were no significant interaction effects.

Past research has found that, at least in Western society, males are perceived as being more skilled than females in specific subjects - particularly in math and science (Raty, Kasanen, Kiiskinen, & Nykky, 2004; Steele, 2003). These implicit social biases often translate into teacher behaviors, with teachers in these subjects often offering more praise to male students and seeming to hold higher expectations of male students (Becker, 1981; Jones & Wheatley, 1990; Kelly, 1988). This extends into the home, with parents often unknowingly perpetuating this stereotype in their interactions with their sons and daughters (Frome & Eccles, 1998; Hyde, Fennema, Ryan, Frost, & Hopp, 1990; Tenenbaum, Snow, Roach, & Kurland, 2005). As such,

boys tend to grow up with the subconscious understanding that they should outperform girls in subjects such as math and science. This may also be true for the subject that was tested in the current study: logical fallacies. If so, males may have performed better than females in this study because of a social bias that causes us to expect males to perform better on logic and reasoning tasks than females.

These findings may have also been a result of stereotype threat – a heavily studied phenomenon in which people tend to fulfill expectations based on stereotypes that are conveyed, either explicitly or implicitly, before completing a task. For example, Shih, Pittinsky, and Ambady (1990) found that Asian-American women performed better than a control group on a math test when their Asian ethnicity was primed, but performed worse than a control group when their female identity was primed, thus fulfilling expectations based on ethnic and gender stereotypes. Stereotype threat is mostly studied as it applies to math and science, as gender stereotypes for math and science performance are strong and persistent in our society. However, as mentioned previously, there may also be a social bias to expect males to perform better on logic and reasoning tasks. If such a stereotype exists, the presence of a male speaker in the video lecture may have primed this stereotype, causing the females in this study to experience a stereotype threat that negatively affected their performance and induced higher levels of mental demand. More studies are needed to determine whether a similar social bias exists for the subjects of logic and reasoning as it does for math and science.

Limitations

As mentioned previously, the current study may not have been sensitive to gesture's effects on learning in that the test material and post-test were too difficult for the population

being tested. As such, the lack of significant differences between conditions might not imply that meaningless gesture has no effect on learning.

In addition, the gesture in the experimental video - while natural - may not have been pronounced enough to have an effect on learning or mental demand.

Finally, participants may have been distracted by the secondary vigilance task, causing them to focus visually on certain parts of the screen other than on the speaker's gestures. As such, gesture may not have had an effect on learning in the current study because the participants were not attending to them as they would in natural learning settings. Future studies should use a less visually demanding secondary vigilance task to avoid this possible issue.

Conclusion

In an attempt to determine whether natural gesture in lecture videos may facilitate learning in online courses, I found that meaningless gestures do not seem to support learning or alleviate cognitive load. According to these results and those of past studies in the literature, meaningful and representative gestures facilitate learning to a much higher degree than gestures that only act to emphasize speech. Further, no evidence was found to suggest that natural and meaningless gestures inherently carry any additional communicative information to verbal language as Noice and Noice (2001) theorized – at least when learning abstract concepts. Natural, emphasizing gestures may only serve to alleviate cognitive load in the producer, as previous studies have found (Graham & Heywood, 1975; Rauscher, Krauss & Chen, 1996; Ravizza, 2003). However, we did not find that meaningless gestures acted to impede learning in any way – even when the gestures were meaningless, the nature of the material was abstract, and the content and post-test proved to be difficult for the sample being tested.

Based on the findings of this study, teachers who aim to improve online classes should incorporate representative gestures into their lecture videos whenever possible, but will not find significant improvements in learning through only incorporating meaningless and emphasizing gestures.

APPENDIX A: POST-TEST

Post-test questions & answers in bold:

1. The following appeared as a letter to the editor from a Central Plaza store owner: "Over the past two years, the number of shoppers in Central Plaza has been steadily decreasing while the popularity of skateboarding has increased dramatically. Many Central Plaza store owners believe that the decrease in their business is due to the number of skateboard users in the plaza. There has also been a dramatic increase in the amount of litter and vandalism throughout the plaza. Thus, we recommend that the city prohibit skateboarding in Central Plaza. If skateboarding is prohibited here, we predict that business in Central Plaza will return to its previously high levels." What fallacy or fallacies are being committed here, if any?

A: Ad hominem

B: Fallacy of the single cause

C: Post hoc fallacy

D: Appeal to authority

E: A + B

F: B + C

G: B + D

H: No fallacy is committed

2. The following appeared in a memo from a vice president of Quiot Manufacturing: "During the past year, Quiot Manufacturing had 30 percent more on-the-job accidents than at the nearby Panoply Industries plant, after the work shifts were changed to be one hour shorter than ours. Therefore, to reduce the number of on-the-job accidents at Quiot

and thereby increase productivity, we should shorten each of our three work shifts by one hour so that employees will get adequate amounts of sleep." What fallacy or fallacies are being committed here, if any?

A: Ad hominem

B: Fallacy of the single cause

C: Post hoc fallacy

D: Appeal to authority

E: A + B

F: B + C

G: B + D

H: No fallacy is committed

3. The following report appeared in the newsletter of the West Meria Public Health Council:
- "An innovative treatment has come to our attention that promises to significantly reduce absenteeism in our schools and workplaces. A study reports that in nearby East Meria, where fish consumption is very high, people visit the doctor only once or twice per year for the treatment of colds. Clearly, eating a substantial amount of fish can prevent colds. Since colds represent the most frequently given reason for absences from school and work, we recommend the daily use of Ichthaid—a nutritional supplement derived from fish oil—as a good way to prevent colds and lower absenteeism." What fallacy or fallacies are being committed here, if any?

A: Ad hominem

B: Fallacy of the single cause

C: Post hoc fallacy

D: Appeal to authority

E: A + B

F: B + C

G: B + D

H: No fallacy is committed

4. “Angelina Jolie is well-known for her political activism and efforts to end violence against women and mass genocides. A couple of years ago, Jolie was in the spotlight again for disagreeing with her fiancé’s mother – Mrs. Pitt – on political issues. Mrs. Pitt wrote a letter urging the public to vote for Mitt Romney and to not support same-sex marriage. Jolie, on the other hand, is a strong supporter of same-sex marriage and has been known to vow to not marry until same-sex marriage is nationally accepted. However, Jolie has also been known to make poor choices in her past. For example, she has personally admitted that she had experimented with every known drug by the time she was 20. Further, she has a past with self-harm (cutting). Being an impulsive woman with little self-control, she cannot be considered a reliable source for political information and her arguments in favor of her political beliefs should not be considered valid.” What fallacy or fallacies are being committed here, if any?

A: Ad hominem

B: Fallacy of the single cause

C: Post hoc fallacy

D: Appeal to authority

E: A + B

F: B + C

G: B + D

H: No fallacy is committed

5. “Vaccines have long been a topic of great debate, with many people arguing that they are more harmful than they are helpful. Recently, some people have begun choosing to not vaccinate their children to avoid any potentially harmful side effects. Although scientists have argued that vaccines are safe and failing to vaccinate children can have devastating effects on the public, some have argued that vaccines may not be as safe as they suggest. In fact, many well-known and prominent figures have come forward as having chosen to not have their children vaccinated, such as Jim Carrey, Mayim Bialik (a neuroscientist and famous actress of The Big Bang Theory), Donald Trump, and even Robert F. Kennedy Jr. With the list of prominent figures against vaccinations growing, we should reconsider whether or not we want our children to be vaccinated.” What fallacy or fallacies are being committed here, if any?

A: Ad hominem

B: Fallacy of the single cause

C: Post hoc fallacy

D: Appeal to authority

E: A + B

F: B + C

G: B + D

H: No fallacy is committed

6. List as many one sentence examples of “ad hominem” in use as you can. You will have 90 seconds (do not use an example from the video). - **Removed from post-test**

Free response

7. List as many one sentence examples of an attempt to “appeal to authority” as you can. You will have 90 seconds (do not use an example from the video). - **Removed from post-test**

Free response

8. Which fallacy is closest in concept to this statement: “Correlation does not equal causation”
- A: Ad hominem
 - B: Fallacy of the single cause
 - C: Post hoc fallacy**
 - D: Appeal to authority
9. What does “ad hominem” translate to?
- A: To the man**
 - B: To the master
 - C: To the argument
 - D: None of the above
10. In the video, what did the business man sell in the example for the post hoc fallacy?
- A: Bidets
 - B: Wiggles
 - C: Widgets**
 - D: Bridges

11. From the post hoc fallacy, “post hoc’ is a shortened version of the latin statement “post hoc ergo propter hoc.” Based on what you know about the post hoc fallacy, what does “post hoc ergo propter hoc” translate to?

A: Before this, therefore because of this.

B: After this, therefore because of this.

C: After this, therefore causing this.

D: With this, therefore because of this.

12. “Don’t listen to Madi about politics. She failed her high school math class and had to retake it.” What fallacy is being committed here, if any?

A: Ad hominem

B: Fallacy of the single cause

C: Post hoc fallacy

D: Appeal to authority

E: No fallacy is committed

13. “You should use Plem brand lotion. I heard Gwyneth Paltrow uses it and has had great results.” What fallacy is being committed here, if any?

A: Ad hominem

B: Fallacy of the single cause

C: Post hoc fallacy

D: Appeal to authority

E: No fallacy is committed

14. Which of the following is an example of ad hominem in use?

A: “Gas prices have been steadily rising since Obama took office.”

B: "Every time I text Courtney after 12PM, my phone freezes"

C: "Elise's argument makes no sense. She hasn't even graduated college!"

D: "I've seen a lot of pictures of well-known actors driving Priuses. They must be good cars."

15. Which of the following is an example of an attempt to appeal to authority?

A: "Whenever I park in Parking Lot C, my car needs a jump to start."

B: "Ever since Mr. Perry took office as mayor, my sales have increased!"

C: "Don't believe Stephanie about politics. Her parents are poor."

D: "Don't buy an Android phone. Both Stephen Hawking and Michio Kaku have iPhones. They must be better."

16. The post hoc fallacy involves _____:

A: Mistakenly assuming that when two events occur sequentially, the preceding event caused the succeeding event.

B: Making an assumption about an event's cause based on poorly done research

C: Mistakenly assuming that when two events occur at the same time, they must be related.

D: Drawing false conclusions about an event's cause based on personal prior experience

17. In the example for the post hoc fallacy in the video, what did the business man do in an attempt to increase sales?

A: Paint his business's sign pink

B: Hire more experienced employees

C: Hang his business's sign higher

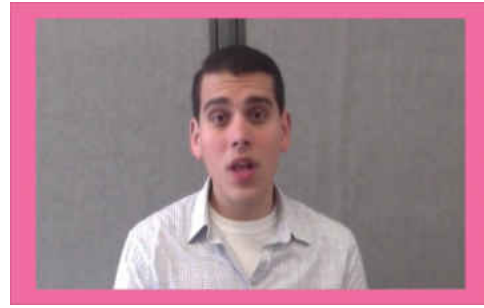
D: Paint his business's sign green

18. Which of the following is an example of the fallacy of the single cause?
- A: "Amanda won't make a good class president. She has only lived in this town for 2 years."
- B: "Gas prices have been steadily rising since Obama took office. This is his fault."**
- C: "Both of the celebrities that I have met refused my request for an autograph. Celebrities are not very nice people."
- D: "Considering James' past with infidelity, I wouldn't trust what he has to say about local politics."
19. Which of the following examples for "ad hominem" was used in the video?
- A: "You can't expect me to trust your opinions at your young age."
- B: "Typical comment for someone so immature."**
- C: "No one will listen to your opinions at your age."
- D: None of the above
20. Which of the following examples for "ad hominem" was used in the video?
- A: "How can I take you seriously when you look like that?"**
- B: "You should work on your appearance if you want to be taken seriously."
- C: "I'm not going to take you seriously until you improve your appearance."
- D: None of the above

APPENDIX B: SCREENSHOTS OF EXPERIMENTAL VIDEOS



Non-gesture video with black background



Non-gesture video with pink background



Gesture video with black background



Gesture video with pink background

APPENDIX C: DEMOGRAPHICS AND FEEDBACK SURVEY

ID Number: _____

Demographics

1. Age: _____
2. Sex: Male Female
3. Year in college: 1st 2nd 3rd 4th 5th Graduate
4. Average college GPA: _____
5. Verbal SAT: _____
6. Quantitative SAT score: _____

Feedback

7. How much did you like or dislike the video?

1	2	3	4	5
Disliked the video a lot		Neutral		Liked the video a lot

9. How effectively did the video convey information?

1	2	3	4	5
Not effective		Not sure		Very effective

10. Please rate the speaker in the video on his clarity

1	2	3	4	5
Not clear at all		Not sure		Very clear

11. Please rate the speaker in the video on his likeability

1	2	3	4	5
---	---	---	---	---

Not at all
likeable

Neutral

Very likeable

12. How difficult or easy was it to understand the speaker in the video?

1

2

3

4

5

Very difficult to
understand

Not sure

Very easy to
understand

APPENDIX D: DIGIT SPAN ANSWER SHEET

Digit Span Answer Sheet

Participant ID: _____

Forward- Practice 1:

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.

Backward- Practice 2:





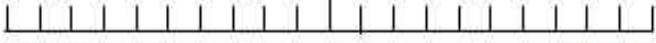
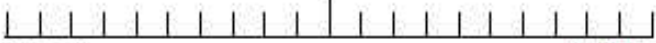
- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.

APPENDIX E: NASA-TLX

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
Mental Demand How mentally demanding was the task?		
		
Very Low Very High		
Physical Demand How physically demanding was the task?		
		
Very Low Very High		
Temporal Demand How hurried or rushed was the pace of the task?		
		
Very Low Very High		
Performance How successful were you in accomplishing what you were asked to do?		
		
Perfect Failure		
Effort How hard did you have to work to accomplish your level of performance?		
		
Very Low Very High		
Frustration How insecure, discouraged, irritated, stressed, and annoyed were you?		
		
Very Low Very High		

APPENDIX F: IRB PERMISSION LETTER



University of Central Florida Institutional Review Board
 Office of Research & Commercialization
 12201 Research Parkway, Suite 501
 Orlando, Florida 32826-3246
 Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1
 FWA00000351, IRB00001138**

To: **Maura M.K. Austin and Co-PIs: Karen W. Verkler, Michele G. Gill, Valerie K Sims**

Date: **January 28, 2015**

Dear Researcher:

On 1/28/2015, the IRB approved the following human participant research until 01/27/2016 inclusive:

Type of Review: UCF Initial Review Submission Form
 Project Title: Gestures and mental models: A triple-coding hypothesis
 Investigator: Maura M.K. Austin
 IRB Number: SBE-15-10939
 Funding Agency:
 Grant Title:
 Research ID: N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form cannot be used to extend the approval period of a study. All forms may be completed and submitted online at <https://iris.research.ucf.edu>.

If continuing review approval is not granted before the expiration date of 01/27/2016, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the [Investigator Manual](#).

On behalf of Sophia Dziegielewska, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 01/28/2015 12:02:33 PM EST

IRB manager

REFERENCES

- Allen, L.Q. (1995). The effects of emblematic gestures on the development and access of mental representations of French expressions. *Modern Language Journal*, 79(4), 521–529.
- Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press.
- Baddeley, A. D. (1999). *Human memory*. Boston: Allyn & Bacon.
- Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. *Journal of Educational Psychology*, 76, 408-417.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 577–609.
- Becker, J.R. (1981). Differential treatment of females and males in mathematics classes. *Journal for Research in Mathematics Education*, 12, 40–53.
- Brucker, B., Ehlis, A. C., Häußinger, F. B., Fallgatter, A. J. & Gerjets, P. (2014). Watching corresponding gestures facilitates learning with animations by activation human mirror-neurons: An fMRI study. *Learning and Instruction*, 36, 27-37.
- Clark, H. H. (1996). *Using language*. New York: Cambridge University Press.
- Cook, S.W., Mitchell, Z. & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, 106, 1047–1058.
- Cook, S.W., Yip, T. & Goldin-Meadow, S. (2010). Gesturing makes memories last. *Journal of Memory and Language*, 63(4), 465–475.
- Cook, S.W., Yip, T. & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language And Cognitive Processes*, 27(4), 594-610. doi:10.1080/01690965.2011.567074

- Cutica, I. & Bucciarelli, M. (2008). The deep versus the shallow: Effects of co-speech gestures in learning from discourse. *Cognitive Science*, 32, 921-935
- DeLeeuw, K. E. & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, 100(1), 223-234.
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91(1), 176–180
- Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology*, 74(2), 435-452.
- Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593-609.
- Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Science*, 3(11), 419-429.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining Math: Gesturing Lightens the Load. *Psychological Science*, 12(6), 516-23
- Graham, J. A., & Heywood, S. (1975). The effects of elimination of hand gestures and of verbal codability on speech performance. *European Journal of Social Psychology*, 5(2), 189-195.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances In Psychology*, 52, 139-183.
doi:10.1016/S0166-4115(08)62386-9
- Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, 137(2), 297-315. doi:10.1037/a0022128

- Hyde, J. S., Fennema, E., Ryan, M., Frost, L.A., and Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly*, 14(3), 299-324.
- Jones, M. G., & Wheatley, J. (1990). Gender differences in teacher– student interactions in science classrooms. *Journal of Research in Science Teaching*, 27, 861–874.
- Kelly, A. (1988). Gender differences in teacher-pupil interactions: A meta-analytic review. *Research in Education*, 39, 1–23.
- Mainieri, A. G., Heim, S., Straube, B., Binkofski, F., & Kircher, T. (2013). *NeuroImage*, 81, 294-305.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. *The Cambridge Handbook of Multimedia Learning*, 31-48.
- Mayer, R. E. (1989). Models for understanding. *Review of Educational Research*, 59, 43-64
- Mayer, R. E. & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, 83(4), 484-490.
- Mayer, R. E. & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, 82, 715-726.
- McNeill, D. (1992). *Hand and Mind*, University of Chicago Press
- Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies. *US Department Of Education*.

- Molenberghs, P., Cunnington, R. & Mattingley, J. B. (2011). Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neuroscience and Biobehavioral Reviews*, 36, 341-349.
- Montgomery, K. J., Isenber, N. & Haxby, J. V. (2007). Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Social Cognitive and Affective Neuroscience*, 2(2), 114-122.
- Noice, H. & Noice, T. (2001). Learning dialogue with and without movement. *Memory and Cognition*, 29(6), 820-827.
- Noyes, J. M. & Bruneau, D. P. J. (2007). A self-analysis of the NASA-TLX workload measure. *Ergonomics*, 50(4), 514-519.
- Novack, M. A., Congdon, E. L., Hemani-Lopez, N. & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychological Science*, 25(4), 903-910.
doi:10.1177/0956797613518351
- Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.
- Piaget, J. (1983). Piaget's theory. In P. Mussen (ed). *Handbook of Child Psychology*. 4th edition. Vol. 1. New York: Wiley.
- Raty, H., Kasanen, K., Kiiskinen, J., & Nykky, M. (2004). Learning intelligence—children's choices of the best pupils in the mother tongue and mathematics. *Social Behavior and Personality*, 32, 303–312.
- Rauscher, F. B., Krauss, R. M., & Chen, Y. (1996). Gesture, speech and lexical access: The role of lexical movements in speech production. *Psychological Science*, 7, 226-231.

- Ravizza, S. (2003). Movement and lexical access: Do noniconic gestures aid in retrieval? *Psychonomic bulletin & review*, *10*(3), 610-615.
- Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.
- Schippers, M. B. & Keysers, C. (2011). Mapping the flow of information within the putative mirror neuron system during gesture observation. *NeuroImage*, *57*, 37-44.
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*(1), 80-83.
- Steele, J. (2003). Children’s gender stereotypes about math: The role of stereotype stratification. *Journal of Applied Social Psychology*, *33*, 2587–2606.
- Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.
- Sweller, J. (2003). Evolution of human cognitive architecture. In B. H. Ross (Ed.) , *The psychology of learning and motivation: Advances in research and theory*, Vol. 43 (pp. 215-266). New York, NY, US: Elsevier Science.
- Tenenbaum, H.R., Snow, C.E., Roach, K.A., & Kurland, B. (2005). Talking and reading science: Longitudinal data on sex differences in mother-child conversations in low-income families. *Applied Developmental Psychology*, *26*, 1–19.
- Wagner, S. M., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of Memory and Language*, *50*, 395-407.
doi:10.1016/j.jml.2004.01.002
- Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*, 159 – 177.

Woodwall, 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(4), 319-333.